On the Go with AR: Attention to Virtual and Physical Targets while Varying Augmentation Density

You-Jin Kim Texas A&M University College Station, Texas, USA Media Arts and Technology University of California, Santa Barbara Santa Barbara, California, USA yujnkm@gmail.com

Tom Bullock Psychological & Brain Sciences University of California, Santa Barbara Santa Barbara, California, USA tombullock@ucsb.edu Radha Kumaran Computer Science University of California, Santa Barbara Santa Barbara, California, USA rkumaran@ucsb.edu

Barry Giesbrecht Psychological & Brain Sciences University of California, Santa Barbara Santa Barbara, California, USA giesbrecht@ucsb.edu Jingjing Luo Media Arts and Technology University of California, Santa Barbara Santa Barbara, California, USA jingjingluo@ucsb.edu

Tobias Höllerer Computer Science University of California, Santa Barbara Santa Barbara, California, USA holl@cs.ucsb.edu



Figure 1: A participant searching for virtual and physical targets in our augmented environment. In this example an image of the participant is superimposed onto a mixed-reality capture of the scene for illustration purposes.

This work is licensed under a Creative Commons Attribution 4.0 International License. *CHI '25, Yokohama, Japan* © 2025 Copyright held by the owner/author(s). ACM ISBN 979-8-4007-1394-1/25/04 https://doi.org/10.1145/3706598.3714289

Abstract

Augmented reality is projected to be a primary mode of information consumption on the go, seamlessly integrating virtual content into the physical world. However, the potential perceptual demands of viewing virtual annotations while navigating a physical environment could impact user efficacy and safety, and the implications of these demands are not well understood. Here, we investigate the impact of virtual path guidance and augmentation density (visual clutter) on search performance and memory. Participants walked along a predefined path, searching for physical or virtual items. They experienced two levels of augmentation density, and either walked freely or with enforced speed and path guidance. Augmentation density impacted behavior and reduced awareness of uncommon objects in the environment. Analysis of search task performance and post-experiment item recall revealed differing attention to physical and virtual objects. On the basis of these findings we outline considerations for AR apps designed for use on the go.

CCS Concepts

Human-centered computing → Empirical studies in HCI;
Computing methodologies → Mixed / augmented reality;
Perception.

Keywords

Mobile Augmented Reality, Extended Reality, Mixed Reality, Perception, Behavior

ACM Reference Format:

You-Jin Kim, Radha Kumaran, Jingjing Luo, Tom Bullock, Barry Giesbrecht, and Tobias Höllerer. 2025. On the Go with AR: Attention to Virtual and Physical Targets while Varying Augmentation Density. In *CHI Conference* on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan. ACM, New York, NY, USA, 16 pages. https://doi.org/10. 1145/3706598.3714289

1 Introduction

Augmented Reality (AR) headsets, such as the Microsoft HoloLens 2, Magic Leap 2 and Xreal Ultra blend virtual content seamlessly with the physical world using sensors and advanced optics and computer vision. AR is projected to be a primary mode of information consumption in the future, thanks to its ability to integrate virtual content into the user's physical surroundings. These unique capabilities of AR could be particularly useful for information consumption on the go, when more powerful but bulkier devices such as PCs, laptops, or tablets are not convenient [29].

AR has demonstrated advantages in supporting search and navigation [3], and could prove useful in scenarios such as emergency search and rescue [52, 84], exploratory learning [25, 31] or tourism [18]. Always-on AR smart glasses are also expected to be widely used for personal everyday use in the near future, projecting virtual content into the physical world as users navigate larger environments. Given that such AR use would make information more conveniently accessible than through smartphone use while walking, which has its own increasingly demonstrated cognitive costs [43, 49, 55, 66], we need to understand the potential impact of AR consumption on walking users' behavior and cognitive processes before designing head-worn AR applications that could support such tasks as navigation and wayfinding in unfamiliar environments, or applications such as entertainment, education, communication, corporate duties, or personal information management during common daily walks [8].

