AUGMENTED MIRRORS: DEPTH JUDGMENTS WHEN AUGMENTING VIDEO DISPLAYS TO REPLACE AUTOMOTIVE MIRRORS

Missie Smith[1], Valerie Kane[1], Joseph L. Gabbard[1], Gary Burnett[2], David R. Large[2]
Virginia Tech[1], University of Nottingham[2]

This study investigates the effects of Augmented Reality (AR) graphics on a drivers' distance estimation and depth perception when using a video-based, AR-enhanced driver's side mirror. Sixteen participants took part in the study, eight in a driving simulator and eight outside in a stationary vehicle. Participants experienced three different AR display image conditions, three different glance patterns, three different target vehicle speeds, and two own-vehicle image conditions. Distance data and confidence data were collected for each participant and analyzed for any correlation between the conditions and performance. The results suggest that various AR images affected depth judgements and confidence levels. In addition, the vehicle speed and glance pattern of the videos also had significant effects.

INTRODUCTION & BACKGROUND

As much as 90% of the information that automobile drivers use is gathered visually (Hills, 1980). Mirrors provide one source of this visual information and are a necessary component for the safety of drivers. Driving on a motorway is one instance where mirror use is extremely important, as drivers must use their mirrors to determine when to merge into lanes or what space constraints exist when overtaking a slower car. These time and distance judgments affect the flow of traffic and, when made poorly, introduce the potential for danger. As such, drivers must be able to quickly assess whether there is sufficient space and time to merge into a lane ahead of another vehicle. Traditional side view mirrors also create drag, therefore making the vehicle less aerodynamic as compared to a car without exterior side view mirrors. A car without exterior mirrors, or a *mirrorless car*, could eliminate this excess drag and increase fuel economy by using interior digital displays to convey driver rear and side visual information.

Large et al. found that driver performance was not affected by using video displays in this manner, although drivers preferred configurations that mimic standard mirror locations (Large et al., 2016). Displays placed in unfamiliar locations (e.g., in the center console) were associated with lower levels of trust, situational awareness, and depth perception. In addition, a dynamic study comparing rearview mirrors to video displays found evidence of more conservative driving judgments using the video display (Flannagan, 2005).

Some depth and judgment studies, as does the study presented herein, have used cars as the target object in the mirror (Flannagan et al., 1997; Flannagan et al., 1996). Participants generally underestimated the distance between target and test vehicles, which goes against the common warning label on mirrors stating: "Objects in mirror are closer than they appear". This proposes the idea that the viewing system is not the cause of inaccurate depth estimation but rather the limited field of view and exclusion of the entire ground plane.

Auto manufacturers are already producing a car that has digital LCD displays on the dashboard in lieu of exterior optical mirrors (Mohamed Ali & Bazilah, 2014). This approach allows for future improvements in terms of image enhancement (NHTSA, 2009). With research into side mirrors potentially favoring digital video solutions, it follows that overlaid registered augmented reality (AR) cues could further differentiate digital mirrors from their traditional glass-based counterparts. When implemented into car mirrors, AR may provide benefits seen in other vehicle AR displays, such as increased situational awareness (Kim et al., 2013) and may further enhance target detection, target speed and safe passing times. While AR has recently received attention in the driving domain (often implemented in head-up or windshield-based displays or in center console rear backup assist systems), it appears that little work has examined how best to augment side view mirrors.

The work presented herein aims to make an initial assessment of whether adding AR images to video side mirrors may assist drivers in making these decisions. Since there are not currently recommendations as to what types of AR graphics would be most beneficial to drivers, examining different AR distance cues was also of interest.

METHODS

Participants

Five females and eleven males participated in this study (ages 21-55 years). Each person had self-reported 20/20 or corrected-to-perfect (where perfect is defined as 20/20) vision and at least one year of motorway driving experience in a right hand drive vehicle.

