Augmented Reality Interface Design Approaches for Goal-directed and Stimulus-driven Driving Tasks

Coleman Merenda*, Hyungil Kim, Kyle Tanous, Joseph L. Gabbard,
Blake Feichtl, Teruhisa Misu, and Chihiro Suga

Abstract— The automotive industry is rapidly developing new in-vehicle technologies that can provide drivers with information to aid awareness and promote quicker response times. Particularly, vehicles with augmented reality (AR) graphics delivered via head-up displays (HUDs) are nearing mainstream commercial feasibility and will be widely implemented over the next decade. Though AR graphics have been shown to provide tangible benefits to drivers in scenarios like forward collision warnings and navigation, they also create many new perceptual and sensory issues for drivers. For some time now, designers have focused on increasing the realism and quality of virtual graphics delivered via HUDs, and recently have begun testing more advanced 3D HUD systems that deliver volumetric spatial information to drivers. However, the realization of volumetric graphics adds further complexity to the design and delivery of AR cues, and moreover, parameters in this new design space must be clearly and operationally defined and explored. In this work, we present two user studies that examine how driver performance and visual attention are affected when using fixed and animated AR HUD interface design approaches in driving scenarios that require top-down and bottom-up cognitive processing. Results demonstrate that animated design approaches can produce some driving gains (e.g., in goal-directed navigation tasks) but often come at the cost of response time and distance. Our discussion yields AR HUD design recommendations and challenges some of the existing assumptions of world-fixed conformal graphic approaches to design.

Index Terms— Mixed-reality, augmented reality, driving, head-up displays

1 INTRODUCTION

Augmented reality's (AR) presence in the automotive market has grown considerably in the last decade. Despite the fact that many questions and challenges remain for applications of AR-based information systems, vehicles that support AR information delivery via head-up display (HUD) technology have now firmly entered the market and are being sold on a large scale consumer basis. The past 5 years alone have seen a continuous growth of vehicles equipped with HUDs, and the market is expected to continue to multiply by a factor of seven through 2020 to an estimated 9.1 million units [1]. A HUD system is generally defined as any transparent display that projects visual information into drivers' forward-looking visual field [2], but most HUDs today function by using optical reflection to project information into drivers' field of view via glass or mirror combiner [3]. While the vast majority of current HUD-equipped automobiles provide only basic information to drivers (e.g. vehicle speed and time), several vehicle setups with more complex functionality have recently emerged. For example, Volvo and BMW now market AR HUDs that deliver more contextual information such as screen-fixed navigation aids like instructions, lane departure warnings, parking space proximity, and collision warnings [4, 5]. Additionally, third-party sellers such as Hudway and Garmin now offer HUD technology packages that host AR graphics via mobile-based applications [6], meaning that even cars without manufacturer-provided HUDs may still be adapted into AR integrated systems. With this clear growth of the AR sector in automotive manufacturing, AR HUD designers have become more pressed to address the critical issues surrounding the development of best interface design practices and their relation to driver attention and safety.

In this spirit, the work described herein contributes to the body of augmented reality and transportation knowledge by providing: (1) a characterization of AR HUD interface elements, (2) delineation of goal-directed and stimulus-driven task scenarios to be supported by AR HUD graphics, (3) two outdoor, on-road user studies that examine multiple AR HUD designs within different driving scenarios, (4) a novel means to assess the effects of designs on human performance that differentiates visual capture from visual guidance, and, (5) AR HUD interface design recommendations for future consideration.

2 RELATED WORK

2.1 Benefits and Challenges of AR Systems in Driving

The potential that AR systems have to improve driving performance and attention while on the road when successfully implemented, is well known. Recently, AR-based displays have been hailed as a better means to deliver information to drivers while still maintaining lower levels of mental workload [7]. When comparing general driving performance via HUD to a traditional head-down display (HDD) (i.e. a visual display mounted in the center of a vehicle's dash), clear advantages emerge. For example, use of a HUD over an HDD allowed drivers to keep their gaze on the road longer, leading to faster reaction times, improved task performance, and less mental stress [8]. AR HUD displays are also able to better mitigate dangers resulting from drivers diverting their view from the road when compared to traditional head-down displays [9]. Active in-vehicle information systems have been shown to help reduce driver task load by presenting real-time information regarding routes, delays, congestion, and warnings of potential hazards [10, 11]. Taken in sum, these findings bolster the promise of spatially co-locating AR graphics with the forward roadway to mitigate the known dangers of divided attention, afford faster visual processing of information at hand [12], and empower drivers to devote more time and mental workload to the primary task of driving.

The described gains in attention can easily translate into significant performance improvements when applied to specific driving scenarios. For example, when navigating, directions presented via AR HUDs can reduce the ambiguity of instructions by directly pointing

- *Coleman Merenda is with Virginia Tech.E-mail: cjm120@vt.edu*
- *Hyungil Kim is with Virginia Tech. E-mail: hci.kim@vt.edu*
- *Kyle Tanous is with Virginia Tech. E-mail: ktanous@vt.edu*
- *Joseph L. Gabbard is with Virginia Tech. E-mail: jgabbard@vt.edu*
- *Blake Feichtl is with Virginia Tech. E-mail: blake96@vt.edu*
- *Chihiro Suga is with Honda Research Institute, USA. E-mail: CSuga@hra.com*
- *Teruhisa Misu is with Honda Research Institute, USA. E-mail: tmisu@honda-ri.com*

Manuscript received 19 Mar. 2018; accepted 10 July. 2018.
Date of publication 14 Sept. 2018; date of current version 28 Sept. 2018.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TVCG.2018.2868531

out intersections and turn locations in the context of what drivers actually see [13]. In a study on cognitive mapping, drivers using pathway style cues projected via HUD in conjunction with traditional top-view navigation (i.e. GPS) experienced reduced navigational errors and distraction measures as compared to using the top-view alone [14]. In more urgent situations such as collision scenarios, drivers have shown preference to augmented views of the road over conventional map views [15]. However, multiple sources of in-vehicle information can also add to drivers' cognitive load by increasing divided attention in cases where task complexity becomes too great to easily handle. Though past work has argued that AR-enabled cues can further guide driver's attention to localize threats earlier [12], others have argued that AR cues do not actually improve the detection of cues related to a maneuver, but rather only change allocation of visual attention by making fixations more numerous and less task-specific [16]. Several other key issues encountered when using AR systems continue to hinder their effective use. For example, individuals operating a HUD-equipped vehicle tend to suffer from both 1) cognitive/attention capture (i.e. unconsciously shifting attention from surrounding world towards the HUD visualization) and 2) perceptual tunneling (i.e. excessively narrowing their peripheral awareness) [17]. AR graphics are also frequently affected by the interaction between real-world backgrounds and natural lighting, and, in outdoor situations where both conditions are dynamically changing, can detrimentally impact users' discriminability of graphics and perception of cue color and meaning [18, 19].

2.2 Technical Challenges of HUDs

The degree to which AR visuals are able to be effectively used in an automotive system is significantly impacted by the means in which they are presented. The level of discretion AR HUD interface designers are given to deliver information to drivers is to a large degree still constrained by the limits of HUD technology. In the past, research has pointed to design and usability limitations imposed by low screen resolution, low luminance levels, and a limited field of view (FOV) [20]. Though the FOV available on HUDs is expected to continue growing, current commercially-available HUDs provide relatively small FOV (6-15 degrees) as compared to head-worn AR displays and thus can display only limited amounts of information [21]. Such displays risk being unable to capture driver attention (in the event the cue is too small) or risk demanding too much visual attention in the event that it is difficult to see or discern important detail in the graphic without prolonged glances [20].

In terms of interface design, there is contention regarding the most effective area to place AR cues in respect to the driver's forward view. On one hand, if virtual images are placed directly in the center of the driver's view, they are likely to occlude imminent information of the real world or reduce a driver's recognition of danger due to an excessive presence of virtual information [21]. In fact, symbology that is placed within a 5-degree radius of the fovea tends to be considered as annoying by drivers and generally ineffective [22]. On the other hand, creating fully immersive AR HUDs is not a solution in itself, as a larger virtual design canvas will likely draw too much visual attention and distract drivers from the forward roadway [21]. To present dynamic information that appropriately informs the situation, HUDs should generally be placed either directly in front of the driver behind the steering wheel or to the right of the steering wheel in the tertiary sector, and no more than 30 degrees below a driver's view [21]. In sum, the limited FOV of current HUDs coupled with HUD placement options limit the practical applicability of conformal AR HUD graphics – that is, a UI designer can only annotate real-world objects that fall within an AR HUD's limited FOV, which arguably are a subset of the most important real-world objects and hazards drivers encounter. Thus, it is likely that conformal approaches to AR HUD design are not inherently optimal. Given these many potential pitfalls, the placement of AR HUD graphics can quickly become a difficult and complex proposition as designers must balance an ability to capture and direct drivers' attention when needed while still minimizing distraction from the driving scene.

