Augmented Reality vs. Street Views: A Driving Simulator Study Comparing Two Emerging Navigation Aids

Zeljko Medenica
University of New Hampshire
Durham, NH, USA
zeljko.medenica@unh.edu

Andrew L. Kun
University of New Hampshire
Durham, NH, USA
andrew.kun@unh.edu

Tim Paek
Microsoft Research
Redmond, WA, USA
timpaek@microsoft.com

Oskar Palinko
University of New Hampshire
Durham, NH, USA
oskar.palinko@unh.edu

ABSTRACT

Prior research has shown that when drivers look away from the road to view a personal navigation device (PND), driving performance is affected. To keep visual attention on the road, an augmented reality (AR) PND using a heads-up display could overlay a navigation route. In this paper, we compare the AR PND, a technology that does not currently exist but can be simulated, with two PND technologies that are popular today: an egocentric street view PND and the standard map-based PND. Using a high-fidelity driving simulator, we examine the effect of all three PNDs on driving performance in a city traffic environment where constant, alert attention is required. Based on both objective and subjective measures, experimental results show that the AR PND exhibits the least negative impact on driving. We discuss the implications of these findings on PND design as well as methods for potential improvement.

Author Keywords

Driving simulation, Augmented reality, Navigation.

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Measurement, Design, Performance, Experimentation, Human Factors.

INTRODUCTION

In some countries driving is the primary mode of commuting. For example, according to the U.S. Census Bureau [12], Americans spend more than 100 hours a year commuting to work. Given the large amount of time that many people spend behind the wheel, and the increasing availability of computational resources that can now operate inside a vehicle, companies have been introducing myriad mobile services and functionalities into the consumer market just for drivers. A few notable examples are hands-free voice dialing, live traffic reports, automated

directory assistance, infotainment systems, and personal navigation devices (PNDs). Unfortunately, the question of how these in-car services impact driving performance is often left unanswered. The focus of this paper is the impact of PNDs on visual attention and driving performance.

In recent work, Kun et al. explored the effects of two PNDs on visual attention and driving performance [9]. One was a standard PND that displays a map and provides turn-by-turn instructions by voice. The other was a PND that only uses voice instructions. They found that participants spent more time looking at the road ahead with the voice-only PND, which in turn translated into better driving performance. Nevertheless, participants preferred the standard PND. So, how can a PND allow drivers to maintain high visual attention on the road for good driving but also generate high user satisfaction? We believe that this can be accomplished with augmented reality (AR) personal navigation devices.

AR PNDs integrate a virtual navigation route into the real-world scene by displaying it directly on the windshield with a head-up-display (HUD). No glancing towards an in-car display is required. Previous research has indicated that HUDs have safety benefits such as shorter reaction time to sudden events [16] in comparison to head-down displays (HDDs), which are today's standard.

With advances in HUD technology and location services, we expect that AR PNDs will soon be commercially available. However, what is available to consumers now are HDD-based solutions that incorporate elements of AR technology. One such solution is the egocentric *street view* (SV) PND. SV PNDs use sequences of still images of the road and surrounding streets, augmented with a virtual navigation route, to help users orient themselves with respect to visual landmarks. While this kind of AR can be beneficial for pedestrians or passenger-side navigators, it is unclear whether an SV PND is appropriate for drivers. In contrast to AR PNDs, SV PNDs do not overlay information on the real world, but rather use sequences of pictures of the real world taken at a prior time; hence, driving performance may be affected by the process of resolving differences between the real world and images displayed by the PND. Furthermore, since commercially available SV PNDs, such as Google Maps Navigation with Street View [3], operate on a HDD, the AR benefits may be offset by the need to look at the display.

In this paper, our goal is to explore the impact of AR and SV PNDs on visual attention, driving performance and user satisfaction. We include a standard PND in this study as a baseline. We focus our exploration on city roads because that is where drivers' interactions with PNDs are most likely to be challenging. With unknown city roads, drivers have to make timely decisions about turns while heeding both pedestrians and traffic. In pursuit of our goal we propose three hypotheses.

1. *With standard PNDs as a baseline, AR PNDs allow drivers to spend more time looking at the road ahead than SV PNDs.* Previous work by Kun et al. [9] indicates that following voice-only turn-by-turn instructions allows drivers to spend about 6.5% more time looking at the road ahead than using a standard map-based PND with the same voice instructions. Similarly, AR PNDs will, by design, allow drivers to keep their eyes on the road. However, we expect that with SV PNDs drivers will need to resolve differences between the real world and PND images and will thus spend even more time looking at the PND display than with the standard PND.

2. *The differences in visual attention between the PNDs are associated with differences in driving performance, with AR PNDs allowing for the best driving performance.* Given that AR PNDs allow drivers to keep their eyes on the road, we expect that these PNDs will allow for better driving performance than standard and SV PNDs. Again, as we expect SV PNDs to incur the highest cost in visual attention, we expect them to exhibit worse driving performance than even standard PNDs.

3. *When comparing different characteristics of AR and SV PNDs, users will express a preference for AR PNDs.* Even if our experiment indicates that AR PNDs are superior to SV PNDs in terms of visual attention and driving performance, will users prefer them? For example, in the experiment of Kun et al., using the voice-only PND resulted in better visual attention and driving performance, but users preferred to see maps along with hearing voice instructions. We expect that users will actually prefer AR PNDs.

