University of Utah
UNDERGRADUATE RESEARCH JOURNAL

EVALUATING THE EFFICACY OF SIMULATED AUGMENTED REALITY CUEING IN VIRTUAL REALITY

Emilee M. Wilson (Jeanine Stefanucci, Sarah Creem-Reghehr, Trafton Drew)
Department of Psychology

Humans possess an innate predisposition for *wayfinding* that requires them to acquire spatial representations of the environment they are traversing. We can define wayfinding as the cognitive process and experience of discovering, locating, or following a route in a given space (Symonds et al, 2017). Wayfinding is directly related to the acquisition of *spatial knowledge*—it is widely accepted that spatial knowledge is acquired through three main phases (Siegel & White, 1975). *Landmark knowledge* is obtained while initially traversing a new environment by observing and memorizing the appearance of landmarks along the way. Built off of this is *route knowledge*, which we learn from internalizing the sequential order of directional changes that encompass a route. Finally, we can develop *survey knowledge* that is essentially a ‘*cognitive map’* representation of the environment.

A completed spatial representation is formed when a navigator can imagine the entire environment independent of their current location, plan out routes in their head, and even create shortcuts to target locations (Siegel & White, 1975). In order to qualify as a sufficient ‘cognitive map’ of an environment, people should be able to attribute a specific landmark to its location from anywhere in the environment. Successfully acquiring adequate spatial knowledge is dependent on the landmark information that people initially encode about their surroundings.

When people utilize landmarks as environmental cues, they can be either distal or proximal cues (O’Keefe, & Nadel, 1978). *Distal cues* are large distinctive landmarks that are seen from a distance as one is traversing an environment—an example might be the mountains seen in the distance. *Proximal* *cues* provide positional information as a person is in motion and establish relative reference points in creating a spatial representation—an example might be a familiar storefront or a park.

Technology provides navigators with the ability to capitalize on these environmental cues that shape spatial knowledge. Spatial cognition research is utilizing the rising technology of *augmented reality (AR)* to study navigation (Martin et al, 2021). AR technology enhances reality by overlaying digital information on to the environment through a device such as a smartphone camera or AR goggles. Augmented overlay beacons could be used to enhance existing proximal landmarks in order to improve a navigator's spatial representation. The use of *beacon cues* as landmark reference points, along with additional AR navigational aids, could supplement spatial knowledge acquisition. Prior research has shown that using navigational aids, such as a compass, can guide wayfinding in a virtual environment (Winn et al., 2011). Due to some of the challenges of implementing AR cues accurately and reliably, augmented spatial cueing could be studied in an idealized immersive *virtual reality (VR)* environment that allows for the utility of cues to be studied more directly. Through capitalizing on well understood concepts of survey knowledge acquisition and active spatial learning, AR cueing could improve a navigator's wayfinding ability.

AR cueing could make wayfinding more efficient by offloading the attentional resources needed for a navigator to form a cognitive map. In order for cueing to be advantageous for efficient navigation, it must not significantly impact other essential cognitive functions. Prior work has shown how utilizing GPS as a navigational aid in laboratory studies can negatively affect a navigator's performance on environmental learning and spatial transformation tasks (Ruginski et al., 2019). Additionally, prior work has established that navigational aids can be advantageous for wayfinding, but create steep costs for spatial memory by dividing attention (Gardony et al., 2013). Analyzing the effects of AR navigational aids in a controlled VR environment will help us understand optimal learning situations by looking at the relative tradeoffs cueing might have on wayfinding and spatial learning.

Another prominent hurdle that must be overcome when implementing AR cues is consideration of how to improve wayfinding without significantly impacting situational awareness. We can look to the three-level model (Endsley, 1995) to define *situational awareness* as 1) perception of environmental elements, 2) comprehension of their meaning, and 3) projection of future status. With this in mind, AR cues must 1) provide spatial information that improves survey knowledge, 2) allow the navigator to offload attentional resources in order to make navigating faster and more efficient, 3) successfully allow for 1 and 2 without substantially decreasing situational awareness.

Prior work indicates that the increase of cognitive load is directly related to a decline in situational awareness (Hollands et al, 2019). We can define *cognitive load* as the information-processing demands imposed on a navigator when asked to perform tasks that require them to interpret and comprehend data (Hollands et al., 2019). Experimental psychology has an extensive history of utilizing a secondary task to assess cognitive workload by testing the spare attentional capacity allowed while a primary task is being completed (Norman & Bobrow, 1975; Wickens, 1984). For the purposes of testing the complex attentional allocation involved with driving, a secondary task called the *Detection Response Task (DRT)* was developed to measure cognitive load (International Standards Organization, 2012).