The introduction of virtual content into environments that are already crowded could lead to cognitive or perceptual overload, potentially distracting from the user's primary tasks and even jeopardizing their safety by reducing their awareness of the physical surroundings. There is a limited understanding of the cognitive phenomena that are important, and most impacted, when users perform AR tasks while walking to navigate an environment. It is therefore crucial that we better understand user responses to increasingly dense augmented environments. Accordingly, the first goal of this paper is to investigate the impact of different levels of environment augmentation density on user performance in a mobile AR search task. Environment search is a suitable generic but representative task for several application scenarios that require some attention to the physical world while benefiting from virtual augmentations.

The main strength of AR is its ability to situate virtual content in the real world, with the intention of creating a blended space that users perceive as a single environment. However, even with the current state-of-the-art technology, there are still visible differences between the real environment and renderings of virtual content. Significant technological advances are still required before the distinctions between virtual and physical worlds can be eliminated, and indiscernibility of real and virtual items may not even be desirable in all applications. In AR-guided search tasks, users must simultaneously focus on both real and virtual environments, so it becomes essential to understand how users attend to and interact with these blended spaces. Therefore, the second goal of this paper is to better understand how users interact with goal-relevant objects in both the virtual and physical worlds, by having them search for both physical and virtual items.

The use of AR in real-world environments has been shown to impair user safety, due to the potential of virtual content obscuring real-world objects, and reduced situational awareness [40, 41]. Future AR on-the-go applications are likely to be used in much less controlled environments than current experiments, and may also display various notifications while in use, potentially distracting users as they navigate their environment. Users may want to use AR on a variety of both destination-driven and more meandering walks. AR may even be the interface of choice guiding a user along a path, similar to how Live View on Google Maps provides navigational guidance on AR-enabled smartphones [26]. Whatever AR guidance is provided, be it for navigation or exploration purposes, ensuring the walker's safety and the safety of others in the vicinity will be essential.

Always-on AR annotations or guidance mechanisms may increase cognitive and perceptual demands on the user, requiring them to dual-task and potentially reducing their situational awareness of the physical world, hence the impact of these mechanisms on behavior must be studied. Our third goal is to investigate the impact of a visual guide that controls the user's walking path while they perform a task that requires scanning the environment for relevant items.

Here we designed a search and classification task in which participants searched the environment for semi-spherical "gems", and classified gems into two categories based on the presence (or absence) of surface markings on the gem that required closer inspection to discern. To address our first goal, we manipulated augmentation density in the environment, such that participants performed the task in both lightly augmented and heavily augmented environments. To address our second goal, we had participants search for both physical and virtual gems, which were matched for perceptual characteristics as closely as possible. Our study is the first to include both real and virtual targets in such a search task, to the best of our knowledge. To address our third goal, in half the experimental trials we introduced a "spotlight" ring annotation at approximately the user's waist height that moved along a predefined path through the environment. Participants were required to stay within the bounds of the ring when it was present.

Our key findings suggest that

- environmental clutter significantly impacts user perception and task performance, as evidenced by
 - increased head rotation in the presence of high clutter and
 - reduced awareness of a highly conspicuous non-target object (a model of Godzilla).
- There appears to be a divide in attention between the physical and virtual scenes, with physical targets being more easily detected.
- In contrast to prior work that reported for AR search tasks involving purely virtual targets that physical scene objects were far less memorable in a recall task than virtual objects, a more balanced recall performance among virtual and physical objects in this task provides new evidence regarding the possible causes of this effect.

Together, these results lead to valuable insights for the future development of AR experiences for walking users.

2 Related Work

Our present study is inspired and informed by three broad areas of relevant literature. In this section, we summarize key findings from studies that explore perception and attention in AR, the blending of physical and virtual worlds in AR, and search for goal-relevant items in AR.