The structure of this paper is as follows. First, we survey related research and describe the experiment we conducted to evaluate the three hypotheses above using a high fidelity driving simulator. Next, we report our findings and discuss their implications on PND design in general. Finally, we conclude with directions for future research.

RELATED RESEARCH

Driving performance is often characterized using the averages of performance over multiple, relatively long, road segments. However, as pointed out by Kun et al. [9], such averages can miss short, consistent sequences where engaging in some secondary action while driving is followed by increased variance in driving performance measures. Kun et al. detected cross-correlation peaks between a sequence representing glances away from the road ahead and sequences representing the variance of lane position and steering wheel angle, where the variances were calculated over a 1 second-long sliding window. These peaks indicate that an increase in lane position and steering wheel variance followed reduced attention to the road ahead. Kun et al. also found that as the percent time spent looking at a PND increases, the magnitude of the peaks also increases. In light of this prior research, we propose testing our second hypothesis using the cross-correlation technique used by Kun et al.

Since displaying navigation instructions using AR and SV PNDs is accomplished using HUD and HDD displays, respectively, it is worth exploring how the two presentation techniques compare with respect to their influence on visual attention, driver safety and driving performance. A plethora of research covers these topics. Besides [16] mentioned in the introduction, the experiment performed by Liu [11] offers additional evidence that HUDs may provide safety benefits; in a driving simulator study, the author found that when using a HUD, drivers tend to respond faster to unexpected road events for both low and high workload conditions. These conclusions were corroborated by Horrey et al. (2003) [6], in which HDDs were associated with decreased performance, as reflected in e.g. slower reactions to hazardous events and side (secondary) tasks. The authors suggest that hazard detection mostly requires focal vision, whereas vehicle control relies on ambient vision. Thus, when drivers direct their focal vision at HDDs, their ambient vision-supported vehicle control does not suffer, though their focal vision-dependent hazard detection does.

Even though HUDs can provide safety benefits, practically all commercially available PNDs rely on a map displayed on a HDD with a navigation route overlaid on top of it. The fact that a screen is present entices drivers to steal glances at it and away from the road ahead [9]. This is a problem. As demonstrated in a driving simulator experiment by Horrey et al. (2006) [7], who explored the effect of interacting with in-car devices on driving performance and visual attention, as users spent less time looking at the road ahead, their lane position variability increased.

In-car HUDs leveraging AR navigation still need to overcome technical challenges before they hit the consumer market. That does not prohibit, however, researchers today from investigating tomorrow's AR navigation technology in HUDs. Burnett found that HUD-based navigation devices can help reduce navigational errors compared to HDD-based devices [1]. More recently, in a driving simulator study, Kim and Dey [8] presented users with a full windshield HUD which overlays directions on the road and also depicts a map of the surrounding area which rises from the road ahead on a plane perpendicular to the driver's eye-gaze. The intention was to provide drivers with both local route guidance and global awareness of the road network, while keeping their eyes fixed on the road. Consistent with our hypothesis that AR PNDs allow drivers to spend more time on the road ahead, Kim and Dey found that for older

Figure 1. 180º field of view driving simulator.

Figure 2. Experimental setup inside the vehicle.

drivers, their approach reduces the effect of divided attention and improves driving in comparison to standard PNDs.

In addition to navigation, in-car AR techniques can be valuable for accomplishing other important tasks, such as alerting drivers to road hazards. Strayer and Johnston [17] used an AR-based 3D arrow displayed on a HUD to direct a driver's attention to imminent danger. In simulator-based experiments, this approach outperformed using an exocentric, bird's-eye schematic of the car with an arrow indicating the direction of the source of danger. Fröhlich et al. explored using an AR HDD for navigational and urgent tasks [2]. They found that the AR HDD had a positive impact on user-perceived safety for urgent tasks.

EXPERIMENT
We now describe our experimental setup and procedure. We also detail each of the three PNDs that were used in our study as well as how we computed the dependent variables.

Equipment
The experiment was performed in a high-fidelity driving simulator (Figure 1) with a 180° field of view screen and a full-width automobile cab. The cab sits on top of a motion base which simulates car movements for braking and accelerating as well as bumps on the road. As shown in Figure 2, the simulator was equipped with two eye-trackers which track subjects' gaze and head position using two pairs of cameras mounted on the dashboard. Figure 2 also shows the location of the in-car LCD screen which was used as the display for the standard PND and the SV PND. The LCD screen was placed on top of the dashboard, which is a common place for contemporary PNDs and smart phones with navigation capabilities. The size of the LCD screen was 3.5 inches (about 9 cm) diagonally which falls within the typical size of commercially available navigation devices.

Method

Participants
Eighteen university students participated in the experiment. They were between 18 and 37 years of age (mean age 20.5 years, standard deviation 4.7 years). As compensation each received a $20 gift card to a popular store chain.

Navigation Aids
Each participant performed navigation with each of the following three PNDs:

1. Augmented reality PND (AR PND): Our AR PND overlays a translucent navigation route on the real world scene. We presented navigation directions using a narrow semi-transparent surface which was suspended above the center of the road at a height of about 2 meters. This created the visual effect of a navigation route hovering above the vehicle, similar to the Virtual Cable™ [20]. In our driving simulator, the route is projected onto simulator screens, as shown in Figure 3, instead of a windshield HUD, which is unavailable.