The DRT has been widely used (Strayer et al., 2015; Gardony et al., 2013), and this widespread use has contributed to it becoming an international standard in assessing selective attention while driving (International Organization for Standardization, 2016). A DRT produces a controlled sensory stimulus directly to a participant during their experimental session that elicits a specific response. For the purposes of this study, the DRT involves a vibrating apparatus that is secured to the participant’s clavicle and vibrates every 3-5 seconds which prompts the participant to respond via pressing a button secured to their pointer finger. The DRT provides a method for quantifying the cognitive load of the primary task, because as the difficulty of the task increases, signals will have slower response times or be missed altogether.

While traversing a new environment, navigators must be actively engaged with their environment in order to develop a substantive spatial representation. In a military context, navigators must be hyper-aware of their surroundings for safety and accuracy of data reported back to higher ranks. Traversing complex environments poses the risk of presenting information at a rate that exceeds the navigator's capacity to understand and interpret it. Successfully implementing AR navigational cues must improve navigation without adding more noise to the already engaged information processing demands of the navigator.

We ran an experiment to test the effects of AR cues on 1) navigational efficiency, 2) spatial memory and 3) cognitive load. We hypothesized that AR cues would positively impact navigational efficiency and supplement survey knowledge. Additionally, we hypothesized that the presence of AR cues would increase cognitive load and negatively impact situational awareness. This increase could impact the potential benefits of AR cueing for spatial memory.

Methods

Participants

15 University of Utah undergraduates (7 male, 8 female, aged x-y years, M=x) were recruited through the Psychology department's participant pool in exchange for course credit or recruited and compensated for their participation. 3 students in total withdrew from their first session due to simulator sickness (not counted in total participants).

Materials

Video Game Software

Immersive virtual reality cities were created using the video game software Unity (version 2018.2.12f1) displayed by the HTC VIVE Pro Eye system. The head-mounted display (HMD) allowed for 1440 x 1600-pixel resolution and 110-degree field of view (FOV) with an adjustable interpupillary distance (IPD) measurement. Participants moved through the virtual environment (VE) by manipulating the controller while seated in an armless swivel chair. Pressing down on the top of the controller's trackpad allowed for forward motion in the initial heading direction of the participant's FOV. Consistently pressing down on the trackpad allowed the participant to move their head and look around while maintaining forward motion in the same initial heading direction.

Actively learning while traversing a VE is accomplished through utilizing visual information and integrating podokinetic information (defined as bodily movement that matches visual stimuli in order to elicit proprioceptive information of displacement from the substrate; Chrastil & Warren, 2013). Physically rotating in the swivel chair allowed the participant to change their heading direction while they were not in motion and allowed for this podokinetic information to match their perceptual stimulation.

Data was collected by utilizing three different buttons on the controller for different tasks—the trigger was used to point at targets, the menu button was used to collect *'be on the lookout' (BOLO)* objects and holding the side button accessed a map. The participants were hooked up to a DRT (attached to the HMD) where a vibrating device taped to their clavicle invoked a sensory stimulus every 3-5 seconds and prompted them to respond via a button attached to their finger (seen in Figure 1). This DRT data was configured to align with the VR system's data to match their navigation performance with cognitive load information.

Figure 1

Example image of a participant secured in the HMD system and attached to the DRT

Virtual Cities

Two distinct 3x3 block virtual cities were designed to resemble downtown Salt Lake City, with each city equivalent in size and overall design. Sidewalks and roads were designed to emulate lengths and widths of blocks in Salt Lake City. Both *City 1* and *City 2* had the same distal landmarks modeled off the mountain ranges seen from downtown Salt Lake City. Both city conditions were populated with cars lining the roads and varying inanimate objects on the sidewalks. Each city contained a full parking lot and a park as large proximal landmarks but in different locations in each city.

City 1 contained a gazebo and a helicopter as distinct proximal landmarks (presented as L1 and L2 in the Figure 2.1 diagram), with a sports car, an ambulance, and a school bus as their targets (presented as T1, T2, and T3 in the Figure 2.1 diagram). City 2 contained a helicopter and a tipped over semi-truck as distinct proximal landmarks (presented as L1 and L2 in the Figure 2.2 diagram), with a cop car, a news van, and a garbage truck as their target landmarks (presented as T1, T2, and T3 in the Figure 2.2 diagram). Duffel bags, mailboxes, and guns (12 in each city) were randomly placed on the sidewalks of both cities and identified to the participants as BOLO objects on their clipboard (an example of a target and the BOLO objects are seen in Figure 3). Identifying and collecting these BOLO objects as they searched for their targets provided an additional measure for quantifying situational awareness while navigating. Additionally, both city conditions displayed the street names on the sidewalk at each intersection, along with reasonably sized road signs.