Materials

Two test vehicles were used for this study: a stationary sedan (located in an outdoor parking lot) or a fixed-base driving simulator. The set of locations was due in part to equipment availability and more importantly to better understand the validity of using a driving simulator for AR HUD perceptual research. It may be possible that *some* perceptual problems can be studied in a simulator without the logistical overhead of outdoor studies. A white sedan was the target vehicle in the study. A 7-inch Lilliput video display was placed in the interior of the car on the driver's side dashboard, to mimic current mirror locations. All other mirrors on the test vehicles were blacked out so that all visual cues related to the approaching target vehicle were seen via video display. However, participants could see out of all windows. Microsoft PowerPoint with Visual Basic for Applications custom code recorded participant response data. Experimental visual stimuli were created using video overlaid with AR images via Apple Motion 5. Participants were given a computer mouse to use as the trigger in the study and when pressed, the identifying video information and time stamp was recorded.

Video. The video displayed to the participants was a single prerecorded video depicting a target vehicle approaching the stationary test vehicle as if to pass. The target vehicle drove 417 feet before passing next to the test vehicle. The target vehicle drove at a constant speed of 20 mph for the entirety of the original video. Because of the difficulty of capturing a video without pedestrians or other vehicles, the same source video was used for all conditions with modifications to the source video. There were no pedestrians or other vehicles in the video.

First, to simulate the three target speeds, the video frame rate was slowed proportionally and was always greater than 30 Hz, ensuring smooth playback. Using the same video source further afforded consistency in known and unknown visual elements across all visual conditions.

Figure 1: Own-vehicle Reflection Present (left) and Not Present

Because the presence of the driver's own vehicle reflection could possibly provide depth cues, a second set of videos were created from the single source video that included the participant's own vehicle reflection (Figure 1). Next, three AR display images were overlaid onto the newly generated videos.

AR Graphics. Figure 2 shows the four AR display conditions: bars, timer, box, and none (control). The *bars* on the ground appeared as conformal distance markers, similar to backup assist camera graphics. The dynamic yellow *box* was registered to the approaching target vehicle (e.g., the box increased in size as the target vehicle moved closer). The *timer* was a screen fixed cue placed in the bottom corner; it did not directly augment a real-world objet per se, but instead provided time as a cue of time until target vehicle passes the test vehicle. The last condition was the control video, a video with no additional AR cues.

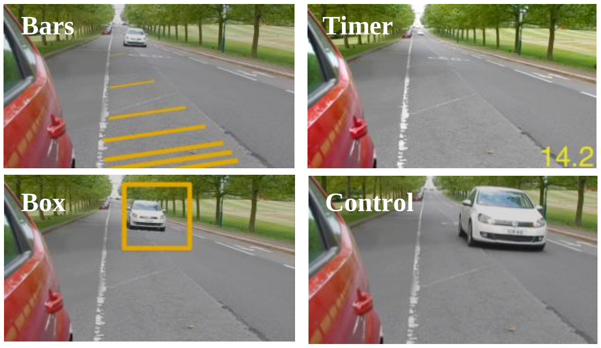

Figure 2: Four Display Image Conditions

Glance Patterns. Since drivers cannot continuously watch side mirrors while driving, the videos were manipulated to systematically control when drivers could "glance" at the virtual mirrors. Previous research indicated that drivers tend to glance at driver side mirrors for durations of 0.94-1.36 seconds (Sodhi et al., 2002). Therefore, a 1 second glance was used to allow drivers to make their distance assessments. Since the same source video was used for all trials, the glance patterns of participants were altered to prevent learning and anticipation effects. Specifically, the videos were divided into three equal segments: beginning (B), middle (M), and end (E), with each trial having a 1 second glance during two of the three segments. The glance order was counterbalanced, and every participant saw every possible glance order (BM, BE, ME) for every speed, display image, and reflection condition.

A total of 24 videos were created from the original source: 3 speeds (10, 18, and 20 mph) x 4 display images (3 AR, 1 control) x 2 own-vehicle reflection conditions (present and absent) x 3 glance patterns (BM, BE, ME), for a total of 72 trial conditions per participant. The experiment was divided into 8 blocks (4 display image conditions x 2 reflection conditions), which were counterbalanced across all participants. Within each block, participants saw nine videos (3 speeds x 3 glance patterns).