2. Street View PND (SV PND): The egocentric street view PND utilizes the LCD to display a sequence of images of the world from the driver's perspective. This sequence is augmented with a translucent navigation route (Figure 4). The navigation route was displayed using a wide, transparent, road level surface, as shown in Figure 4. We used this road-level surface because of its similarity to commercially available HDD-based PNDs (e.g. [3]). Simulating the SV PND required having a stream of images of the surrounding environment. This stream was generated by a driving simulator running in parallel with the one that was operated by our participants. An SV PND uses images of the world taken at some earlier point in time, and this was faithfully represented in our experiment. Specifically, fixed entities (roads, buildings, etc.) were the same in both simulations, while parked and moving vehicles and

Figure 3. AR PND from the driver's perspective.

Figure 4. LCD screen displaying SV PND.

Figure 5. LCD screen displaying SPND.

Figure 6. Simulated two-lane city road. The environment included ambient traffic and unexpected events, such as this pedestrian emerging from behind a parked vehicle.

pedestrians were different. The SV PND displayed a new image every 15 meters, which is the distance at which images are also taken for Google Street View [4] in a city environment. Thus, participants experienced an SV PND similar to Google Maps Navigation, which uses Google Street View data. Note that the season, the weather and the time of day were identical in our two simulations. In real-life scenarios, these variables can be quite different between the outside world and Google Street View data.

3. Standard PND (SPND): Similar to the most basic commercially available PNDs, our LCD screen (Figure 5) presented users with a real-time map of the surrounding environment as well as the position of the vehicle in a simulated world. We used the same LCD for both SV and SPND. The 2D map was presented in a dynamic, exocentric, forward-up view. The car (represented by a small triangle on the map) always remained in the center of the screen, while the road moved about it. Note that our baseline of using the 2D map may be more difficult than the 3D angular map utilized by many PNDs today since transforming a 2D map into 3D has been shown to require mental effort [15].

Since most commercially available PNDs ship with spoken directions enabled, for our experiment all three PNDs also utilized identical turn-by-turn spoken directions. In order to eliminate potential problems with comprehension of synthesized speech (e.g., Lai et al. [10]), we used pre-recorded voice directions by a female voice talent.

Procedure
After filling out the consent forms and personal information questionnaires, participants were given an overview of the driving simulator and descriptions of the three navigation conditions. Next, they proceeded to complete three navigation experiments, one with each of the PNDs. Before each condition, we provided subjects with about 5 minutes of training using that PND. For training, users followed PND navigation instructions in a city environment. In order to circumvent order effects, we counterbalanced the presentation order of the PNDs between subjects.

As shown in Figure 6, subjects drove on two lane city roads which included ambient vehicles, moving pedestrians, traffic signs and lane markings. Lanes were 3.6 meters wide. Subjects were instructed to drive as they normally would in real life and to obey all traffic laws. They were also instructed (and trained) to pay attention to unexpected

events, such as pedestrians emerging from behind parked vehicles (Figure 6) or vehicles braking suddenly. These unexpected events are not uncommon in city driving. Furthermore, the ability to avoid collisions when such unexpected events occur is a valuable measure of driving performance.

For all three PNDs, the participants drove a different route with two unexpected events in each case. Figure 7 shows the first route used in this experiment. For the second route we reversed the direction of travel, and the third route was the mirror image of the first route. In short, all three routes were of the same length (about 10 km) and complexity. However, the turn-by-turn directions for each route were different. Thus, there was no risk of subjects remembering navigation instructions from the previous route. Each route had both long (400 and 800 meter long) and short (200 meter long) segments with many intersections on the given path. On average it took about 15 minutes to traverse a route. The presentation order of routes was the same for all subjects.

After each PND condition, subjects filled out a NASA-TLX questionnaire. Finally, at the end of the experiment, subjects ranked their level of agreement with various statements pertaining to the PNDs and provided written and verbal feedback about the experiment.
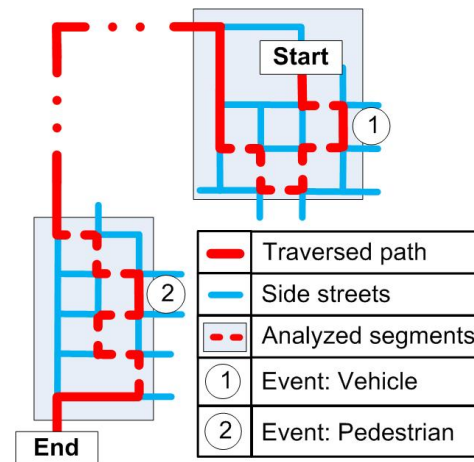

Figure 7. One virtual route navigated by subjects and the short segments included in the analysis. The other two routes were created by travel direction reversal and mirroring.

We chose a within-subjects factorial design experiment with PND type, *Nav*, as our independent variable.

We considered conducting a 2x2 study with the display type (HUD or HDD) and the type of navigation instructions (AR or SV) as the independent variables. This would have allowed us to directly compare the effects of the independent variables and investigate their interaction. However, such a 2x2 study would have included two confounding factors, both related to visual attention. First, HUD SV PNDs would have required drivers to focus on a portion of the windshield in order to view navigation instructions. Unfortunately, that shift of focus away from the road would have made HUD and HDD SV PNDs too similar. Second, HDD AR PNDs introduce the possibility of driving exclusively using the live feed on the HDD. This would have changed the way drivers use focal and ambient vision, making that condition markedly different from the others. Because we still wanted to compare differences in visual attention, we decided to focus on HUD AR and HDD SV PNDs as holistic conditions, using SPNDs as a baseline.