Figure 2.1
Example diagram of an aerial view of the City 1 VE

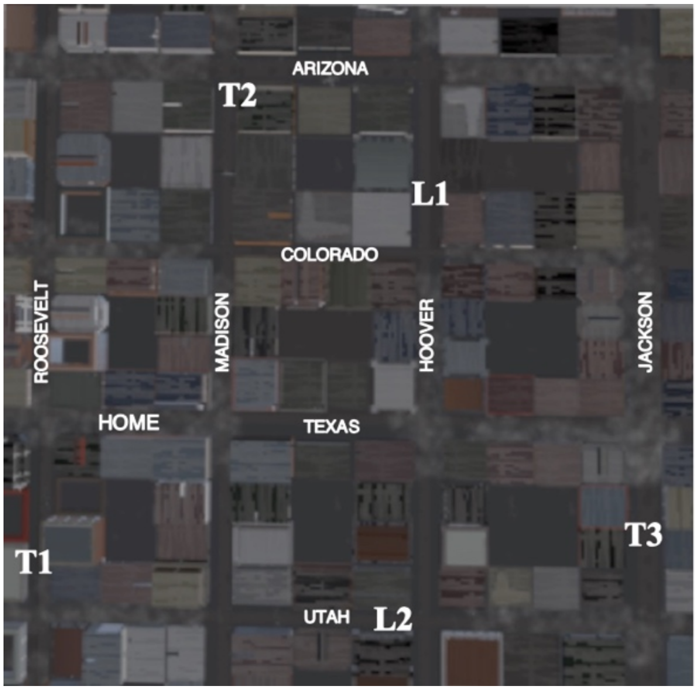


Figure 2.2

Example diagram of an aerial view of the City 2 VE

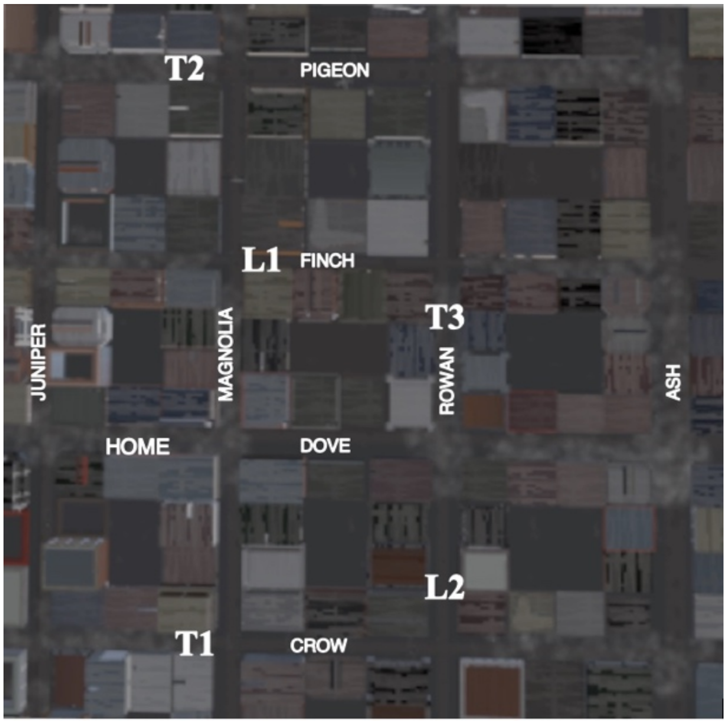
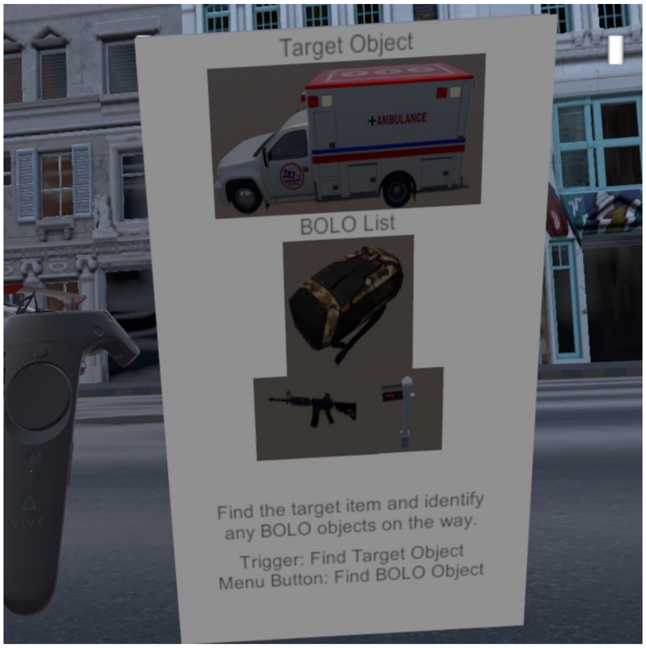


Figure 3

Example image of the controller’s clipboard displaying the target object and the list of BOLO objects.