2.3 AR HUD Presentation Methods

Drivers using AR systems are influenced not only by the system's technical design and limitations thereof but also by the complexity of the graphics delivered. Currently, many HUD interface designs provide simple fixed-position, two-dimensional graphics (e.g., current speed). These designs thus are limited in their ability to provide depth or spatial cues to drivers concerning their environment [23]. Fixed-focal plane displays with relatively short focal-depths (e.g., 3m) increase the cost of attentional switching from HUD to primary task, which can result in drivers missing more external hazards, delaying responses to external events, or suffering from longer transition times to return attention from the HUD to the road scene [24].

We can define the effort required to shift to and from the physical space (i.e. the real world) to the information space (i.e. virtual objects overlaid in the world) and back as "cognitive distance". This distance increases as AR graphics are perceived as less integrated with the environment and causes driving to become, in effect, a dual task [14]. The cognitive distance can also increase when AR graphics are projected at a different focal depth than their real-world referents, forcing drivers to accommodate (and shift attention) back and forth between two distances. If continued at length, this mismatch can lead to eye strain and diminishing ability to correctly perceive and perform in-vehicle tasks [25]. More recent work has shown that the inclusion of screen-fixed alerts alone does not necessarily provide significant reaction gains in gaze behavior towards a road hazard, yet still produces improvements in concurrent braking performance [26]. This suggests that AR cues that do not provide strong localization may encourage fast, but "blind" reactions to threats without actually guiding drivers' attention effectively.

There is a significant need to explore how driver behavior changes as HUD graphic presentation moves from traditional static 2D graphic types to a true conformal, or "world-oriented" approach. Gabbard et al. [20] helped conceptualize the design space for automotive AR interfaces from the users' point of view: *Screen-fixed* AR graphics are rendered at a fixed location on the HUD and are generally not spatially (or perceptually) "attached" to any specific objects in the scene, while *world-fixed (or conformal)* AR graphics are rendered such that they are perceived to exist in specific locations in the real world using real-time geolocation, pose-estimation, and 3D rendering software. In simulation, conformal graphics have been shown to direct attention to real-world referents that fall within the HUD FOV. For example, a study that highlighted pedestrians with gaze-contingent cues helped reduce the number of collisions when tested in a simulator [27]. However, truly conformal presentation is often difficult to research in on-road studies as it requires a technical vehicle-based implementation that supports real-time, geo-referenced 3D rendering of graphics outdoors at speed. Recent work has shown promising results when comparing 3D AR graphics to traditional 2D HUD views, including task performance gains [13] as well as faster braking reaction time and smoother braking profiles [28, 29]. However, few studies yet exist that effectively explore the performance of conformal graphics in real-world driving environments.

3 Objectives

Given the limitations of current HUD technology, it is important to further explore the extent to which AR interfaces may be leveraged without requiring significant or unfeasible demands from manufactures. For example, can conformal AR graphics provide enough driver performance gains to merit the investment in developing more advanced AR HUDs with, for example, larger FOVs and multiple focal planes? In examining these questions, we propose that available AR HUD interface design approaches not be limited to traditional 2D (i.e. screen fixed) or conformal (i.e. world-fixed) cues alone. Figure 1 presents a proposed classification for AR HUD interface design approaches that distinguish between the behavior of an AR HUD interface element and the nature of the space in which that element exists. For simplicity, we present the classification as four discrete cells, but one can imagine AR HUD designs that fall between these representative endpoints. The behavior dimension is anchored on one end by *fixed interface elements* that are statically positioned at a specific location, and on the other by *animated interface elements* that rotate and/or translate through a volume. The orthogonal dimension specifies the frame of reference, or location of the "designer's canvas" – that is, the space in which the AR HUD designer can position fixed or animated interface elements. On the *screen-centric* end of the continuum, the designer's canvas is attached to the HUD screen, moving through the world along with the driver as if the canvas is physically attached (at a distance) to the front of the vehicle. On the *world-centric* end of the continuum, the designer's canvas is attached to the environment (or the world). This set of AR HUD interface design approaches contains what is commonly termed conformal graphics.

Note that this classification of interface design approaches is, by design, independent of the AR HUD hardware used to display a given AR interface design. While often presumed true, it is an open question whether AR graphics for driving, presented at fixed depth or animated through a volume, are better-supported by volumetric head-up displays as compared to single, fixed focal plane HUDs. Moreover, the interface design approaches represented in the classification are not mutually exclusive, and indeed an AR HUD application could very well include a meaningful mashup of interface elements from each design approach. To this end, the proposed classification may aid future AR HUD researchers and practitioners in organizing their interface design work.

One of the main objectives of this work is to explore the impact of the previously described AR HUD interface design approaches on drivers' performance and behavior when used in representative driving scenarios. To this end, we first suggest that a large number of representative driving activities fall into one of two categories. (1) In *goal-directed tasks*, drivers are aware and actively seeking a goal-state and managing their attention to the means-end of the task. Examples of these tasks include navigating from point a to point b, finding a parking space, and parallel parking. (2) In *stimulus-driven tasks*, drivers are unaware of an incoming event and must react to stimuli accordingly by managing attention by adapting to a change in situation. Examples of stimulus-driven tasks include responding to vehicles in blind spots, swerving around animals entering the roadway, and braking for pedestrians stepping out into the roadway. These two scenario categories can be drawn from top-down vs. bottom-up human information processing, which are also regarded as the two major paths to human comprehension [30]. We believe that the cognitive process (and subsequent performance) of drivers may be significantly impacted by which mode of processing they are engaging in during a task. Thus, this work aims to examine differences in the effectiveness of different AR interface design approaches in goal-directed (user study 1) and stimulus-driven (user study 2) tasks. Through these studies, we aim to answer the following questions:

1. Do drivers experience driving performance gains using conformal (as compared to traditional HUD) AR interface designs?
2. In addition, do drivers experience visual attention gains when using conformal (as compared to traditional HUD) AR interface designs?
3. Given that world-fixed, conformal approaches may not always be feasible, practical or desirable, to what extent can other design approaches that use A) screen-based frames of reference, and B) dynamic conformal cues, leverage similar gains in driving performance and visual attention?

4 Study 1: Goal Directed Navigation

In the first user study, we were interested in comparing animated AR interface designs to traditional screen-fixed as well as world-fixed conformal designs within the context of a goal-directed task. We chose navigation as a representative goal-directed task and defined an experimental task that required participants use AR cues to locate a pre-designated target parking space.

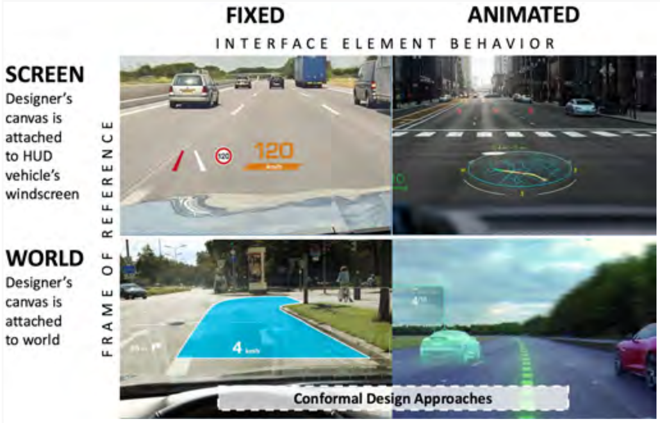

Fig. 1. We consider four classes of AR HUD interface design approaches ranging from traditional 2D HUD graphics (e.g., speed and symbology, top-left) to highly dynamic, animated conformal graphics (e.g., virtual car metaphor for navigation, bottom-right).

4.1 Participants

Twenty-four drivers (17 male, 7 female) took part in the user study, with an age range of 18-40 years (mean = 27.35 yr.). Qualified participants were adults with perfect or perfect-corrected vision and at least two years of driving experience.

4.2 Experimental Design

We conducted a two-factor within-subjects repeated measure experiment, where participants completed tasks using four types of AR *graphics* at three different *distances* from the navigation target. Figure 2 depicts the four AR graphics conditions inspired by the four interface design approaches in Figure 1.

Screen-fixed (screen-fixed): displayed a static blue turn-arrow symbol indicating direction of target parking space and a dynamic counter of remaining distance to target space.

Conformal (world-fixed): was a blue turn-arrow "painted" on the road and perceptually anchored adjacent to the target parking space.

Compass (screen-animated): featured a flat circular area virtually attached to the participant's car approximately eight meters in the forward-looking field. The circular area contained a blue navigation arrow that oriented itself dynamically in accordance with the changing position of the participant's car relative to the target parking space.

Pathway (world-animated): employed motion-based spatial ¬cues vis-à-vis an animated blue arrow that moved across the ground and pointed to the target parking space.