We measured multiple dependent variables:

- *Percent dwell time (PDT) on the road ahead,* which measures the percent of time drivers spend looking at the road ahead. A low value indicates that a driver is distracted, which in turn can lead to collisions.
- *Cross-correlation peaks* to detect short periods of deterioration in driving performance following glances away from the road. The existence of statistically significant peaks indicates that such deteriorations do indeed follow reduced visual attention to the road [9]. The larger the size of a peak, the larger and/or more frequent the deteriorations.
- *Average driving performance measures,* which included the absolute values of first difference and variances of lane position, steering wheel angle and velocity. The absolute value of first difference (AVFD) is defined as:

$$AVFD\{x[n]\} = |x[n] - x[n-1]|$$

Equation 1. Absolute value of first difference.

where $x[n]$ is a discrete time sequence and n is an integer. In each case, higher values for driving performance measures indicate deterioration. We also evaluated mean velocity. A low mean velocity for a portion of the road may indicate that the driver was concerned about safety or was otherwise distracted.

- *Number of collisions* is a rather coarse measure of driving performance, which represents the number of times that the subjects came into contact with another car or a pedestrian while driving. A collision is more likely to happen when the driver is confronted with unexpected events. However, it is important to note that our unexpected events were specifically designed to be avoidable by an alert driver. Hence, if a collision occurs, it is most likely due to driver distraction.

- *NASA-TLX score* [5], which measures subjective opinion about the perceived cognitive workload while using each of the three navigation devices.
- *Level of agreement with preferential statements,* which reflects subjective opinion about PNDs, was collected using a 5-point Likert scale.

Using an eye-tracker we were able to automatically classify gazes as being directed at the road ahead, at the LCD, or somewhere else inside the cabin (e.g. the speedometer). The sampling frequency of the eye-tracker was 60 Hz. In rare occasions when the eye-tracker was not able to resolve where the subject was directing his/her gaze, the experimenter made classifications using video footage from a camcorder installed on the dashboard.

Lane position, steering wheel angle, vehicle speed and collisions with other entities in the simulation were recorded by our driving simulator software at a frequency of 10 Hz.

Data segmentation: The city routes in our experiment can be broken up into segments by treating roads between two intersections as separate segments. We calculated the PDT and all of our driving performance results (except the number of collisions), such as the absolute value of first difference and mean velocity, using data from 13 short segments (dashed lines in Figure 7).

All 13 short segments had the same characteristics, thereby controlling factors that could potentially confound our results. In particular, the segments were 200 meters long measured from the centers of the adjacent intersections. Furthermore, at both the beginning and end of each segment, there was a four-way intersection where participants made either a right or left turn. Finally, participants did not encounter any unexpected events (represented by 1 and 2 in Figure 7) in the 13 segments we used to analyze visual attention and driving performance. Unexpected events often require sudden braking and steering wheel motion, which in turn can result in very large first differences and variances for these measures, making comparisons with other segments difficult. For the purpose of counting the number of collisions only, we used 15 short segments, including the ones with unexpected events, since collisions are more likely to occur there.

In analyzing all of the segments, we excluded data collected over the first 60 meters and the final 40 meters of a segment, and analyzed data generated over (200–60–40) = 100 meters. This was done because driving performance is different between the excluded and analyzed portions of the segments. For example, at the beginning of a segment, drivers are completing the turning maneuver that is necessary to get through the previous intersection. And at the end of a segment, they are decelerating before entering the next intersection. Thus, the resulting first differences and variances can be much larger than those encountered

away from intersections, which makes it difficult to compare excluded and analyzed portions of segments.

Using the aforementioned short segments, we calculated the following measures.

Visual attention: For each participant p and navigation aid *nav*, we calculated the average percent dwell time, $PDT_{p,nav}$, on the road ahead by finding the ratio of the sum of dwell times for all 13 segments and the sum of the total time spent traversing all 13 segments. We defined "looking at the road ahead" as looking at one of the three projection screens of the simulator. Similarly, we calculated the percent dwell time on the LCD screen when participants used the SV and SPND. Finally, we calculated the percent dwell time on the rest of the cabin (speedometer, rear view mirror, etc.).

Cross-correlation: We estimated the cross-correlation between the eye-gaze vector (*EGV*) and two driving performance vectors (*DPV*). *EGV* is a sequence of 0s and 1s, where 1s represent moments when the driver's gaze returns to the road ahead (e.g. from the LCD). The two *DPV*s are the absolute values of first difference (defined in Equation 1) for lane position and steering wheel angle.

Before calculating the *EGV* and cross-correlations, we brought both correlates to the same sampling frequency. Since the sampling frequency was 60 Hz for the eye-tracker and 10 Hz for the simulator we down-sampled the eye-tracker signal to 10 Hz. Thus, cross-correlation peaks (if any) are located at lag values that are integer multiples of 100 milliseconds (1/10Hz).