Find the target item and identify any BOLO objects on the way.

Trigger: Find Target Object
Menu Button: Find BOLO Object

Navigational Aids

Two different conditions of the cities were presented—a *Cues ON condition* and a *Cues OFF condition*. Both conditions allowed the participant to access a *dynamic map* that showed an aerial view of the city with an arrow that tracked their current location in the environment. Figure 4.1 shows an example image of the controller’s clipboard displaying the dynamic map in the Cues ON condition. Access to the map was the only navigational aid that was consistent among both conditions. The Cues OFF condition only allowed for the participant to utilize the map, the existing landmarks and street names on the ground as navigational aids.

The Cues ON condition displayed superimposed neon colored *beacon cues* on top of the existing proximal landmarks (seen as L1 and L2 in Figure 2). Figure 4.2 provides an example image of these proximal landmark beacon cues in the experiment’s training environment. These augmented beacon cues were visible through the walls, suggesting that if the buildings were not there, they could walk directly to the landmark in the cue’s angular direction and make judgements about its distance by assessing the cue's size. The beacon cues acted as reference points to provide a relative spatial orientation to the navigator and were intended to aid survey knowledge acquisition and spatial memory.

Beacon cues appeared on targets after they were located, along with an augmented line attaching the participant to the located target as they moved around. Augmented above the attached line was the name of the located target and the distance (in meters) the participant was from the target (Figure 4.3 displays an example of this on City 1's first target). The Cues ON condition also overlaid the name of the road they were currently on, cardinal directions, and degrees away from that direction in the upper third of their FOV. Table 1 lists all of the navigational aids present in each condition.

Figure 4.1

Example image of the dynamic map in the Cues ON condition.

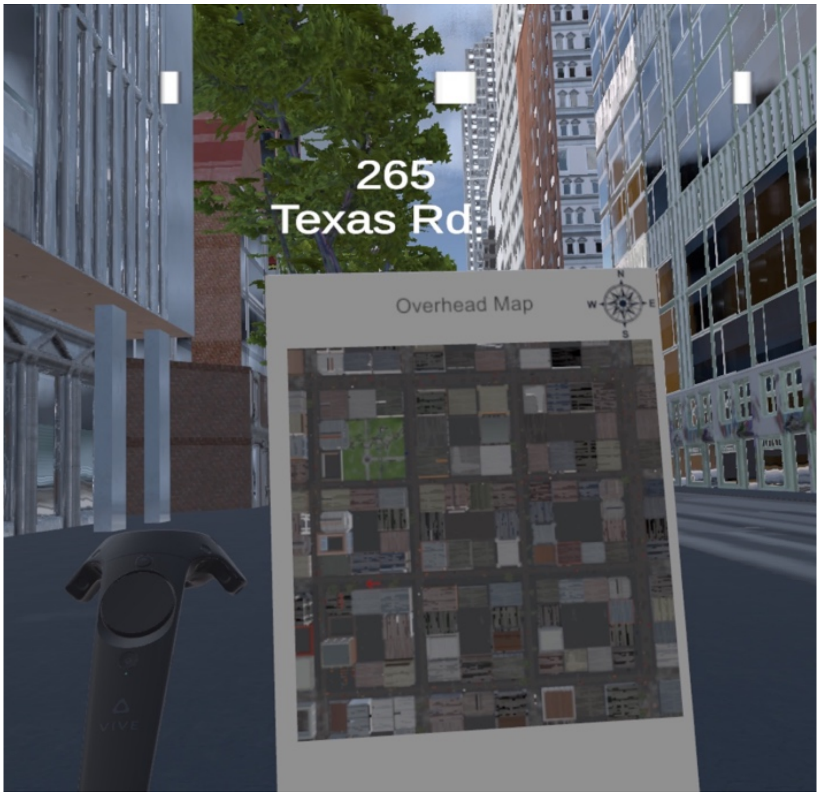


Figure 4.2

Example image of a landmark beacon cue in the Cues ON condition's training environment.

Figure 4.3

Example image of the target beacon cue, augmented line, distance to target, and target name.

Table 1
Table showing distribution of navigational aids in each condition.