2.1 Perception and Attention in Mixed Reality Environments

In recent years, standalone AR headsets have become more powerful and capable, enabling them to overlay increasingly rich and complex layers of augmentation into the physical environment. However, the current technological constraints of AR headsets prevent them from projecting truly photo-realistic virtual information into the physical world, meaning that the user can always still clearly differentiate between real and virtual, thus preventing seamless blending of the two worlds. Optical see-through head-mounted displays (OST-HMD), such as the Hololens 2 used in this study, have a restricted field of view [62], poor color resolution [51, 63], do not effectively render black or dark annotations without carefully crafted halo or highlighting techniques [53] and cannot function in brightly lit environments due to their use of an additive light model [20]. These lighting and display limitations can induce eye fatigue [23, 51, 57] and can impact color perception [16, 24] and depth perception [1, 71, 72] in the user. Fusion of virtual information with the physical world, especially when the virtual information is perceptually inaccurate or degraded, raises concerns about the

potential impact of AR technologies on how humans perceive and attend to information in both worlds. To better understand these potential impacts, we draw upon modern theories of attention, visual search, and investigations into attention within mixed reality environments. Given our paper's central theme, we narrow our discussion to selective attention and visual search. Recent study [8], which explores AR interfaces in real-world walking scenarios, provides valuable insights, emphasizing the importance of our research into how AR influences perception and attention.

Search of a natural visual scene requires that we deploy selective attention [13], because we cannot process everything in the scene all at once. Contemporary models of selective attention assert that selection can be guided by a wide range of different factors including but not limited to current goals, physical salience of items in the scene and prior search history [2, 69, 82]. In other words, attention can be deployed voluntarily across a room to scan for something specific like a set of keys; captured involuntarily by a sudden movement; and drawn towards items associated with a positive outcome. The perceptual difficulty of a primary task can also determine whether an individual will become aware of task-irrelevant items in a visual scene, with higher task demands leading to reduced competition from competing distractor items [44, 45]. Furthermore, in the absence of focused attention we can fail to detect highly salient and unexpected objects that appear in our environment, as demonstrated in the well-known "Invisible Gorilla" study [70]. Together, these models and examples from traditional lab-based cognitive studies confirm that attention is highly susceptible to a range of different goal and environment related influences.

Both voluntary and involuntary modes of attentional control can potentially be disrupted when a user is placed in a mixed reality environment. For example, in dense augmented environments, newly entering goal-relevant target items can be missed due to occlusion by virtual "clutter" in the scene [19, 36, 77], and this clutter can impair user interactions with both physical and virtual objects [11, 65]. Furthermore, when compared with user experiences in physical environments, AR can have a negative impact on users' attention [4, 21], particularly during the decision-making phase of a driving task [21]. Studies on the use of mixed reality in an outdoor wide-area search task also indicate that attention to virtual objects may reduce situational awareness of objects in the physical environment [34, 41] (see Section 2.3 for discussion of wide-area search).

In summary, this related work suggests that AR can have a profound impact on user's attention. The three main goals of this study are designed to assess aspects of AR that may impact user attention: augmentation density, search for both physical and virtual task-relevant items in the environment (see next section), and the additional dual-tasking demands of navigational guidance on attention.

2.2 Physical and Virtual Objects in AR Experiences

As Mixed Reality (MR) technologies become more advanced (e.g. higher resolution, computing power), tangible user interface systems, which integrate virtual content into physical worlds [33],

such as ARTable [56], have blurred the lines between the physical and virtual. In Cross-Reality (CR) environments, transitional navigational techniques such as MagicBook [5], which transport users into the virtual environment of book pages, and the creation of virtual environments from live video feed of the physical environment [74] are examples of how there is increased interest in blending physical and virtual worlds and transitioning between them. Studies on this relationship have measured traditional approaches like "step into" or "reach into" as well as newer methods like the "blended space," where users can see the source and destination environment simultaneously [14].

Mixed reality environments have been explored in various studies and projects [41, 54, 61, 76]. Notable examples of these enhanced experiences include projects such as Remixed Reality, Reality Distortion Room, and Room2Room [37, 50, 58]. These projects utilize the physical layout of a room and employ spatial augmented reality to extend user interaction and experience within these blended spaces, where interaction is expanded through imagination. Additionally, the Blended Whiteboard project [27] suggests interacting with a physical object, in this case, a whiteboard, from each user's perspective, allowing them to experience seamless interaction as if they are in the same location, even when they are not.

Assessing the relationship between physical and virtual objects, studies have examined the ways virtual objects adapt to shifting physical environments, such as those modeling illusionary episodes [10, 32, 74]. Separate from the environment, virtual objects can also be manipulated by going through recoloring, teleportation, moving, copying, scaling in real time, and customization such as in MineXR [12, 50].