We examined driver performance and behavior at three *distances* (9m, 13.5m, and 18m) termed near, medium and far respectively, and defined as the distance between the onset of AR graphic and the target parking space location. We varied distance to better understand how different interface design approaches help drivers localize targets at different distances, and whether there is an interaction effect between design approach and distance to target. We used a 4x3 nested counterbalanced design to mitigate order effects across four AR graphics by three distance conditions.

4.3 Apparatus & Testbed

During the user study, participants drove a 2009 Honda Odyssey test vehicle equipped with a volumetric head-up display provided by Honda Research Institute. It is a projection-based volumetric display with a swept-volume technique [28] using fast switching image planes within a range of focal distance between *8m* and infinity (0.125D ~ 0D); affording flicker-free appearances of virtual objects in about 17° circular field of view.

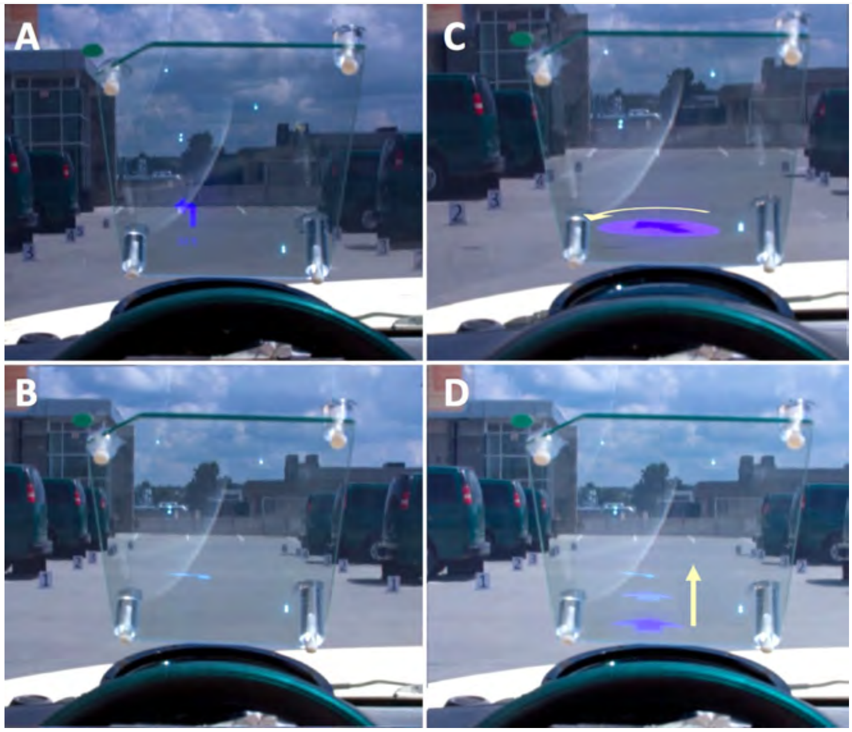

Fig. 2. Navigation user study AR graphics: (A) screen-fixed, (B) conformal, (C) compass, and (D) pathway. Since the compass and pathway interfaces changed appearance and/or location over time, figures 2c and 2d are approximate of what participants actually experienced (with yellow arrows denoting animation direction).

We authored testbed software in C++ and QT5 to render the AR graphics. The testbed software was designed to allow changes to the testbed parameters and AR interfaces in real-time, allowing iterative design and improvements to be made while conducting pilot testing directly in the outdoor experimental driving environment. GPS (uBlox EVK-M8N) was used to activate AR interfaces based on the test vehicle's geolocation and used to track and record vehicle position via GPSD application. Additionally, an SMI ETG (80°*60° tracking range) recorded participants' gaze behavior at 30Hz.

All experimental trials were conducted in an outdoor parking lot testbed. The parking lot was a cordoned off area of approximately 100 meters by 110 meters, or an area with enough width to allow participants to experience at least 10 seconds of straight driving at 10mph. Participants' tasks took place within a straight, preplanned event zone with known geolocations for target parking spaces. The event zone contained six cargo vans of identical color, parked in three of nine parking spaces on each of the left and right sides. The remaining six empty spaces on each side were numbered using 1x1 ft. upright signs placed at the opening of each space and facing participant's approach (see Figure 2).

4.4 Procedure

Trials were completed in ample daylight to afford safe vision by ensuring that light levels fell within useful daylight luminance (approximately 1000-2000 lux) [31]. Each participant first engaged in a practice drive with sample AR HUD graphics and tasks to become accustomed to the test vehicle dynamics and graphics. Participants drove around the parking lot in a pattern similar to that of actual testing. During this time, the researcher instructed the participant to drive as close as possible to 10 mph at all times unless they felt a situation warranted the correct action to brake, in which case they should freely brake in a manner as they would when normally driving. A researcher was present in the vehicle at all times, and constantly monitored participant's speed. If participants' speed drifted beyond 10±2 mph for more than 5 seconds, the researcher would verbally remind participants to hold a 10 mph speed. Participants were allowed to practice driving until they felt comfortable driving and viewing sample AR graphics. Once comfortable with the procedure, a researcher equipped participants with eye-tracking glasses and performed a calibration procedure to ensure accurate eye gaze tracking.

For the experimental task, participants drove in a looping pattern around the parking lot. When entering the event zone, a navigation graphic would appear that participants used to localize the target parking space (which were randomly assigned per trial). Once participants felt confident of their choice, they would verbalize the parking space number aloud. Participants were instructed to localize the spot as quickly as possible without losing accuracy.

After each trial, participants completed a short five question survey in which they evaluated their perceived effectiveness of the AR graphics shown in the specific trial. The survey (taken from [32]) was designed to explore several key stages of the human cognition and was organized using a seven-point Likert scale ranging from strongly agree to strongly disagree. Participants rated their agreement with the following statements: (1) The visuals caught my attention quickly; (2) The visuals helped me detect objects in the real world relevant to the task at hand; (3) The visuals helped me predict the movement and the future action of real-world objects; (4) The visuals help me react appropriately to the urgency of the situation; and; (5) I felt confident using the system.

Following the conclusion of trials, participants were given the opportunity to provide detailed and contextual verbal feedback on any of the AR graphics they experienced. For brevity, we do not extensively report the qualitative data of verbal feedback in this paper, and instead, focus on reporting feedback via subjective ratings.

4.5 Dependent Measures

To support an integrative view of the effect of AR interface designs on drivers, we collected measures for both task performance as well as driver eye behavior. We evaluated task performance via:

Verbal response time: difference in time (in seconds) between the onset of AR graphic and the moment participants provided their verbal response denoting which spot they believed the graphic was directing them towards.

Verbal response accuracy: defined as the severity of judgement error, defined in the number of parking spaces from the correct target parking space.

In assessing the effects of AR HUD designs on human performance, we believe it is critical to also examine driver eye behavior at both the "detection" phase of an event (i.e., when the driver must perceive an alert or direction from an AR interface element or target, respectively) and the "decision" phase of an event (i.e. when the driver must direct attention towards executing a response). Thus, to explore the dynamics of visual attention, we established the following measures related to drivers' eye behavior:

Time to fixation on AR graphic (FG): the time elapsed between the onset of an AR graphic and fixation on the AR graphic.

Time to fixate from AR graphic to task-related target (GT): time difference between participants' fixation on AR graphic and target parking space.

Time to fixation on task-related target (FT): time elapsed between onset of AR graphic and fixation on target parking space. FT can be represented as FG+GT.

4.6 Analysis & Results

All collected measures were analyzed using either JMP or R code software packages. Inferential statistical comparisons were performed for each collected measure across all four AR graphic conditions and three distance conditions using two-way repeated measures Analysis of Variance (ANOVA) with post hoc contrast tests for planned comparisons among experimental conditions with Tukey's adjustment for multiple comparisons. In figures graphing statistical results, significance between experimental conditions are reported by *'***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$, and* '•' $p < 0.1$. Results shown in all figures are additionally annotated using either a *square* symbol to denote a *screen-centric* graphic approach or a *circle* to denote a *world-centric* graphic approach. Similarly, results shown in all figures use either *solid lines* to denote *fixed (or static)* graphic approaches, or *dashed lines* to denote *animated* graphic approaches. Error bars indicate standard errors of the means. In cases of missing data, Satterthwaite approximations were used to compensate for lost degrees of freedom [33]. Unless otherwise stated, all data met normality assumptions initially or after log transformation.

4.6.1 Task Performance

Repeated measures ANOVA tests found main effects of both interface ($F(3,51) = 4.755$, $p = 0.005$) and distance to target parking space ($F(2,32) = 8.793$, $p < 0.001$) on *verbal response time*. No interaction effects were found. Post-hoc contrast tests revealed that drivers' verbal response times were longer in the compass condition as compared to conformal ($t(49)=3.264$, $p = 0.010$) and pathway ($t(50)=3.270$, $p=0.010$). Examining interactions, we saw longer verbal response times for compass as compared to conformal ($p = 0.008$) and pathway ($p = 0.008$) at near distances and as compared to conformal graphics at medium distances ($p = 0.026$), with no significant difference between AR graphic types at the far distance (Figure 3).