Procedure

After sitting in the driver's seat of the vehicle, participants buckled their seat belt, adjusted their seat so they were comfortable and put their hands on the steering wheel to simulate a real driving scenario. In the outside vehicle, participants were told to look straight ahead at a lamppost that was approximately 20 yards ahead. Looking at the lamppost accounted for the accommodation switching required when drivers look from the road to their mirrors. In the driving simulator, participants looked ahead at a lead vehicle depicted on a screen positioned approximately 3 yards ahead, although the perspective of the image implied that the car was further ahead.

Depth and Timing Judgment of Passing Vehicle. Participants were told that they would receive two glances prior to making the depth/timing judgment and they were instructed to look at either the lamppost (outside) or the lead vehicle (driving simulator) between all glances. Each trial began with a black screen on the video display while drivers focused on the lamppost/lead vehicle. As prescribed by the pre-determined glance pattern, a bell rang to cue participants to switch focus (i.e., glance) to the video display. Bell sounds occurred one second prior to glance videos to ensure participants saw entire glances of passing vehicles. After the second glance, participants signaled when they thought that the target object was directly next to their car by pressing a trigger (Figure 3). When pressed, the video was stopped and the time stamp was recorded. After the time stamp was recorded, participants stated their judgment confidence on a scale of 0-10 (0 indicated no confidence, while 10 indicated perfect performance).

RESULTS

This driving task was a combination of distance and temporal judgments, which are linearly related and intrinsically intertwined. Therefore, all error was measured in units of time. Figure 3 shows that a participant who waited too long to press the button (e.g. the target car passed prior to their trigger) over-estimated the time until the target vehicle passed. Conversely, pressing the trigger too early was an under estimation and resulted in a negative signed error. Signed error indicates directionality and quantifies participants' propensity to over or under estimate. The absolute value of the signed error (absolute error) provided the magnitude of error, with performance improving as absolute error approached 0. Both signed and absolute errors were used to show the differences in judgment accuracy and consistency. Subjective measures included confidence scores after each trial and post-hoc preference questionnaires.

All outliers beyond 1.5 times the inter-quartile range were removed; however, the data were not normally distributed. Therefore, unless otherwise stated below, Wilcoxon was used for all pairwise analyses and the Kruskal-Wallis test for groups of three or more. Steel-Dwass post-hoc tests were used when initial analysis indicated significant differences.

Display Image

Signed Error. There was a significant difference in signed error of the four display images (H(3)=13.641, *p*=0.0034). The post-hoc test showed two significant differences; Box and Bar conditions differed (*p*=0.0236) as did the Control and Box conditions (*p*=0.0073); see Figure 4.

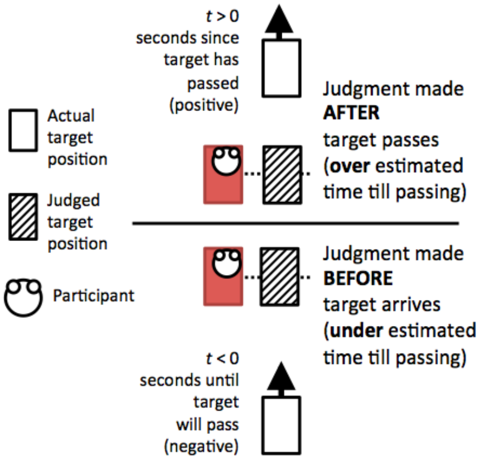

Figure 3: Conceptual representation of participant judgments (hatched boxes): both underestimating and overestimating arrival times of approaching target vehicle (white)

Absolute Error. Initial tests indicate differences in absolute error with regards to display image (H(3)=29.3730, *p*<0.0001). The subsequent test found differences between Timer and all other display image conditions: Timer-Box (*p*=0.0022), Timer-Control (*p*=0.0004), and Timer-Bar (*p*<0.0001); see Figure 4.

Confidence. There was a significant difference in reported confidence between the four displays (H(3)=96.6957, *p*<0.0001). The post-hoc analysis found differences between the timer and all other display images (*p*<0.0001 for all pairs); see Figure 4.