With increasingly complex blended spaces, virtual augmentations such as AR cues and adaptive clustering have been implemented to direct user attention to target objects and simplify distracting environments respectively [60, 75]. AR cues given from a first-person view (FPV) have particularly shown increased ease in following instructions and performing tasks [35, 46, 48]. Direct manipulation of objects also allows users to better navigate spatial recordings of AR environments.

Using current augmented reality (AR) technologies, users perceive two worlds due to incongruencies of display, rendering, blend between physical and virtual content, and the ability of AR to modify and create environments that do not exist in the real world [9, 50, 78, 79]. In the real world, people experience source confusion, which is when memories are attributed to the wrong source (e.g. confusing pictures seen with real experiences) [7]. Since virtual objects have greater media richness than physical objects, researchers have speculated that VR could provide more memory suggestions in ways that are autobiographical (real life) rather than episodic (shallow, isolated events) [7, 67, 68].

Previous work has shown that the phenomenon of attentional tunneling, which is where users focus more on virtual objects rather than physical objects [80], is affected by the number of virtual tasks [73], accurate cues to complete tasks [83], and level of engagement the tasks require [17].

As the second goal of our study, we explore the phenomena of attention tunneling and source confusion during natural walking while exploring an AR environment. We compare attention to physical and virtual target objects during a search task, as well as investigating attention tunnelling and source confusion through a memory task involving task-goal irrelevant objects in the environment (which participants are only informed of after the experiment).

2.3 Search and Mobile Information Browsing in Augmented Reality

AR's ability to overlay digital information into the physical world unlocks many potential use cases for assisting with search tasks during various different scenarios. For example, emergency search and rescue, using always-on AR to explore a new place, or locating a target in a crowded space. For this reason, search has been studied in mixed reality environments, with two broad types of search tasks often used - informed search tasks, where participants are guided to the location(s) of the target(s) using virtual cues [6, 37–39, 64], or uninformed search tasks, where participants are not given any information about the location of the target(s) [47]. We note that in the present study participants were not cued to the locations of the targets, because the focus of the experiment was to observe natural attention patterns in the different experimental conditions.

Much of the literature on search in XR does not involve any natural locomotion, likely due to technological limitations. Given the portability of AR and enormous future potential for it to be used on the go, the lack of locomotive AR studies represents a critical gap in the literature. Initial studies have explored mobile information browsing via window-based AR workspaces that are either oriented with walls in the environment or following the user in various ways [8, 42], but the impact of 3D AR scene density has not yet been formally explored. Recent advances have enabled studies on users navigating and searching larger real-world environments while wearing mixed reality headsets, scouting for virtual target items under different lighting conditions [34] and with the assistance of different navigational aids [41]. The present study makes several important novel contributions to our understanding of locomotionbased search in AR, by assessing the impact of augmentation density, physical and virtual items and path guidance in a locomotion-based search task.

3 Experimental Design

In the present study, participants performed a search task, where they searched for 12 target gems in an indoor area (208.54 m2, or 2,244 ft2) augmented with virtual and physical furniture, wall augmentations and other virtual objects.

3.1 Task

Participants performed two tasks in the experiment - a search and classification task (8 trials total) and a surprise object recall task designed to probe memory for goal-irrelevant objects in the environment, which was administered after all the search task trials were completed.

In each trial, participants had to walk through the L-shaped environment and classify each gem they found into one of two categories - marked or unmarked (Figure 2) - by clicking buttons on a handheld controller. If a participant classified the gem and their response was registered by the system, a small yellow sphere appeared above the gem and remained visible throughout the trial. On the Go with AR



Figure 2: Four classes of gems encountered by participants. From left to right: Unmarked-Physical, Unmarked-Virtual, Marked-Physical, Marked-Virtual. All gems emit purple light, and the virtual gems were designed to closely resemble the physical gems.



Figure 3: Participants experienced two levels of virtual augmentation (environmental clutter) in the hallway the experiment was conducted in. Images depict the environment with low (a) and high (b) levels of augmentation.

Participants were instructed to walk through the environment in a single direction, ensuring that they explored new areas without retracing their steps or revisiting any previously encountered sections of the environment. On trials when path guidance (via a spotlight ring annotation moving above the floor) was provided, participants were instructed to maintain a consistent position within the ring throughout the trial as they performed the task.