We defined *verbal response accuracy* as the signed localization error as measured in number of parking spaces. Positive localization error indicates an overestimation of distance while negative localization denotes distance underestimation. ANOVA tests (Figure 4) revealed significant main effects for both AR graphic ($F(3, 51) = 10.246$, $p<0.001$) and distance ($F(2, 35) = 41.861$, $p < 0.001$), as well as interaction effects ($F(6, 99) = 5.340$, $p < 0.001$). Post hoc tests show that screen-fixed graphics were overestimated more than conformal ($t(52)=5.378$, $p < 0.001$) and pathway ($t(53)=4.149$, $p<0.001$) graphics. Also, compass graphics were consistently overestimated as compared to conformal graphics ($t(50)=3.390$, $p = 0.007$). 95% confidence interval (CI) testing shows that conformal graphics were underestimated most severely at medium distance (-1.05 spaces [-1.65, -0.46]) and far distances (-1.5 spaces [-2.09, -0.91]). All graphic types trended towards an underestimation at the far distance, a finding consistent with other AR-based depth judgement research (e.g., [16]).

4.6.2 Driver Eye Behavior

We analyzed participants' eye fixation patterns using SMI BeGaze 3.5 software suite. We defined the AR graphic and the target parking space as significant areas of interest and coded fixations in these areas as such (all other fixations were labeled as "other"). After coding fixations in areas of interest, fixation data was further annotated with specific experimental condition and relevant dependent measure. Annotations were exported with timestamps and integrated with data logs from the testbed software using NumPy and Matplotlib.

Gaze data was excluded for two participants due to poor eye-tracking quality, and for another two participants after the integration process revealed the incorrectly logged data. In addition, a small subset of trials (approximately 12-24 instances, or once or twice per participant) were removed in cases where participants did not actually fixate on either the AR graphic or the target parking space.

ANOVA tests found a significant main effect for time to fixation on graphic (FG) by AR graphic ($F(3,6) = 3.218$, $p=0.024$) but not for distance. Post-hoc analysis revealed that conformal graphics were fixated on more slowly than were screen-fixed graphics ($t(53)=3.022$, $p=0.019$). Results are shown in Figure 5a.

For fixation from graphic to target (GT), we found significant main effects of graphic type ($F(3,98)=5.4956$, $p=0.001$) and distance ($F(2,61) = 5.139$, $p=0.007$). Post-hoc analysis revealed that

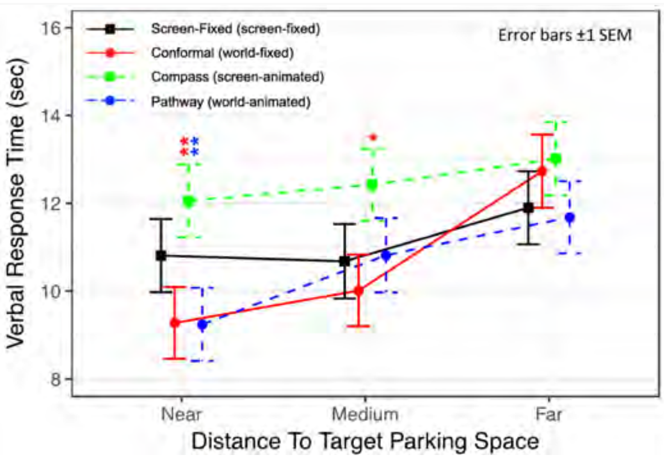

Fig. 3. Overall, participants needed more time to verbalize a target parking space using the compass as compared to conformal and pathway graphics, especially at near and medium distances.

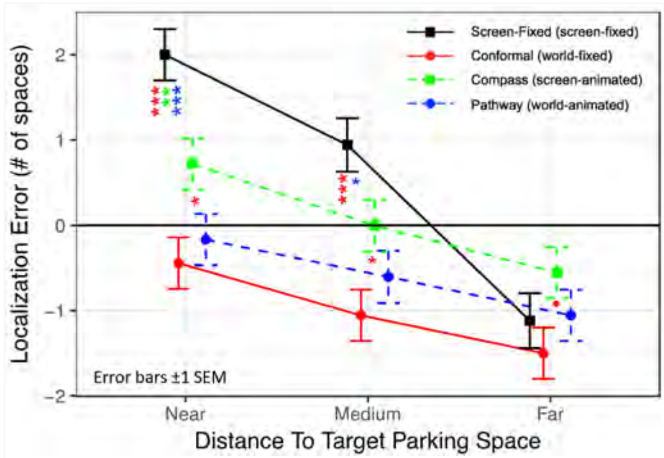

Fig. 4. Participants overestimated the position of the target parking space using screen-fixed graphics as compared to using compass and pathway graphics. Participants worst underestimations occurred at medium and far distances using conformal graphics.

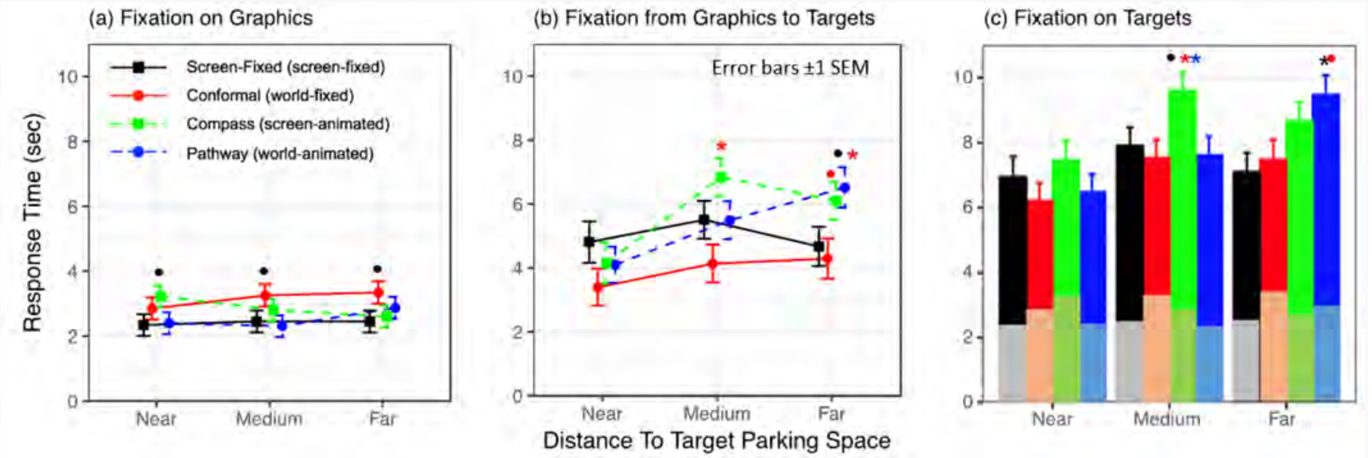

Fig. 5. (A) Participants' took longer to fixate on conformal cues as compared to screen-fixed. **(B)** Once seen (and fixated upon), participants were able to more quickly fixate on the target parking space using conformal AR cues (perhaps suggesting that conformal cues may guide attention toward a target more effectively than other graphic types). **(C)** When examining the total task time participants generally took longer locating a target parking space when using compass graphics, and compared to using screen-fixed and conformal graphics. Note that in (c), the lightly shaded portion of bar graphs represents FG, the fully saturated portions GT, and the overall height FT.

participants were able to more quickly shift their gaze from conformal graphics to target parking spaces as compared to switching from the compass graphics ($t(176)=3.904$, $p<.001$) and pathway graphics ($t(177)=2.805$, $p=.028$) to target parking spaces (Figure 5b).

Analysis of time to fixation on target (FT) revealed main effects for both AR graphic ($F(3, 102)=5.580$, $p=0.001$), and distance ($F(2, 64)=5.254$, $p=0.006$). Post-hoc tests show that it took longer for participants to eventually fixate on parking space targets using compass graphics as compared to both screen-fixed ($t(179)=2.759$, p-0.032) and conformal ($t(177)=3.453$, $p=0.003$) graphics. Analysis of interactions shows that at far distances, participants needed more time with the pathway graphic to fixate on the target as compared to using screen-fixed ($p=0.013$) and conformal ($p=0.058$) graphics (Figure 5c).

4.6.3 Subjective Feedback

We analyzed participants' subjective ratings for the awareness survey completed after each trial. As the data was not normal, we used Friedman tests to examine ratings for significant differences but found no main effects for AR graphic.