For each navigation aid *nav*, $Rlp_{nav}[lag]$ is the cross-correlation between *EGV* and DPV_{lp} for lane position. For each PND $Rlp_{nav}[lag]$ was calculated as the average of cross-correlations for each of the 13 segments and each of the 18 participants. Similarly, $Rsw_{nav}[lag]$ is the cross-correlation between *EGV* and DPV_{sw} for steering wheel angle and was calculated analogously to $Rlp_{nav}[lag]$. The *lag* variable indicates how much the change in driving performance measure lags behind the return of the gaze to the road ahead. Thus, any peaks in the cross-correlation functions might represent drivers' corrective actions taken when their attention returns back to the road ahead.

Average driving performance measures: For each participant, segment and navigation type we calculated the average absolute values of first differences and the variances of the following three driving performance measures: lane position, steering wheel angle and velocity. We also calculated the average velocity for each segment. For each participant and navigation aid we then calculated the averages for these variables over 13 segments.

Number of collisions: We reviewed the simulator log files for collisions between the participants' vehicle and surrounding objects.

	AR	SV	S	p-value
road	96.4	86.8	89.3	<0.0001
LCD	■	9.9	7.4	<0.0001
cabin	3.5	3.2	3.2	0.788

Table 1. Percent dwell time (PDT) on road, LCD and other parts of the cabin as a function of PND type.

Subjective assessment: NASA-TLX scores were calculated using NASA TLX for Windows [13]. Agreement levels with preferential statements and written comments were transcribed from paper forms. We also solicited qualitative verbal comments about the experiment from participants.

RESULTS

Visual Attention

Table 1 shows the PDT on the road ahead, the LCD screen and the rest of the cabin for each PND type. We conducted a repeated measures ANOVA to assess the effect of different PNDs on visual attention using PDT on the road ahead as the dependent variable. The analysis revealed a significant main effect on PDT for *Nav* ($F_{2,24}$=121.647, $p < 0.0001$). PDT for AR was the highest at 96.4%, while SPND and SV PND had 89.3% and 86.8%, respectively. Post-hoc comparisons indicated significant differences between AR and SPND (p<0.0001), AR and SV (p<0.0001), and SV and SPND (p=0.008).

As hypothesized, PDT dropped significantly when participants used either of the HDDs. More interestingly, SV PND required more visual attention (as measured by PDT) than using the simple, 2D SPND. To corroborate this finding, we conducted a repeated measures ANOVA with PDT on the LCD as a dependent variable comparing SPND and SV PND. Again, we found a statistically significant difference ($F_{1,12}$=23.091, p<0.0001) with average PDT on LCD being much higher for SV (9.86%) than SPND (7.42%). We also confirmed that PDT on the rest of the cabin was not significantly affected by PND type. In short, our results seem to indicate that resolving differences between SV PND images and the observed world takes time and effort.

Driving Performance

Cross-correlation

With respect to cross-correlation, we observed a relationship between the eye-gaze vector and both lane position and steering wheel angle DPVs, as shown in the peaks of the two cross-correlation functions $Rlp_{nav}[lag]$ and $Rsw_{nav}[lag]$ in Figure 8.

Following Kun et al. [9], we evaluated the statistical significance of these peaks by employing a randomization test inspired by Veit et al. [19]. The procedure operates as follows. When estimating cross-correlations, we employ EGV and DPV vectors from matching segments. In the case of a randomization test, we employ an EGV from one segment and a DPV from a different segment. Since an EGV and a DPV from different segments are not related, the resulting cross-correlation estimate represents a chance outcome. We repeated this procedure 1000 times, thus calculating 1000 chance cross-correlation values for each value of lag. To determine if $Rlp_{nav}[lag]$ (or $Rsw_{nav}[lag]$) is statistically significant for a given value of lag, we assessed whether $Rlp_{nav}[lag]$ (or $Rsw_{nav}[lag]$) is larger than the $(p \cdot 1000)^{th}$-largest cross-correlation value generated using mismatched segments, where p is the desired significance level. For example, if $p=0.01$, $Rlp_{nav}[lag]$ (or $Rsw_{nav}[lag]$) is statistically significant if it is larger than the 10th–largest $(0.01 \cdot 1000 = 10)$ cross-correlation value as estimated using mismatched segments.

As we can see in Figure 8 (top) there are statistically significant peaks in $Rsw_{nav}[lag]$ for all three PNDs, at the $p=0.01$ level. The most prominent peaks appear at 0.6 sec for SV and SPND and at 0.8 sec for AR PND. These peaks indicate that on average, the periods of looking away from the road ahead are followed by a larger change in the steering wheel angle (possible corrective actions) than in usual circumstances. Note that there is also a significant peak for AR. Such peaks occur when participants cast occasional glances towards the speedometer, steering wheel or dashboard. Similarly, the bottom graph of Figure 8 shows the most prominent peaks for $Rlp_{nav}[lag]$ at 0.6 sec for SV PND and SPND and at 3.6 sec for AR PND.

According to our results, glancing away from the road ahead does seem to influence driving performance for all three PNDs. We can also rank the size of the effect for the three PNDs by comparing the magnitudes of the cross-correlations for the most prominent peaks. In Figure 8, we observe that the effect size is the smallest for the AR PND and the largest for the SV PND. The relatively large difference in effect size between AR PND, on the one hand, and SV and SPND on the other, might be attributed to the difference in display type: HUD for AR and HDD for SV and SPND. However, we also see that the SPND cross-correlation peaks are consistently smaller than SV PND peaks. Again, this indicates that resolving differences between SV PND images and the observed world may be cognitively taxing (certainly time consuming), even more so than receiving navigation instructions from a 2D map.