At the end of all trials participants answered an object memory questionnaire, where they were given a list of twelve objects (six absent and six present in the environment). They had to indicate which objects were actually present in the environment, and if present, whether they were physical or virtual.

3.2 Design

In keeping with our three key goals, we manipulated augmentation density, gem type and path guidance in a locomotion-based search task. Specifically:

Augmentation Density. Participants performed the search task under two different augmentation levels: low augmentation (Figure 3a, physical environment with few virtual additions) and high augmentation (Figure 3b, approximately equal amounts of physical and virtual content). For the purposes of this study, physical content was defined to include both objects in the environment, as well as physical structures such as walls and pillars. The ceiling and



Figure 4: Environment navigation. (a) A depiction of a participant walking within the "spotlight" ring - a virtual green ring that was used to guide participants' path through the study environment in the guided conditions. (b) An example of the spotlight from a participant's perspective (captured directly from the headset). (c,d) Heatmaps depicting the locomotive activity of all participants during spotlight-present (c) and spotlight absent (d) trials. Higher temperatures indicate increased dwell times.

floor, which were intentionally excluded from full augmentation, were not considered. In the low augmentation scenario, no walls or objects underwent skin wrap alterations that would change their appearance. Conversely, in the high augmentation scenario, approximately two-thirds (66%) of the wall surfaces were augmented throughout the environment. Physical objects, such as stools, were given virtual skin wraps to enhance their visual appeal, while still retaining their functionality and visibility with amplified virtual surfaces. This was done intentionally to create a distinct difference in the appearance of the two conditions. As the future of AR leans towards more extensive augmentation, we wanted to explore the effects of a highly augmented, nearly virtual environment where the physical layout and content is still preserved.

Gem Type. In each trial, participants searched for six virtual and six physical gems scattered throughout the environment, which were matched as closely as possible in terms of their physical characteristics (Figure 2). Participants were not explicitly asked to detect whether the gems were physical or virtual, but gem type (virtuality) was included as a factor in our analyses.

Path Guidance. In half of the trials, participants were guided on a set trajectory determined by the movement of a green ring ("spotlight") that moved at a steady pace of 0.92 m/s at 0.75m height above the floor from the entrance of the corridor to the back exit,

CHI '25, April 26-May 01, 2025, Yokohama, Japan



Figure 5: Twenty-five different light renderings were tested to create virtual gems that most resembled the physical gems. Eight of the twenty-five renderings are shown, with the second gem (from left) in the top row being the most accurate and thus chosen for the study. The image shown was captured directly from Unity, not from the headset.

taking 60 seconds in total to complete the journey (Figure 4a, 4b). In the other half of the trials the spotlight was absent, and participants were able to walk through the environment at their own pace, with the only constraint being that they were not permitted to retrace their steps. While effective in providing a predefined path, as shown in (Figure 4c), maintaining position within the spotlight placed an extra burden on the user's attention, essentially creating a dual task that served as an additional challenge.

We also manipulated the presence of specific task-irrelevant objects in the environment, with the goal of probing participants' recall for these items once the search task was complete. Six beachthemed objects were placed throughout the environment in each trial to assess user attention and memory, including three physical ones (mini ice box, camping chair, beach umbrella) and three virtual (coconut, hammock, grill). The positions of the objects were not changed between trials or conditions, and they were distributed throughout the environment as evenly as possible while still ensuring a safe walking path. After completing the eight search trials, participants were given a surprise object recall test (see Procedure). Furthermore, we included a highly salient virtual object (a large "Godzilla" figure, standing 1.9m tall. See Figure 9) in every trial, with the specific goal of probing participants' awareness for an object that was incongruent with the surroundings. In the absence of focused attention, it is possible for individuals to fail to become aware of highly salient items in the environment - a phenomenon known as inattentional blindness [70]. An additional goal of of our study was to test the impact of AR on susceptibility to inattentional blindness. Probing participants' awareness of this highly salient item at the end of each trial meant that we could assess if/when they first noticed the item, and if they did notice it, under what conditions of augmentation and path guidance. Godzilla was placed in the same position during all trials and conditions, in a passageway branching off from the main environment (as seen in Figure 8). Godzilla was occluded by pillars and walls from some angles, but was positioned right behind a stool that sometimes held a target, making it highly likely that Godzilla fell within the participant's field of view during the study.