Passing Vehicle Speed

Signed Error. The passing speed of the vehicle correlated to different signed error judgments (H(2)=99.940, *p*<0.0001). A post-hoc analysis indicated that the 10 mph speed was significantly different from both the 18 mph (*p*<0.0001) and 20 mph (*p*<0.0001) conditions (Figure 5).

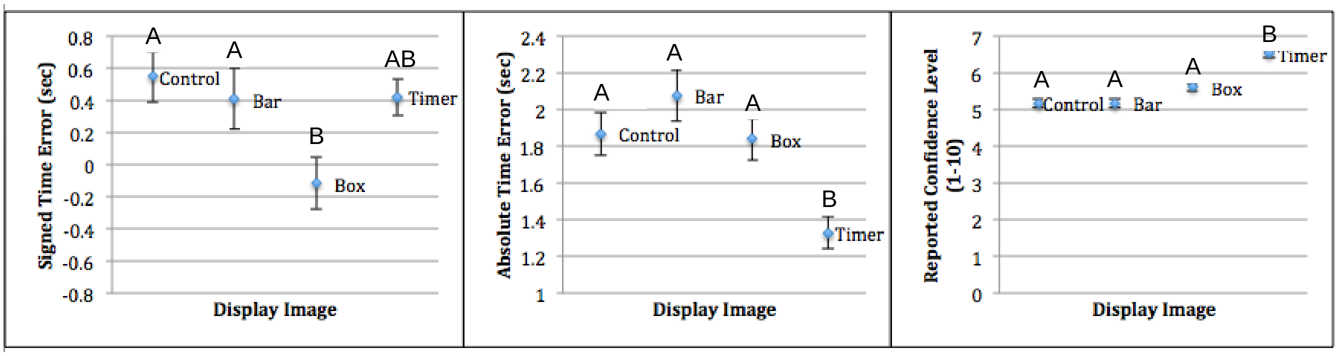

Figure 4: Signed Time Error (a), Absolute Time Error (b), and Reported Confidence Level (c) by Display Image

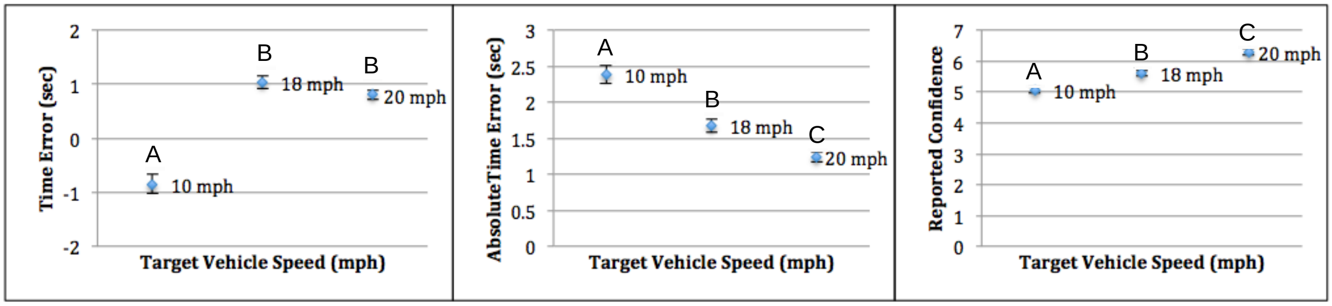

Figure 5: Signed Time Error (a), Absolute Time Error (b), and Reported Confidence (c) by Vehicle Speed

Absolute Error. There were significant differences in participants' absolute error across the three target vehicle speeds ($H(2)=55.184$, $p<0.0001$). The post-hoc analysis indicated differences between all pairs (20-18: $p=0.009$, 18-10: $p<0.0001$, 20-10: $p<0.0001$); see Figure 5.

Confidence. The initial test indicated a significant difference between reported confidence levels and target vehicle speed ($p<0.0001$). Further analysis found differences in reported confidence for all pairs (20-10: $p<0.0001$, 20-18: $p<0.0001$, and 18-10: $p=0.0006$); see Figure 5.