Average Driving Performance Measures
We conducted a repeated measures ANOVA to assess the effect of different PNDs on each of the average driving performance measures. As hypothesized, based on [9], we found no significant difference between the three PNDs regarding any of the average driving performance measures. As in [9], averaging performance measures over entire
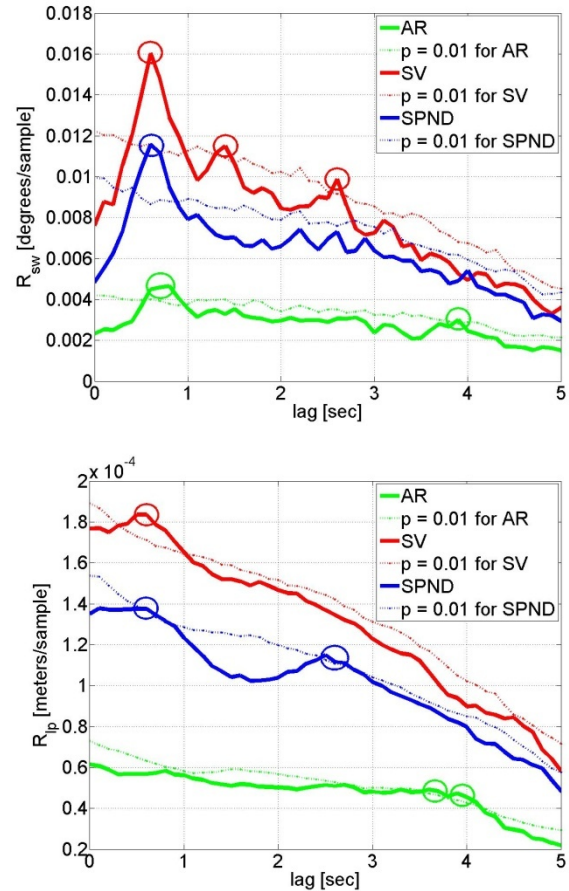

Figure 8. Cross-correlation between eye-gaze vector (EGV) and driving performance vector (DPV) for steering wheel angle (top) and lane position (bottom). Circled peaks are statistically significant at the p=0.01 level.

segments hides the immediate effects of glancing away from the road. However, these effects are evident in our cross-correlation analysis.

Collisions
There were no collisions with pedestrians or ambient traffic for any PND on segments without unexpected events. There were 8 collisions with vehicles on segments with unexpected events: 2 for AR, 3 for SV and 3 for SPND. Clearly, the occurrence of collisions did not depend on the PND type.

Subjective Assessment

NASA-TLX
Users' average NASA-TLX ratings were 28.7, 38.7 and 33.4 for the AR, SV and SPND, respectively. We performed a one-way ANOVA to examine the effect of PND on these subjective workload ratings. Our analysis revealed a significant main effect of *Nav* on workload ($F_{2,24}=6.759$, $p<0.005$). Post-hoc comparisons indicated that subjects experienced less load using the AR than the SV PND ($p<0.0001$).

Statement	Agreement	AR	SV	SPND	p (χ^2)
My driving was best when using _____ interface.	highly agree or agree	72.2	11.1	38.9	**0.014** (8.49)
	highly disagree or disagree	16.7	61.1	50	
I prefer to have a _____ for navigation.	highly agree or agree	66.7	22.2	38.9	**0.023** (7.53)
	highly disagree or disagree	16.7	72.2	27.8	

Table 2. Level of agreement with preferential statements.

Preferential Statements
Table 2 shows the percent of participants who agreed (white cells) or disagreed (shaded cells) with two preferential statements. Note that the percentages do not always sum up to 100% since some subjects were undecided. For each statement we performed a Friedman non-parametric test with respect to *Nav*.

Table 2 shows a significant main effect on the subjective judgment about best driving performance (p=0.014). Participants ranked AR PND very highly (72% highly agreed or agreed) in comparison to others, while both SV and SPND were perceived as detrimental to driving (61% and 50% disagreed or highly disagreed, respectively). Using the Wilcoxon Signed Rank test for pairwise comparisons, we found significant differences between AR and SPND (p=0.027) and AR and SV (p=0.003). Clearly, most participants felt that the AR PND allowed for the best driving performance.

Responses to the second preferential statement in Table 2 indicate that subjects liked the AR PND. Using the Wilcoxon Signed Rank test, we found that participants significantly preferred the AR PND to both the SV (p=0.007) and SPND (p=0.045) and that participants significantly preferred the SPND over the SV PND (p=0.038).

Written and Verbal Comments
AR PND: Ten of the 18 participants (P) provided either written or verbal feedback indicating a preference for the AR PND over the other two PNDs. Quote 1 (Q1) of a written comment by participant 6 shows that P6 found the AR PND easy to use, presumably because it did not result in high cognitive load.

Q1 (P6): "*The windshield cable system* [AR PND] *was the easiest to use, and least obtrusive. … very intuitive to use.*"

On the other hand, two participants explicitly indicated (in verbal comments) that they liked the AR PND the least because the PND was always present in their field of vision.

SV PND: While four participants mentioned liking the SV PND, none singled it out as the best of the three PNDs.