Navigational Aids Present	*Cues ON Condition*	*Cues OFF Condition*
Dynamic Map	✓	✓
Distal Landmarks	✓	✓
Proximal Landmarks	✓	✓
Street Signs	✓	✓
Augmented Street Names	✓	
Landmark Beacons	✓	
Target Beacons	✓	
Augmented Line to Target	✓	
Augmented Distance to Target	✓	
Augmented Compass	✓	

Navigation Experience

Participants were asked to complete a brief questionnaire before they started the study. This included a Simulator Sickness Questionnaire (SSQ) that addressed their overall comfort levels and cognitive functions before and after their time spent in VR (Saredakis et al., 2020). The first part of the survey asked them to report on their prior video game experiences and provide self-assessments of their navigational strategies in everyday scenarios. This information was requested with the intention of analyzing the additional ease a seasoned video gamer might feel in the VR environment, along with the tendency to use traditional navigational strategies when one finds themselves in an unfamiliar environment.

Design and Procedure

The city version and cue condition were varied within-participants and counterbalanced for order. Participants were asked to return to complete their second session within 14 days of their first session. The cue condition was randomly assigned for their first session.

After giving consent, the participant was asked to complete the brief survey. Once completed, the participant measured their IPD, and then watched the experimenter demonstrate how to manipulate the controller. To prevent simulator sickness, thorough instruction was given on how they were to move in the virtual environment. They were hooked up to the DRT and placed in the HMD (set to accommodate their IPD measurement).

The participant was first placed in a training version of the virtual world where they were able to practice responding to the DRT while getting comfortable with the movement. They were instructed to look down at the clipboard on their controller to see where their targets would be displayed, along with the map when they held down the side button. The function of the map and the conditional navigational cues were thoroughly explained before they entered the actual city. Once the finish line was crossed in the training environment, the participant was instructed that they would be teleported to the starting location of the actual experiment.

Once in the actual city, the participant was instructed to look down at their clipboard to view the picture of their first target (to be collected with their controller's trigger) along with the pictures of the BOLO objects they were instructed to collect whenever noticed (to be collected with their controller's menu button). They were told to take note of their starting location because they would be asked to navigate back to *home* (a purple pedestal on the sidewalk) after they found the three targets. They were instructed to remember where the located targets were at because they would be tested on it at the end of the experiment. After reiterating that they were able to utilize the dynamic map, all conditional navigational aids, and any other navigational strategies, the participant was instructed to start searching for their first target on the main roads.

Once a target was found, they were encouraged to look at their clipboard to view the new picture of their next target. When all targets were located, the participant was instructed to return back to home without having to worry about collecting BOLO objects along the way. Once home, the DRT was turned off and the participant was instructed on how to perform a *point-to-origin task* to assess their spatial memory among the targets they found. The participant was prompted to point at each target's relative location by directing an augmented line from their controller to where they remember the target being. Once the participant pointed at all three targets from home, they were teleported to their subsequent target locations and asked to complete the task for the other targets and the home location. Pointing accuracy was measured by the degree of error from their augmented line to the actual location of the target. After completion of the pointing task and removing the HMD and DRT, the participant was asked to complete the second part of the SSQ. After their second session, the participant was debriefed and compensated for their time.

Results

Navigational Efficiency

Search step

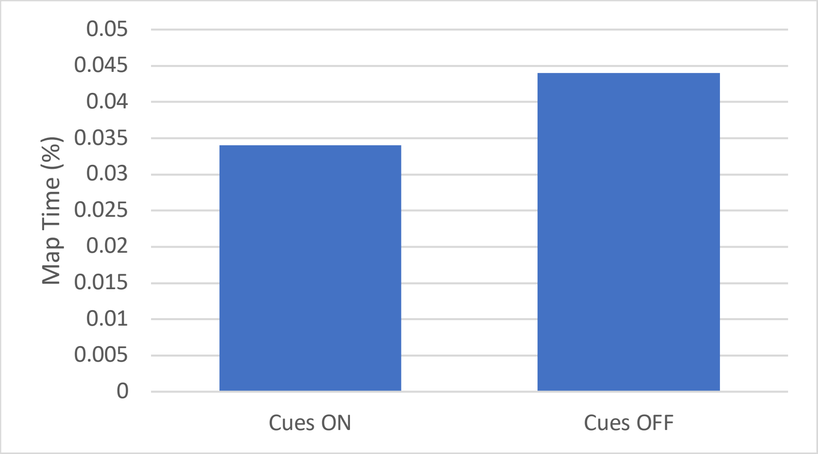
Search Time. The participant spent the duration of the experiment traveling from the initial home location to each of the three targets. The search step excludes the time spent traveling from the final target back to the home location. We ran a repeated-measures ANOVA and found a main effect of cues for the overall time spent on the search step ($F(1,13)=4.829$, $p=0.047$, $\eta^2=0.271$). Search time for the Cues ON condition ($M=1138.512$, $SD=581.2817$) was significantly longer than the search time in the Cues OFF condition ($M=1078.48$, $SD=367.5661$). This suggests that people spent more time searching the virtual environment for their targets when they had the presence of navigational cues.