Glance Pattern

Signed Error. Significant differences in signed error exist between the three glance conditions ($H(2)=9.4885$, $p=0.0087$). Therefore, subsequent testing determined that the BM condition is different from both the BE ($p=0.0336$) and the ME conditions ($p=0.0140$); see Figure 6.

Absolute Error. There were significant differences between glance conditions when examining absolute error ($p<0.0001$). Further testing showed differences between the BM-BE condition ($p<0.0001$) and the ME-BM condition ($p<0.0001$); see Figure 6.

Confidence. There were also significant differences in reported confidence based on the glance pattern ($p<0.0001$). Subsequent testing indicated that ME-BM ($p<0.0001$) and BM-BE ($p<0.0001$) glance patterns were different (Fig. 6).

Glance and Speed Interaction. Figure 7 suggests there is likely an interaction between glance pattern and speed. Specifically note the difference in drivers' performance during the 10 mph condition across the BM curve relative to the other conditions on both absolute and signed error.

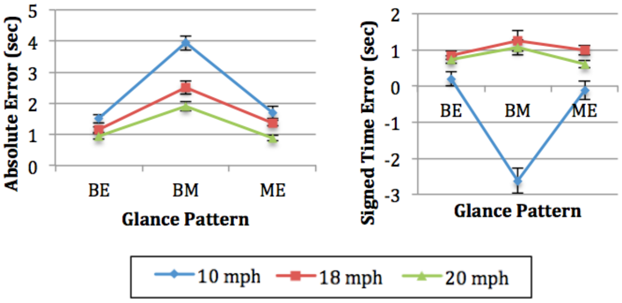

Figure 7: Absolute and Signed Time Error interactions with Glance Patterns

Own-Car Reflection

Signed Error. The presence of the car reflection in the video had no significant impact on the participant signed time error ($p=0.147$, $Z=1.452$).

Absolute Error. There was no significant difference in the absolute error based on the presence or lack of presence of the car in the video display ($p=0.109$, $Z = 1.601$).

Confidence. The reported confidence level was not significantly different regardless of the presence of the vehicle reflection in the display image ($p=0.536$, $Z=-0.619$).

Study Location

Signed Error. There was no significant difference in signed error when comparing data from the simulator and data from the outside car ($p=0.173$, $Z=-1.361$).

Absolute Error. The location of the study did not significantly effect on the absolute error ($p=0.225$, $Z=-1.214$).

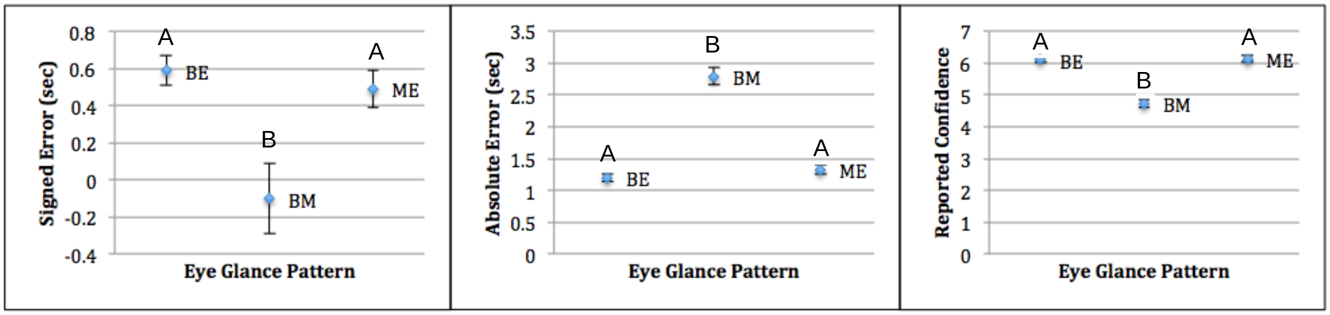

Figure 6: Signed Time Error (a), Absolute Time Error (b), and Reported Confidence (c) by Eye Glance Pattern

Confidence. The location of the study did not have a significant effect on the reported confidence (p=0.445, Z=0.763).