However, fully ten participants reported disliking the SV PND, with P1 calling it "horrible." P15 felt there was "too much going on" when using the SV PND. Even more interestingly P10 and P16 indicated that images in the SV PND did not correspond well to the (simulated) real world.

Q2 (P10): "*The cars in it [SV PND] were different than the real ones and so that was distracting.*"

Q3 (P16): "*The 3D display did not relate well mentally to the on-road task and is less preferred than even the 2D GPS display.*"

SPND: Six participants reported liking the SPND. No participants explicitly indicated disliking it, although two said they did not like having to look at the two HDD-based PNDs. Three liked that the SPND provided information beyond the upcoming turn.

Spoken instructions: Four participants explicitly reported liking spoken instructions. Three (P1, P10 and P14) reported relying on spoken instructions and ignoring the visual information. Eye tracker data indicates that P10 and P14 had some success in ignoring the HDD PNDs and focusing on the road ahead. However, all three participants had the lowest PDT on the road ahead when using the SV PND and the highest when using the AR PND.

DISCUSSION
We started this study by proposing three hypotheses. We now consider each hypothesis in light of our results.

1. *With standard PNDs as a baseline, AR PNDs allow drivers to spend more time looking at the road ahead than SV PNDs.*

Participants' average percent dwell time (PDT) on the road ahead was highest for the AR PND (96.4%), followed by the SPND (89.3%) and the SV PND (86.8%). Clearly, the HUD-based AR PND allows users to keep their eyes on the road more than the HDD-based SPND and SV PND. This result was also supported by NASA TLX scores which showed that participants found the SV PND more difficult to use than the AR PND.

The fact that we observed a difference in PDT between SV and SPND suggests that PDT is not solely a function of display modality. Rather, it is likely that participants found it difficult to resolve differences between the real world and SV images. This explanation is supported by the significantly more frequent glances at the display in the SV condition than with the SPND (19.1 and 15 on average per participant, respectively), despite no significant difference in mean glance durations between the two (0.47 sec and 0.45 sec, respectively). Subjective assessments also support the explanation. In preferential statements, participants indicated that they preferred the SPND over the SV PND. Furthermore, ten participants indicated disliking the SV PND while only two indicated disliking both HDD-based PNDs. Finally, two participants specifically echoed our

suggestion that resolving differences between real and SV PND images is difficult.

2. *The differences in visual attention between the PNDs are associated with differences in driving performance, with AR PNDs allowing for the best driving performance.*

We observed statistically significant correlations between returning the driver's gaze to the road and changes in two driving performance measures: lane position and steering wheel angle. This suggests an association between glances directed away from the road and degradation of driving performance. Furthermore, ranking the amplitudes of prominent peaks in the cross-correlations for each PND we can see that AR PND consistently allows for the best driving performance, while using the SV PND results in the worst performance. Comparing peak magnitudes, SV PND appears to be even worse than our baseline SPND. However, ranking of amplitudes does not allow us to draw conclusions about how using the different PNDs relates to the risk of a collision. In fact, under our experimental conditions, we did not find any differences in the number of collisions for the three PNDs.

3. *When comparing different characteristics of AR and SV PNDs users will express a preference for AR PNDs.*

In evaluating preferential statements, our participants indicated that AR PND allowed for the best driving performance and that they preferred it to the other two PNDs. Furthermore, in written or verbal comments, ten of 18 participants told us they preferred the AR PND over the other two PNDs. Only two participants indicated disliking the AR PND. The SV PND faired much worse, with four participants liking it but fully ten considering it to be the worst of the three PNDs.

Design Implications

Designing AR PNDs
In order to keep drivers' visual attention fully on the road, our results and the previous research literature suggest that interface designers consider one of two choices: AR PND and voice-only PND [9]. Only AR PND however offers visual confirmation that users are still on the right route. Furthermore, users expressed a preference for AR PND over SPND whereas in the study of Kun et al. [9] they expressed a preference for SPND over voice-only PND.

While AR PND stands out as a safe and agreeable PND, our participants brought to our attention two concerns that merit further study. First, our implementation of an AR PND did not provide global navigation information; it only informs drivers about the current route to follow. Three participants in fact indicated they would have appreciated receiving information about upcoming turns. One approach to address this issue is that proposed by Kim and Dey [8]. Second, overlaying routes for long stretches of road may be distracting. Two participants stated that they disliked the elevated surface in the AR PND because it was always present in their peripheral vision. Showing AR directions only when a turn is coming may alleviate this problem.

Designing SV PNDs
Even though SV PNDs may look visually pleasing and efficient, our results indicate that using them might have a detrimental effect on visual attention and driving performance. SV PNDs may even require more visual attention than SPNDs, presumably because of the cognitive load of having to match the real world with SV images. Interface designers should exercise caution when taking the SV type of approach.

One way that SV PNDs might be useful for navigation is when they are being used by a passenger. Is there a way to make SV PNDs act more like a co-present human navigator and less like a distracting visual display? Should SV PNDs add visual aids or special verbal instructions that help drivers understand the scene displayed on the PND? We suggest that exploring such questions, perhaps by learning from human navigator behaviors, might allow us to design better SV PNDs.