Additionally, we found an interaction with order for search time with the Cues ON condition ($F(1,13)=61.374$, $p<0.001$, $\eta^2=0.825$). Participants who had the Cues ON condition as their first session (Cues ON: $M=1745.452$, $SD=381.3884$, Cues OFF: $M=790.3998$, $SD=122.5716$) took longer overall than those who had the Cues OFF condition first (Cues OFF: $M=1270.533$, $SD=350.6291$, Cues ON: $M=733.8839$, $SD=238.4723$). This order effect suggests that when experiencing the environment for the first time, having navigational cues present actually made searching for the targets take longer.

Map Time. We analyzed the percentage of time users spent looking at the map for Cues ON and Cues OFF with a repeated-measures ANOVA. We found a main effect of having the cues off for the time spent using the map ($F(1,13) = 6.776$, $p = 0.022$, $\eta\ 2 = 0.343$). As displayed in Figure 5, users spent significantly less time looking at the map in the Cues ON condition (M = 0.034, SD = 0.015) than in the Cues OFF condition (M = 0.044, SD = 0.018).

Figure 5

Map Time Data



Return Step

After the search step, participants were asked to find their way back home from their final target. We did not find any significant differences between the Cues ON and Cues OFF conditions when it came to this return step. There were no significant effects on return time, path length, map time or map calls.

Survey Spatial Knowledge

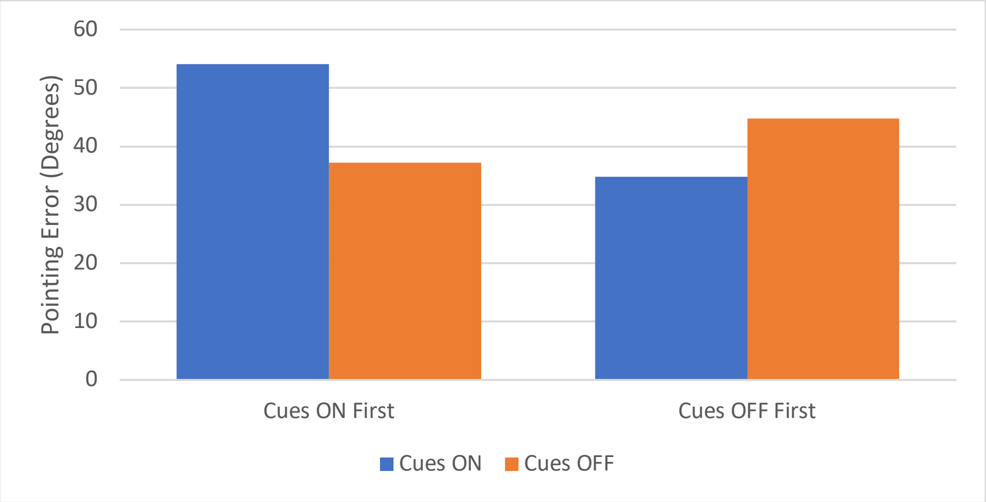
Judgement of Direction to Targets

We utilized point-to-origin tasks to assess the judgement of relative direction among each located target. Pointing error was measured as the difference of degrees between the direction of the target and the direction the participant pointed. The Wilcoxon's Signed Rank Test revealed no significant difference between the Cues ON and Cues OFF conditions (Z = -0.123, p = 0.902). We found an order interaction for overall pointing error when the Cues ON condition was in their first session (F(1,13) = 11.311, p = 0.001, η^2 = 0.060). As displayed in Figure 6, participants were more accurate with the cues off when they had the Cues ON condition first (Cues ON: M = 54.135, SD = 49.897, Cues OFF: M =

37.183, SD = 43.799). When they had the Cues OFF condition first, participants were more accurate with the cues on (Cues OFF: M = 44.796, SD = 42.454, Cues ON: M = 34.772, SD = 40.301).