DISCUSSION

The timer was the most preferred condition by 50% of the participants, yet it was identified as the most distracting condition. Participants said that although it helped them accurately determine when the target car was to pass, they focused only on the timer. Thus, such an AR design might increase distraction and raise safety concerns. The location of the timer (bottom right of display) may further draw attention away from other information seen in the mirror display. The nature of the graphic (temporal only) might also encourage participants to ignore other in-mirror cues regarding distance and depth judgments. Therefore, while the timer condition was associated with lower absolute error and higher confidence, it may have negative effects in driving situations requiring high levels of situational awareness.

Target vehicle speed had a noticeable effect on driver performance, especially when paired with the glance patterns. While a relative passing speed of 10 mph is valid, empirical observations suggest that the combination of a beginning-middle (BM) glance pattern with a slow passing speed resulted in an unnatural experience for participants. The glance and speed interaction suggests that future studies should randomize glance patterns, and present glance opportunities when the target object is relatively closer to the participant (vehicle). Furthermore, perhaps the starting vehicle distance should be shorter when speeds are slower to keep participants engaged throughout each trial. Participants consistently pressed the trigger prematurely in the 10 mph-BM condition and remarked about the lack of glances towards the end of the trial. Conversely, the beginning-end and middle-end glance patterns were associated with better performance, suggesting that the end glance is the most important glance in making depth and timing judgments. Such glances provide the most recent and timely location information for the participant prior to completing the task. Overall, the forced two-glance structure used herein may not provide a natural experience relatable to driving, and should be refined for future studies.

The absence or presence of the own-car reflection did not seem to affect participants' judgments. The presence of the reflection most accurately represents current vehicle mirror displays. Conversely, the lack of reflection mimics what might be seen in a video captured from the rear of the vehicle.

Participant location (in simulator vs. outside) did not affect the outcome of the study. Participants gave feedback on both locations saying that the external noises outside were distracting. Participants inside the simulator said that having no noise was off setting. Therefore, perhaps noise cues play a larger role in vehicle passing judgments than previously anticipated. At a minimum, it was something that was frequently noted by participants.

CONCLUSION & FUTURE WORK

Overall, the study suggested that a purely temporal AR display could improve performance some time and distance judgment tasks. However, the task did not require driving or anticipating unexpected events, which may result in negative consequences in the timer condition. The interaction between glance pattern and speed had strong effects on performance, and should be carefully considered when undertaking any future research of this type. The presence of an own-vehicle reflection and physical location (simulator or outside) both had little effect on performance and perceptions. Therefore, future studies such as this can exclude vehicle reflection as a variable and can be performed either in a naturalistic environment or a simulator.

REFERENCES

Flannagan, M. J. (2005). Distance perception with a camera-based rear vision system in actual driving.

Flannagan, M. J., Sivak, M., Schumann, J., Kojima, S., & Traube, E. (1997). *Distance perception in driver-side and passenger-side convex rearview mirrors: Objects in mirror are more complicated than they appear*. Retrieved from

Flannagan, M. J., Sivak, M., & Traube, E. C. (1996). *Driver perceptual adaptation to nonplanar rearview mirrors* (0148-7191). Retrieved from

Hills, B. L. (1980). Vision, visibility, and perception in driving. *Perception, 9*(2), 183-216.

Kim, H., Wu, X., Gabbard, J. L., & Polys, N. F. (2013). *Exploring head-up augmented reality interfaces for crash warning systems.* Paper presented at the Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.

Large, D. R., Crundall, E., Burnett, G., Harvey, C., & Konstantopoulos, P. (2016). Driving without wings: The effect of different digital mirror locations on the visual behaviour, performance and opinions of drivers. *Applied Ergonomics, 55*, 138-148.

Mohamed Ali, J., & Bazilah, F. F. (2014). Mirrorless Car: A Feasibility Study. *Applied Mechanics & Materials*(663).

NHTSA. (2009). Traffic safety facts: 2007 data: pedestrians. *Annals of Emergency Medicine, 53*(6), 824.

Sodhi, M., Reimer, B., & Llamazares, I. (2002). Glance analysis of driver eye movements to evaluate distraction. *Behavior Research Methods, Instruments, & Computers, 34*(4), 529-538.