4.7 Discussion

Recall that an objective of this work was to better understand whether conformal interface approaches can improve driver performance (i.e., finding target parking space) and visual attention (as measured by graphics ability to capture and direct attention), as compared to traditional vehicle-fixed 2D interface designs. Further, we wanted to understand if adding animation to both vehicle-centric and world-centric design approaches would result in additional performance gains. Based on our findings, world-fixed (conformal) and screen-fixed (traditional HUD) design approaches were associated with similar response times needed to localize a target parking space and provide a verbal response to the perceived correct target space. However, when designating the target parking space, participants notably overestimated distances when using the screen-fixed designs as compared to conformal designs (especially at near and medium distances). At far distances, we saw no differences in accuracy between the two.

The compass screen-animated design approach allowed drivers to make more accurate judgements than the world-fixed, conformal design at all distances. However, these gains in accuracy judgments came at the cost of time (by which it was associated with the slowest performance of all other design approaches). These findings support the idea that AR graphic designs that provide dynamic but relative directional cues (i.e., relative between drivers' own-vehicle and target positions) may require drivers to "close the distance" to the target before the directional cue is effective. That is, at a distance, drivers may not be able to discern which of a cluster of empty parking spaces a virtual compass is pointing to, but as drivers approach the group of parking spaces, the diagnosticity of the compass increases. Therefore, this presentation style may not be effective in goal-directed tasks that require high granularity at distance (e.g., locating a parking space or townhouse address), but instead may be effective for goal-directed tasks with lower granularity such as street turning or highway ramp exiting that require much less precision for successful completion.

Albeit under mixed distance conditions, both animated design approaches were generally associated with less overestimation and underestimation of distance as compared to world- or screen-fixed graphics. We expect that the extra feedback afforded through motion cues helped drivers make more accurate judgements. These results suggest that there are AR design opportunities beyond what is typically examined in the literature (i.e., screen-fixed and world-fixed conformal designs) that can leverage AR graphics' motion to provide drivers with useful feedback on their position in space relative to a real-world target of importance.

In examining the effect of AR design approaches on driver eye behavior, we propose that while ultimately the total time between graphic onset and target fixation is important (FT, total time needed to locate a target), researchers and practitioners can gain insight on candidate designs by teasing apart how much of the total time is needed to detect the AR graphic (e.g., using FG as a measure of *attention capture*) and how much time is needed to locate the target once the AR graphic is detected (e.g., using GT as a measure of how well an AR graphic can *direct attention*). Indeed, examining Figure 5 closely reveals that screen-fixed and world-fixed conformal cues have no significant difference in total time to locate a target, but what is very different is *how* these two graphics support the task. Specifically, the world-fixed cue is quick to capture attention but slow to guide, whereas the world-fixed conformal cue is slow to capture but quick to guide.

Our results suggest that motion-based interface design approaches can be more effective than fixed-approaches at capturing attention, but in fact are slower to guide drivers' attention to real-world targets. World-fixed conformal cues may not be effective at capturing attention since they naturally become less distinct at increasing distance due to perspective and apparent size depth cues. Moreover, conformal world-fixed navigation graphics that adopt a strict real-world metaphor (e.g., painting a turn arrow on the road), can further suffer from low saliency at distance since its height in the visual field at distance can create a worst-case viewing perspective. Conversely, screen-fixed cues of adequate size and saliency will likely capture

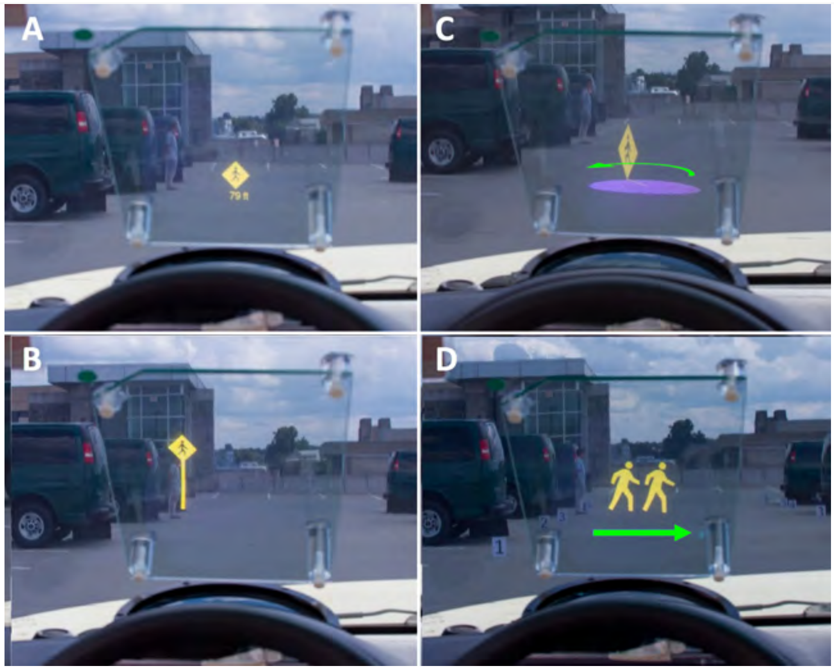

Fig. 6. In study 2, AR graphics aimed to guide drivers' attention to an actual pedestrian stepping out from behind a parked van. We tested four AR graphics types: A) screen-fixed, B) conformal, C) compass, and, D) pathway. Green arrow annotations were not visible and are shown here to merely denote animation direction.

drivers' attention independent of distance to their real-world referent due in part to the fact that the graphics are often oriented perpendicular to drivers' line of sign and have a stable visual footprint independent of distance to target.

The fact that the world-fixed conformal graphics were able to quickly guide attention to targets very likely occurred because the position of the world-fixed conformal graphic consistently matched the spatial location of the target for the duration of the task. In contrast, symbolic screen-fixed cues are arguably the most spatially disconnected from targets, requiring further cognitive processing to map the information conveyed into the forward-looking scene. For different reasons, the motion-based AR interfaces were also slower to guide attention to real-world targets than the conformal design. From observation, we noted that these design required time and/or distance to elapse before the cue could effectively guide drivers' attention. Specifically, for the compass graphic, drivers had to spend time approaching (driving towards) the target, and in the case of the pathway, drivers had to wait for an animated world-fixed graphic to traverse towards an accurate but eventual target.

Thus, our work suggests that while the spatial properties of world-fixed conformal graphics may effectively direct attention to real-world targets, more work is needed to better understand how to design them such that they are easier to detect at distance (e.g., by increasing saliency and altering the viewing perspective). Similarly, work is needed to understand how best to design motion-based AR cues for driving. In the interim, we suggest that care be taken to ensure that the motion characteristics that afford attention capture are not overly salient and accidentally *demand* too much of driver attention. And similarly, that the time and distance requirements for an animated graphic to be effective be taken into consideration to ensure that the driving task in which the AR is to support can adequately support those timelines.

5 Study 2: Stimulus-Driven Pedestrian Collision

In the second user study, we were also interested in comparing animated AR interface design to traditional HUD and conformal designs but within the context of a stimulus-directed task. We defined a pedestrian collision event scenario (while driving through a parking lot) as a representative stimulus-driven task, whereby participants used AR cues to help identify a pedestrian stepping out from between two parked vehicles.

5.1 Participants

The same twenty-four drivers (17 male, 7 female) from study 1 also took part in study 2, with an age range of 18-40 years (mean = 27.35 yr.). Qualified participants were adults with perfect or perfect-corrected vision and at least two years of driving experience.

5.2 Experimental Design

We conducted a two-factor within-subjects repeated measure experiment, where participants completed tasks using four types of AR *graphics* (Figure 6) corresponding to the four interface design approaches in Figure 1.

Screen-fixed (screen-fixed): displayed a static yellow pedestrian symbol indicating location of the real pedestrian (left or right) and a dynamic counter of remaining distance to the pedestrian.

Conformal (world-fixed): presented a pedestrian crossing signpost perceptually anchored adjacent to the target pedestrian.

Compass (screen-animated): featured a flat circular area virtually attached to the participant's car approximately eight meters in the forward-looking field. The circular area contained a yellow pedestrian sign that oriented itself dynamically in accordance with the changing position of the participant's car relative to the target pedestrian.

Pathway (world-animated): was realized as an animated virtual pedestrian symbol that simulated the actual pedestrian's predicted path into the street.

To better understand how AR interface designs may interact with the time- and distance-pressured demands of driving and ultimately driver performance and judgement, participants performed tasks at three different distances represent differing levels of *urgency*; high, medium, and low. Urgency levels were based on time to collision (TTC) measures established in the literature (e.g., [26, 28, 34]) and included: high urgency (~2s TTC), medium urgency (~3s TTC), and low urgency (~4s TTC). Resulting TTCs translated to approximately 9m, 13.5m, and 18m from a target pedestrian given an instructed driving speed of 10 mph.