Figure 6

Pointing Error Data



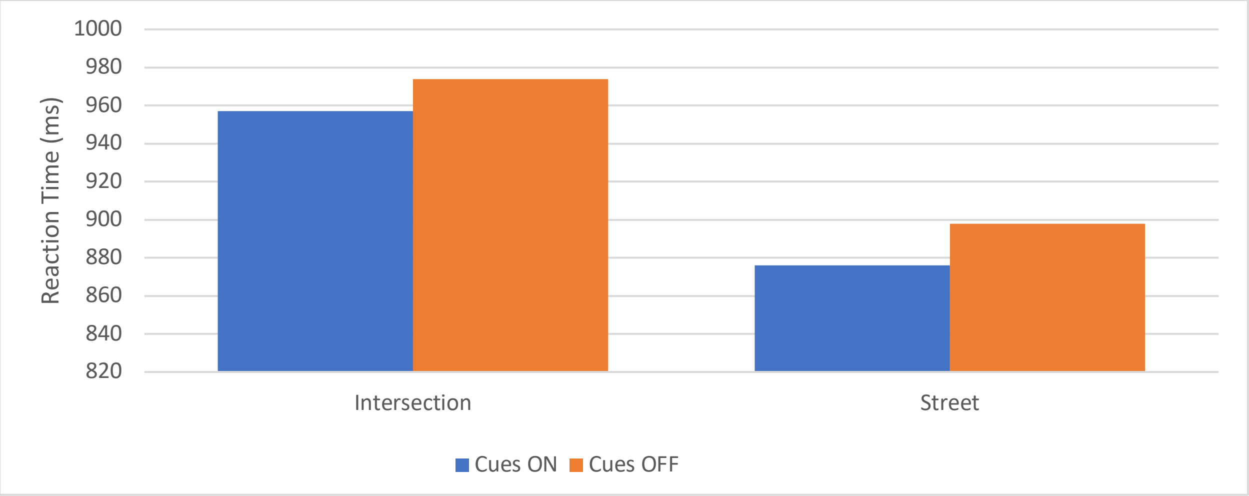
Situational Awareness

Cognitive Load

We analyzed response times measured by the DRT using a Non-Parametric Related-Samples Friedman's test between cues and location, and we evaluated any significant effects via post-hoc pairwise comparisons using the Wilcoxon's Signed Ranks test. We found significantly different reaction times across cues and location (χ 2 (3) = 478.667, p < 0.001). As seen in Figure 7, DRT response times were significantly faster in the Cues ON condition compared to the Cues OFF condition. Additionally, we analyzed the response times when the participant was in a street versus an intersection. As seen in Figure 7, response times were faster when participants were in the street compared to when participants were in intersections.

Figure 7

Cognitive Load (DRT) Data



Discussion

In this study, participants navigated immersive virtual cities in two different sessions to assess the impact of AR cues on navigational ability, survey spatial knowledge and situational awareness. Participants were asked to locate three targets, collect BOLO objects, and eventually navigate back to their starting location. The time spent searching and returning home were analyzed for navigational efficiency. The navigational cues did not improve the overall search time—in fact, participants spent significantly longer searching for targets while in the Cues ON condition. This was supported by prior research that suggests visual aids do not improve a 2D search (Drew & Williams, 2017).

We hypothesize that fixed beacon cues may have been guiding participants instead of acting as a reference point for gaining landmark knowledge. This is consistent with prior work that showed how salient cues that are supposed to aid search can actually pull attention from important environmental information (Cunnigham et al, 2016; Drew et al, 2012). In future studies we will integrate eye-tracking hardware that allows us to incorporate gaze-contingent beacon cues that could alleviate this problem. Additionally, future studies could incorporate augmented 'breadcrumb' cues on roads to guide participants more directly while still enhancing landmarks.

At the end of the experiment, participants completed point-to-origin tasks to assess their survey spatial knowledge. They indicated the relative direction of each target from the home location and then teleported to another target in the city, where they were asked to point at the other targets and their prior home location. Survey knowledge was assessed by measuring their degree of pointing error from the actual target. Our analysis showed no difference in survey spatial knowledge between the Cues ON and Cues OFF conditions, implying that beacon cues were not beneficial in supplementing landmark knowledge.

We hypothesize that due to the nature of the dynamic map, environmental cues were overlooked because participants relied heavily on the map for navigation. Because the map could indicate the user's position as they moved around the world, they primarily learned the environment from a bird's eye view. This encouraged the development of an allocentric cognitive representation of their environment instead of a spatial representation that landmarks could provide. Without utilizing the beacon cues for spatial learning, the cues could not play a significant role in the development of survey knowledge. Future work could provide participants with a static map to further investigate the reliance on the dynamic feature and how that could affect spatial learning and survey knowledge.

Interestingly, navigational cues reduced the amount of time that participants spent consulting the map. This suggests that AR cues could be a useful supplement to utilizing a dynamic map. It is possible that the presence of the AR cues and the dynamic map allowed the participant to integrate egocentric and allocentric spatial information to help them offload attentional resources while searching for targets. Future work could investigate this relationship further to assess their role on acquiring survey knowledge more efficiently.