CONCLUSION & FUTURE DIRECTION

This paper presents a thorough driving simulator-based comparison of two novel PNDs, HUD AR and HDD SV, using a standard PND as a reference. Using a simulated city environment in a high fidelity driving simulator, we found that an AR PND provided for more visual attention at the road ahead as compared to both SV PND and SPND. On average, when using an AR PND, participants spent about 5.7 sec and 4.2 sec more each minute looking at the road ahead in comparison to SV and SPND, respectively. Furthermore, we demonstrated how this increased visual attention was associated with better driving performance over both SV and SPNDs through peaks in cross-correlation functions. Importantly, participants' subjective assessment of the AR PND was largely positive.

SV PNDs use HDDs to show images of the world, which was associated with lower visual attention and worse driving performance. However, our results indicate that these effects are not solely due to using a HDD. After all, participants fared better with the HDD-based SPND. We suggest that a key issue is the difficulty in matching real world and SV images. Note that our real world and SV images were very similar, as the season, the weather and the time of day were identical in the two simulations that generated these images. In real-life scenarios these variables are likely to be different between the outside world and street view data. The effects of the differences on visual attention and driving performance are yet to be explored.

Finally, the commercial availability of HDD AR PNDs for cars is imminent. Already in 2005 Narzt et al. reported on creating a prototype system, although they did not conduct controlled user studies to evaluate its effects on visual attention and driving performance [14]. And at least one downloadable mobile application promises to turn select

mobile phones into HDD AR PNDs (Wikitude Drive [21]). These PNDs introduce the possibility of driving exclusively using a camera feed of the real world displayed on the HDD. As future research, we plan to explore how drivers would employ focal and ambient vision with such a PND, and how such visual behavior would relate to driving performance.

ACKNOWLEDGEMENTS

REFERENCES

1. Burnett, G. E. A Road-Based Evaluation of a Head-Up Display for Presenting Navigation Information. In *Proc. HCI Int'l Conference* (2003).

2. Fröhlich, P., Schatz, R., Leitner, P., Mantler, S., and Baldauf, M. Evaluating realistic visualizations for safety-related in-car information systems. In *Proc. CHI 2010*, ACM press (2010), 3847-3852.

3. Google Maps Navigation. http://www.google.com/mobile/navigation/.

4. Google Street View. http://maps.google.com/.

5. Hart, S. G. and Staveland, L. E. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*. P. A. Hancock and N. Meshkati, Eds. North Holland Press (1988), 239-250.

6. Horrey, W. J., Wickens, C. D., and Alexander, A. L. The Effects of Head-Up Display Clutter and In-Vehicle Display Separation on Concurrent Driving Performance. In *Proc. HFES 47th Annual Meeting* (2003).

7. Horrey, W. J., Wickens, C. D., and Consalus, K. P. Modeling Drivers' Visual Attention Allocation while Interacting with In-Vehicle Technologies. *Experimental Psychology: Applied 12*, (2006), 67-78.

8. Kim, S. and Dey, A. K. Simulated augmented reality windshield display as a cognitive mapping aid for elder driver navigation. In *Proc. CHI 2009*, ACM Press (2009), 133-142.

9. Kun, A. L., Paek, T., Medenica, Ž., Memarović, N., and Palinko, O. Glancing at personal navigation devices can affect driving: experimental results and design implications. In *Proc. AutomotiveUI '09*, ACM Press (2009), 129-136.

10. Lai, J., Cheng, K., Green, P., and Tsimhoni, O. On the Road and on the Web? Comprehension of synthetic and human speech while driving. In *Proc. SIGCHI 2001*, ACM Press (2001), 206-212.

11. Liu, Y.C. Effects of using head-up display in automobile context on attention demand and driving performance. *Displays 24*, 4-5 (2003), 157-165.

12. Lowe, S. Many Workers Have Long Commutes to Work. US Census Bureau Press Release. http://www.census.gov/Press-Release/www/releases/archives/american_community_survey_acs/004279.html. March 29, 2005. Retrieved 5/30/2009.

13. NASA TLX for Windows. http://humansystems.arc.nasa.gov/groups/TLX/computer.php.

14. Narzt, W., Kolb, D., Muller, R., and Hortner, H. Pervasive Information Acquisition for Mobile AR-Navigation Systems. In *Proc. WMCSA 2003*, IEEE Press (2003).

15. Porathe, T. and Prison, J. 2008. Design of human-map system interaction. In *Ext. Abstracts CHI 2008*, ACM Press (2008).

16. Steinfeld, A. and Green, P. Driver Responses to Navigation Information on Full-Windshield, Head-Up Displays. *Int'l Journal of Vehicle Design 19*, 2 (1998), 135-149.

17. Strayer, D. L. and Johnston, W. A. Driven to Distraction: Dual-Task Studies of Simulated Driving and Conversing on a Cellular Telephone. *Psychological Science 12*, 6 (2001), 462-466.

18. Tönnis, M. and Klinker, G. 2006. Effective control of a car driver's attention for visual and acoustic guidance towards the direction of imminent dangers. In *Proc. ISMAR 2006*, IEEE Computer Society (2006), 13-22.

19. Veit, R.R., Pyle, P., and McGowan, J.A. Ocean Warming and Long-Term Change in Pelagic Bird Abundance within the California Current System. *Marine Ecology Progress Series*, 139 (1996), 11-18.

20. Virtual Cable^{TM}. http://www.mvs.net/.

21. Wikitude Drive. http://www.wikitude.org/en/drive.