5.3 Experimental Design

Study 2 used the same test vehicle, volumetric HUD, testbed software, GPS hardware/software, and eye-tracking equipment as Study 1. For Study 2, however, the testbed software also communicated with the test vehicle's CANBUS to log velocity and ultimately derive deceleration profiles. The pedestrian study was performed in the same outdoor parking lot, with the same arrangement of six cargo vans. The numbered signs used in the navigation study were removed, and black cloth was hung at the rear of each cargo van to prevent participants from seeing the pedestrian's feet before stepping out (see Figure 6).

5.4 Procedure

Trials were completed under monitored daylight conditions between approximately 1000-2000 lux [31]. As in Study 1, participants experienced a practice drive with sample AR HUD graphics, and were instructed to drive 10 mph (with similar contingencies in place when participants would deviate from 10 mph). Once participants were comfortable with the procedure, a researcher equipped participants with eye-tracking glasses and performed a calibration procedure.

For the experimental task, participants were instructed to drive as they normally would, and respond to any "sudden events" using aids provided via HUD graphics. Periodically a participant could experience a pedestrian event, in which a pedestrian actor would walk out from behind a parked cargo can and face the road in a manner that indicated they were about to cross into the path of the vehicle (the pedestrian actor did not actually walk into the intersecting path of the participant's test vehicle). As pedestrian actors stepped out from behind a parked van, the HUD displayed a pedestrian alert AR graphic and the participant would need to brake in order to avoid a theoretical collision. To simulate a believable scenario, the pedestrian actor's behavior was timed carefully to match the graphic appearance using pre-defined locations of pedestrians and a trigger point for pedestrians to begin movement using a "Wizard of Oz" technique similar to those developed in past work [28, 35]. Both the researcher in the car and the two researchers acting as pedestrians were in communication with each other at all times via standard localized short-range radios. When the test vehicle passed the pre-defined trigger point, a GPS data feed would trigger the onset of the visual warning, while the pedestrian actor was signaled to step out via radio (with an approximate 0.5 second lead time for actor signals to compensate for actor reaction time based on practice and pilot testing). We used trigger mapping with a 0.5m tolerance to ensure that triggers would still occur given a GPS sampling rate of approximately 20Hz.

After each trial, participants completed a short five question survey on perceived effectiveness of the AR graphics, followed by the opportunity to provide qualitative feedback on the AR graphics.

5.5 Dependent Measures

For the pedestrian study, we focused on braking performance as a means to evaluate participant's task performance. Specifically, braking performance was evaluated using the following metrics:

Time to stop: the difference in time (measured in seconds) between the instant the pedestrian actor began stepping out into the road and an AR graphic simultaneously appeared, and the moment that participants' test vehicle reached zero velocity and came to rest.

Vehicle stop gap: calculated by measuring the remaining longitudinal distance (in meters) between the location of participants' stopped test vehicle and pedestrian actor.

Peak deceleration during braking: derived from the acceleration profile logged during braking (in meters/second2), and is useful to account for the risk of rear-end collision (e.g., by a closely following "tailgating" vehicle) that could theoretically occur when drivers urgently brake.

As with the navigation study, we also evaluated driver gaze behavior using measures of:

Time to fixation on AR graphic (FG): the elapsed time between the onset of an AR graphic and fixation on the AR graphic.

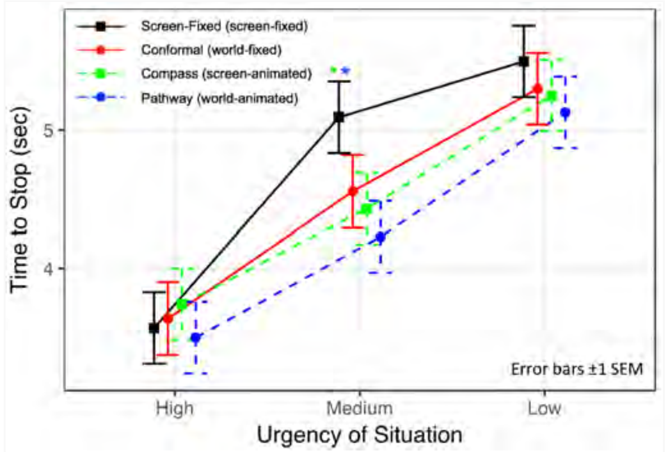

Fig. 7. Time to stop during pedestrian scenarios remained similar for most AR graphics (except for screen-fixed in a medium urgent situation), but decreased in tandem with increasing urgency.

Time to fixate from AR graphic to task-related target (GT): time difference between participants' fixation on AR graphic and pedestrian actor.

Time to fixation on task-related target (FT): time elapsed between onset of AR graphic and fixation on the pedestrian actor.

5.6 Analysis and Results

5.6.1 Task Performance

For braking response time, ANOVA testing found a significant main effect for urgency ($F(2,34)=28.818$, $p<0.001$) but no main effect for AR graphic type or interaction effects were found. Post hos analysis revealed that time to stop was significantly faster for pathway graphics ($p=0.013$) as compared to screen-fixed AR graphics, but only at medium urgency. Differences in stop time are shown below in Figure 7.

ANOVA tests on vehicle stop gap found a significant main effect for both graphic type ($F(3,151)=334.06$, $p<0.001$) and distance ($F(2,34)=58.020$, $p<0.001$), but no interaction effects. Generally, stop gaps decreased with urgency and were similar across all graphic types except for the compass, which showed significantly shorter stop gaps as compared to other AR graphics (all pairwise comparisons $p<0.001$). Of particular note is that in high urgency

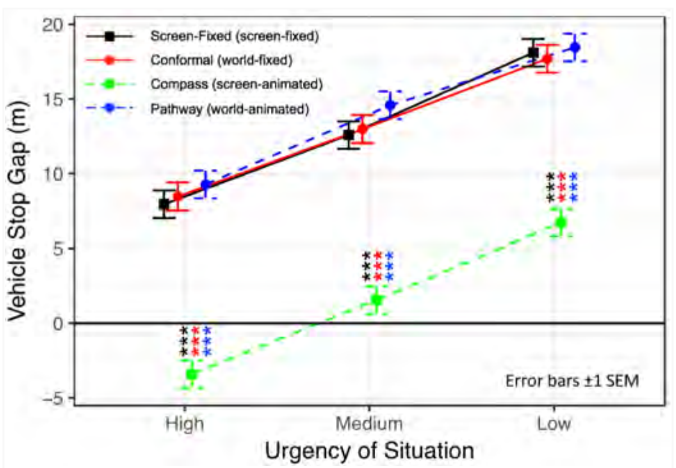

Fig. 8. Final vehicle stop gap by AR graphic and urgency. Participants using the compass displayed much smaller stop gaps than when using any other AR graphic. In the high urgent conditions participants actually stopped on average beyond the pedestrian, meaning that a collision could have occurred.

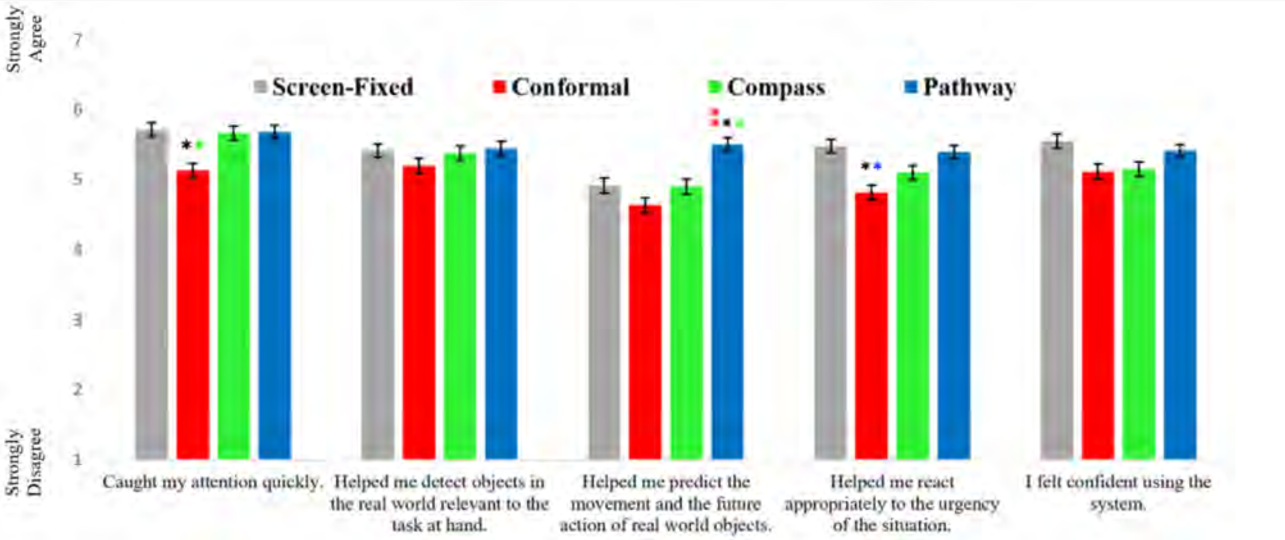

Fig. 9. Subjective feedback ratings for the five question awareness survey from the pedestrian study. Conformal graphics are generally rated more poorly than others, with significance indicated by colored asterisks.

situation, the compass graphic was associated with negative stop gaps, which implies that pedestrian collision could have occurred (but of course did not in this controlled experiment). Results for vehicle stop gap are shown in Figure 8.