While navigating the city environment, participants were to respond to the DRT stimulus throughout the duration of the experiment. The presence of navigational cues significantly reduced the response time for the DRT. This suggests that AR cues allowed participants to offload cognitive resources due to the information being readily available for them in the environment. For example, keeping track of the road name or the cardinal direction was not necessary when it was augmented in the field of view. This cognitive offloading allowed them to respond more quickly to the DRT and implies that augmented cues do not increase cognitive load. If AR cues do not overly tax cognitive resources, then there are many implications for the future use of AR cues without significantly impacting situational awareness. Future work should further investigate the role of AR cueing for assisted navigation and training. Additionally, significant results from our DRT data suggests that future navigational work can integrate the DRT as a legitimate methodology for assessing cognitive load.

References

Chrastil, E.R., Warren, W.H., 2013. Active and passive spatial learning in human navigation: acquisition of survey knowledge. J. Exp. Psychol. Learn. Mem. Cogn. 39, 1520–37.

Drew, T., Cunnigham, C., & Wolfe, J. M. (2012). When and why might a Computer-aided Detect6ion (CAD) system interfere with visual search? An eye-tracking study. *Academic Radiology, 19*(10), 1260-1267.

Drew, T., & Williams, L. H. (2017). Simple eye-movement feedback during visual search is not helpful. Cognitive Research: Principles and Implications, 2(1), 1-8.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. Human factors, 37(1), 32-64.

Gardony, A. L., Brunyé, T. T., Mahoney, C. R., & Taylor, H. A. (2013). How navigational aids impair spatial memory: Evidence for divided attention AU—Gardony. *Aaron L. Spatial Cognition & Computation, 13*(4), 319-350.

Gardony, A. L., Brunyé, T. T., Mahoney, C. R., & Taylor, H. A. (2013). How navigational aids impair spatial memory: Evidence for divided attention AU—Gardony. *Aaron L. Spatial Cognition & Computation, 13*(4), 319-350.

Hollands, J. G., Spivak, T., & Kramkowski, E. W. (2019). Cognitive load and situation awareness for soldiers: Effects of message presentation rate and sensory modality. *Human Factors,* 0018720819825803

International Organization for Standardization. (2016). Road vehicles: Transport information and control systems—Detection-response task (DRT) for assessing selective attention in driving (ISO 17488:2016). Geneva: Author.

International Standards Organization. (2012). Road vehicles: Transport information and control systems. Detection Response Task (DRT) for assessing selective attention in driving (ISO TC 22 SC 13 N17488). Working draft under development by Working Group 8 of ISO TC22, SC 13.

Martin, A., Cheriyan, J., Ganesh, J., Sebastian, J., & V, J. (2021). Indoor Navigation using Augmented Reality. *EAI Endorsed Transactions on Creative Technologies*, *8*(26), 168718. https://doi.org/10.4108/eai.17-2-2021.168718

Norman, D., Bobrow, D. (1975). On data limited and resource limited processing. Cogntive Psychology, 7, 44-60.

O’Keefe, J., & Nadel, L. (1978). The hippocampus as a cognitive map. Oxford: Clarendon Press.

Ruginski, I. T., Creem-Regehr, S. H., Stefanucci, J. K., & Cashdan, E. (2019). GPS use negatively affects environmental learning through spatial transformation abilities. Journal of Environmental Psychology, 64, 12-20.

Saredakis, Dimitrios, Ancret Szpak, Brandon Birckhead, Hannah A D Keage, Albert Rizzo, and Tobias Loetscher. "Factors Associated With Virtual Reality Sickness in Head-Mounted Displays: A Systematic Review and Meta-Analysis." *Frontiers in Human Neuroscience* 14 (2020): 96. Web.

Siegel, A. W., & White, S. H. (1975). The Development of Spatial Representations of Large-Scale Environments. Advances in Child Development and Behavior, 10, 9-55.

Strayer, D. L., Turrill, J., Cooper, J. M., Coleman, J. R., Medeiros-Ward, N., Biondi, F. (2015). Assessing cognitive distraction in the automobile. Human Factors, 57, 1300-1324.

Symonds, P., Brown, D. H. K., & Lo Iacono, V. (2017). Exploring an Absent Presence: Wayfinding as an Embodied Sociocultural. Experience. Sociological Research Online, 22(1), 48–67. https://doi.org/10.5153/sro.4185

Wickens, C. D. (1984) Processing resources in attention. In Parasuraman, R., Davies, D. R. (Eds), Varieities of attention (pp, 63-101). Orlando, FL: Academic.

Winn, B. M., Peng, W., & Pfeiffer, K. (2011, November). Player guiding in an active video game. In 2011 IEEE International Games Innovation Conference (IGIC) (pp. 107-108). IEEE.