Estimates for peak deceleration during breaking were derived by using velocity and time differentials between data points sampled from the test vehicle's CANBUS. Due to some levels of noise in the initial data set, the deceleration feed from CANBUS velocity logs was first filtered using a moving average filter for five samples (given a GPS sampling rate of 20 Hz) before obtaining peak deceleration values. As to be expected, a significant main effect was found for urgency of situation ($F(2, 34)=31.631$, $p<0.001$), with participants expectedly breaking harder in more urgent situations. No main effects were found for AR interfaces.

5.6.2 Driver Eye Behavior

We analyzed participants' eye fixation patterns using a similar method to that described in the navigation study. We defined the AR graphic and the pedestrian actor as significant areas of interest, and coded fixations in these areas. Fixation data were further annotated to specify the experimental condition and dependent measure associated with each fixation event. We calculated FG, GT, and FT and performed an ANOVA. However, no significant differences were found across any of these measures for AR graphic type or urgency.

5.6.3 Subjective Feedback

We analyzed participants' subjective awareness survey responses collected after each trial. As the data did not follow a normal distribution even after logarithmic transformation, we conducted non-parametric Friedman tests of differences among repeated measures and found a main effect for AR graphic style for *capturing attention quickly* (X^2 (3, N = 239) = 12.304, p <.01), *reacting to situation urgency* (X^2 (3, N = 239) = 11.801, p <.01), and *predicting future motion of objects* (X^2 (3, N = 239) = 8.012, p <.05). Post-hoc analysis conducted for main effects using Wilcoxon signed-rank tests revealed that participants gave lower ratings for conformal graphics than screen fixed (p=0.034) and compass graphics (p=0.043) for capturing attention quickly, lower ratings for conformal graphics than screen fixed (p=0.015) and pathway graphics (p=0.038) for reacting to situation urgency, and higher ratings for pathway graphics than conformal (p=0.002), screen fixed (p=0.031), and compass graphics (p=0.030) for predicting future motion of objects. Results are shown in Figure 9.

5.7 Discussion

For stimulus-driven pedestrian tasks, we see that some animated interface designs, such as the compass, may be significantly less effective at facilitating safe vehicle stop gaps as compared to other AR interface design approaches. As mentioned in the navigation study, the nature of the compass graphic requires drivers to cover distance (and consume time) to activate and provide salient spatial cueing, which is in direct conflict with the required action the AR graphic is meant to provoke (i.e., effective braking and stopping). At close range and relatively high urgency situations, drivers may actually have been encouraged to continue moving in order to "complete" the function of the animated compass. Indeed, from empirical observations, we believe this phenomenon explains why participants using the compass in urgent situations exhibited dangerous behavior resulting in negative stop gaps (Figure 8). Such finding suggests that when designing animated AR graphics to communicate a real-world hazard to drivers, it is recommended that such graphics employ animations that are independent of the vehicle's motion itself. The world-animated pathway pedestrian graphic is such an example, which exhibited animation independent of vehicle speed and/or distance travelled still providing important spatial cues about the hazard and even offering a prediction of the hazard's future behavior.

Compass measures for stop gap notwithstanding, for the most part, time to stop, stop gap, and peak deceleration remained similar across all interface design approaches. Additionally, visual attention measures were not necessarily improved by any particular AR interface design. Rather, in cases where urgent action is required, drivers may simply leverage the appearance of a graphic to facilitate improved reaction to threats. Moreover, in our study, the pedestrian actor was a single pedestrian stepping out into an otherwise lifeless scene. The motion and uniqueness of the single pedestrian actor could have caused a pop-out effect, resulting in immediate attentional capture and thus requiring little need for participants to use the AR graphic other than for the basic directional alert. Additionally, since the world-fixed conformal, as well as the world-animated pedestrian graphics, were presented perpendicular to drivers' forward view, it could be argued that all four AR interface designs served as equally visible and salient alerts. These observations could explain why no significant differences were found for AR graphic type or urgency across FG, GT and FT. Thus, designers could consider a simple but salient world- or screen-fixed directional cue to be sufficient for single pedestrian hazard use cases. But it should be cautioned that this recommendation is difficult to generalize beyond the lone pedestrian use case since many driving scenes include a number of moving and potentially hazardous pedestrians.

Lastly, when coding the eye-tracking video, we noticed that world-fixed conformal AR graphics used to cue highly urgent (near) pedestrians would often quickly appear and then move out of the left or right edge of the AR HUD field of view. That is, since nearby pedestrians entering the roadway from the side are by definition positioned to the left and right of the vehicle, it can be difficult if not impossible to conformally annotate these pedestrians given HUD's currently limited horizontal FOV. Thus, in cases where real-world referents are located to the side of the driver, or even on a sidewalk that will eventually track to the side of a driver, care should be taken when designing with world-fixed conformal graphics since they are unable to guide drivers' attention if they are not rendered within the HUD FOV. This finding is supported by the relatively low subjective ratings that participants gave conformal AR graphics on their ability to capture attention, facilitate reaction, and provide prediction on future behavior of the target threat.

6 Comparative Insights Across the Two User Studies

By design, the two user studies presented herein systematically investigated similar, yet different AR interface designs across a navigation and pedestrian usage scenarios. Through analysis and discussion, we can extract some high-level insights that are best presented within the context of both user studies.

First, when considering the arguable merits of using AR to guide drivers' visual attention, designers should consider the usage context; specifically, the type of real- and virtual visual stimuli available. For example, in a goal-directed task such as navigation, an AR graphic may serve as the primary, salient visual cue required to accomplish the target-localization task. Conversely, in a stimulus-driven environment, the AR cue is by definition competing with a real-world visual stimulus. In these cases, the AR cue is not required to complete the task, and in worse cases may compete for attentional resources that otherwise could be allocated to critical real-world visual stimuli.

The effectiveness of conformal graphics in driving is likely strongly tied to the relationship between the shape and orientation of the cue as well as the distance to the cue. For example, AR graphics that are planar in nature will be more salient if presented perpendicular to drivers' line of sight, even though such placements will often not make practical sense (e.g., nearby upright conformal AR graphics may occlude other vehicles in the forward roadway). Conversely, planar conformal AR graphics that lay on the horizon at a distance should be avoided since they are difficult to perceive. It is possible that the effectiveness of some conformal cues could be improved by instead placing nearby conformal cues near-flat on the ground (e.g., rendering a virtual pedestrian walkway on the forward roadway), and far cues upright for easy visibility at distance (e.g., highlighting landmarks, or placing turn arrows on the side of distant buildings instead of the road.

Lastly, the left/right orientation of symbology as well as the direction-of-animation, should be carefully considered when supporting both goal-directed and stimulus-driven tasks. To direct attention in navigation scenarios, clearly, all symbols (e.g., arrows) should point from the driver to the target. However, when designing screen- and world-fixed pedestrian cues (as shown in Figure 6a and 6b), should the symbol face away from the actual pedestrian (to convey direction of actual pedestrian travel), or towards the actual pedestrian in hopes of better guiding drivers' visual attention? Similarly, when animating AR cues to support a goal-directed navigation task, it is reasonable to assume an animation path that, for example, mimics the planned driver route (from the driver to the target). However, motion cues for stimulus-driven tasks may be best rendered in the opposite fashion, since movement from the actual pedestrian to the projected collision point can convey future movements, despite the animation's start position being rendered near the HUD's periphery (and possibly outside the HUD FOV).

7 Limitations

As with many research endeavors, challenging outdoor AR studies that involve human-subjects driving real vehicles are subject to limitations. The use of a GPS system operating at a sampling rate of 20Hz was sufficient to produce results with reasonable granularity, but also functioned as a source of uncertainty and noise for measures relating to driving performance. It is possible that a lack of differences found in peak braking deceleration could have been aggravated by the relatively large amount of noise introduced when deriving the measure from raw data logs. Additionally, the pedestrian study used GPS and Wizard of Oz to synchronize the onset of AR graphics and pedestrian actors, which could introduce small errors despite practice and pilot testing. A similar study using machine vision to detect the test vehicle position, cue pedestrian actors, and render AR graphics could likely produce more pronounced differences in the stimulus-driven study.

Bona fide driver reactions and behavior in real threat scenarios is difficult to induce and authentically re-create in experimental testbeds, whether it be simulated or real-world. The mere fact that drivers were partaking in an experiment likely influenced their perceived level of threat, risk, and expectancy of upcoming experimental events. Limits in available time and resources meant road events were required to occur more frequently than would naturally occur on-road. A future study conducted in a larger test area with more naturalistic conditions (for example, in a city instead of a parking lot) would likely better approximate the true conditions in which AR graphics would be used.

8 Conclusions and Future Work

In this work, we explore the impact of interface design approach on driver performance and visual behavior when engaging in top-down and bottom-up processing. We demonstrate that vehicle- and world-animated approaches using directional and motion-based cues can support accurate judgments with less spatial error of localization than either traditional screen-fixed or world-fixed conformal cues anchored in place. Though compass AR graphics consistently produce the highest judgement accuracy, this design approach does not always produce fast response times and are likely not appropriate for goal-directed tasks that require fine-grained or rigorous localization. AR graphics that employ animation based on a vehicle's continued movement should be avoided in stimulus-driven scenarios that require a driver to cease movement quickly in order to avoid consequences.

Designers can leverage unique gains from both directional and spatially-based conformal cues but should consider carefully in which situations their inclusion is appropriate or safe. In this work, we explore only a small subset of AR design approaches and exclude numerous variations such as temporal dependencies (i.e. task completion-based cue duration vs. predetermined presentation duration) and saliency-related animations (such as graphics blinking or fade-in and -out). Further work is needed to fully conceptualize and understand this application area for AR, particularly to examine the limits of what is achievable within the still relatively small AR design canvas afforded by current HUD technology. Still, this work demonstrates that the design space of automotive AR applications need not be limited to traditional 2D screen-fixed or world-fixed graphics alone. The future effectiveness of AR driving interfaces will depend on the extent to which designers can creatively exploit the strengths of augmented reality while managing the perceptual and attentional ramifications associated with human use of such interfaces.

9 Acknowledgements

The authors wish to thank Alex Miranda for his consultation on the Honda Odyssey vehicle CANBUS operation as well as Joey Wu, Jordon Horengic, Sarah Lee, and Cole Benedict for assisting in coding eye-tracking data, and Matt Chau, Deborah Asabere, Youngjae Lee, and Joey Lee for their work as pedestrian actors during data collection. We further appreciate Ethan Larsen's excellent photography work capturing realistic views of driver perspective.

REFERENCES

[1] I. M. Relations. "Automotive Head-Up Display Market Goes into High Gear," 2018; http://news.ihsmarkit.com/press-release/design-supply-chain/automotive-head-display-market-goes-high-gear.

[2] H. Guo, F. Zhao, W. Wang *et al.*, "Analyzing Drivers' Attitude towards HuD system using a stated Preference survey," *Advances in Mechanical Engineering*, vol. 6, pp. 380647, 2014.

[3] M. Ablassmeier, G. Mcglaun, and G. Rigoll, "Evaluating the potential of head-up displays for a multimodal interaction concept in the automotive environment."

[4] B. Holt. "10 Reasons the 2018 Toyota Prius is the AutoWeb Buyer's Choice Best Alternative Fuel Vehicle," 2018; https://www.autobytel.com/auto-news/awards/10-reasons-the-2018-toyota-prius-is-the-autoweb-buyer-s-choice-best-alternative-fuel-132364/.

[5] J. Meiners. "BMW Vision ConnectedDrive Concept," 2018; https://www.caranddriver.com/news/bmw-vision-connecteddrive-concept-photos-and-info.

[6] R. Glon. "The best head-up displays (HUDs) for your car," 2018; https://www.digitaltrends.com/cars/best-head-up-displays/.

[7] M. Tonnis, C. Lange, and G. Klinker, "Visual longitudinal and lateral driving assistance in the head-up display of cars." pp. 91-94.

[8] Z. Medenica, A. L. Kun, T. Paek *et al.*, "Augmented reality vs. street views: a driving simulator study comparing two emerging navigation aids." pp. 265-274.

[9] W. Narzt, G. Pomberger, A. Ferscha *et al.*, "Augmented reality navigation systems," *Universal Access in the Information Society*, vol. 4, no. 3, pp. 177-187, 2006.

[10] A. May, T. Ross, and Z. Osman, "The design of next generation in-vehicle navigation systems for the older driver," *Interacting with computers*, vol. 17, no. 6, pp. 643-659, 2005.

[11] M. Blanco, "Effects of in-vehicle information systems (IVIS) tasks on the information processing demands of a commercial vehicle operations (CVO) driver," Virginia Tech, 1999.

[12] Y.-C. Liu, and M.-H. Wen, "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*, vol. 61, no. 5, pp. 679-697, 2004.

[13] K. Bark, C. Tran, K. Fujimura *et al.*, "Personal Navi: Benefits of an augmented reality navigational aid using a see-thru 3D volumetric HUD." pp. 1-8.

[14] S. Kim, and A. K. Dey, "Simulated augmented reality windshield display as a cognitive mapping aid for elder driver navigation." pp. 133-142.

[15] P. Froehlich, R. Schatz, P. Leitner *et al.*, "Evaluating realistic visualizations for safety-related in-car information systems." pp. 3847-3852.

[16] R. Eyraud, E. Zibetti, and T. Baccino, "Allocation of visual attention while driving with simulated augmented reality," *Transportation research part F: traffic psychology and behaviour,* vol. 32, pp. 46-55, 2015.

[17] M. R. Savino, "Standardized names and definitions for driving performance measures," Tufts University, 2009.

[18] J. L. Gabbard, J. E. Swan, and D. Hix, "The effects of text drawing styles, background textures, and natural lighting on text legibility in outdoor augmented reality," *Presence: Teleoperators & Virtual Environments,* vol. 15, no. 1, pp. 16-32, 2006.

[19] J. L. Gabbard, J. E. Swan, D. Hix *et al.*, "An empirical user-based study of text drawing styles and outdoor background textures for augmented reality." pp. 11-18.

[20] J. L. Gabbard, G. M. Fitch, and H. Kim, "Behind the Glass: Driver challenges and opportunities for AR automotive applications," *Proceedings of the IEEE,* vol. 102, no. 2, pp. 124-136, 2014.

[21] M. Tonnis, V. Broy, and G. Klinker, "A survey of challenges related to the design of 3d user interfaces for car drivers." pp. 127-134.

[22] Y. Inuzuka, Y. Osumi, and H. Shinkai, "Visibility of head up display (HUD) for automobiles." pp. 1574-1578.

[23] D. R. Tufano, "Automotive HUDs: The overlooked safety issues," *Human factors,* vol. 39, no. 2, pp. 303-311, 1997.

[24] K. W. Gish, and L. Staplin, "Human factors aspects of using head up displays in automobiles: A review of the literature," 1995.

[25] D. Gupta, "An empirical study of the effects of context-switch, object distance, and focus depth on human performance in augmented reality," Virginia Tech, 2004.

[26] C. Merenda, H. Kim, J. L. Gabbard *et al.*, "Did You See Me?: Assessing Perceptual vs. Real Driving Gains Across Multi-Modal Pedestrian Alert Systems." pp. 40-49.

[27] L. Pomarjanschi, M. Dorr, and E. Barth, "Gaze guidance reduces the number of collisions with pedestrians in a driving simulator," *ACM Transactions on Interactive Intelligent Systems (TiiS),* vol. 1, no. 2, pp. 8, 2012.

[28] H. Kim, A. Miranda Anon, T. Misu *et al.*, "Look at Me: Augmented Reality Pedestrian Warning System Using an In-Vehicle Volumetric Head Up Display." pp. 294-298.

[29] H. Kim, J. L. Gabbard, A. M. Anon *et al.*, "Driver Behavior and Performance with Augmented Reality Pedestrian Collision Warning: An Outdoor User Study," *IEEE transactions on visualization and computer graphics,* vol. 24, no. 4, pp. 1515-1524, 2018.

[30] A. Paran, "Bottom-up and top-down processing," *English Teaching Professional,* vol. 3, no. 11, 1997.

[31] P. R. Boyce, "Illumination," *Handbook of Human Factors and Ergonomics, Fourth Edition*, pp. 673-698, 1997.

[32] H. Kim, J. D. Isleib, and J. L. Gabbard, "Virtual Shadow: Making Cross Traffic Dynamics Visible through Augmented Reality Head Up Display." pp. 2093-2097.

[33] F. E. Satterthwaite, "An approximate distribution of estimates of variance components," *Biometrics bulletin,* vol. 2, no. 6, pp. 110-114, 1946.

[34] J. H. Lim, O. Tsimhoni, and Y. Liu, "Investigation of driver performance with night vision and pedestrian detection systems—Part I: Empirical study on visual clutter and glance behavior," *IEEE Transactions on Intelligent Transportation Systems,* vol. 11, no. 3, pp. 670-677, 2010.

[35] J. J. Anaya, P. Merdrignac, O. Shagdar *et al.*, "Vehicle to pedestrian communications for protection of vulnerable road users." pp. 1037-1042.