Figure 6: Four different study conditions. Examples from the full factorial combination of augmentation (low, high) and spotlight (absent, present), which creates four distinct conditions. Condition order was fully counterbalanced, and participants completed two trials of each condition, resulting in a total of eight trials for each participant. The participant's perspective at the start of a trial is depicted in each of the images.

In summary, each trial was a factorial combination of augmentation density (low, high) and path guidance (spotlight ring present, spotlight ring absent) (Figure 6). Participants searched for six physical and six virtual gems in each trial. In each of the eight trials participants encountered a distinct arrangement of 12 gems, with the positions randomly chosen from 21 possible locations. The eight gem arrangements (curated to ensure a balanced distribution of gems throughout the scene) were randomized across conditions for all participants using a set of Latin Square permutations. Participants completed four conditions in total, and each condition was repeated twice (eight trials total). Condition order was fully counterbalanced between participants.

3.3 Pilot Considerations and Formative Experiments

We conducted several pilot experiments to determine parameters of the experiment design such as the appearance of the virtual gems and pace of the gliding ring for the spotlight.

3.3.1 Gem Design and Rendering. We tested various types of lights for the environment, including Point Lights, Spot Lights, Directional Lights, and Area Lights. Area Lights were the most effective due to their ability to illuminate objects from multiple directions and create softer, more subtle shading, making them ideal for blending physical and virtual gems without casting harsh shadows. To capture the subtle light bleed seen in physical gems, we used the MRTK Hover Light Shader, which mimics the effect of an area light and light bleed that may cast to the surface. The rendering of six virtual gems in this fashion did not negatively impact frame rate on the HoloLens 2.

We created 25 versions of the virtual gems and compared them with physical gems using an AR headset to determine the best match, eventually resulting in a remarkably similar appearance.

Kim et al.

On the Go with AR

Some of the candidate rendering models are shown in Figure 5. The chosen design specifically matches shadow and light bleeding on the surface of the object it was mounted on, as validated in a pilot study with 5 participants.

3.3.2 Gem Marking and Distance. In order to emulate the stripe markings of a physical gem as precisely as possible in our virtual gems, and to ensure that rendering closely mirrored the original gem, we developed a custom UV mapping based on a physical gem's scan. We performed multiple tests and comparisons to achieve the best possible match, through early pilots with team members. We then conducted a pilot study with five users who did not subsequently participate in our main study. Each pilot session lasted approximately 30 minutes, and helped us select the most suitable gem rendering model from 25 pre-designed options, with each participant narrowing their choices to their top 5 and then top 2.

We then conducted tests for gem identification at distances ranging from 1 to 5 meters. During these tests, each participant was required to identify 10 virtual gems, both marked and unmarked, presented in a random sequence. Participants achieved nearly perfect scores at distances of 1 meter (100% accuracy) and 2 meters (98% accuracy). Accuracy was still high at a 3-meter distance (96% accuracy), but dropped significantly beyond 4 meters (72% accuracy), plummeting to 28% at 5 meters. Participants noted that the images appeared pixelated at these greater distances, making it difficult to discern the presence of stripes. We therefore positioned all gems within 3 meters (at their shortest distance) of the planned walking path. During the main user study, 99.7 percent of all classifications occurred within a range of 2.6 meters.

3.3.3 Gliding Spotlight. The choice of this "spotlight" path guidance mechanism, which indicates a region that the user must stay within (much like real-life spotlights in stage performances), was driven by our goal to investigate visual cues that indicate safe walking areas in an environment during a timed navigation task, akin to real-world tasks such as emergency response, evacuation training, or navigation and wayfinding. The mechanism also loosely models self-imposed casual walking speed along a familiar path, which is meaningful for future always-on AR opportunities.

There are different ways to enforce a specific path on the user, ranging from directional arrows (that are merely suggestions of the path to follow) to a visibility bubble (that penalizes a user if they do not remain in the bubble). We chose to implement a mechanism with an intermediate level of enforcement: a moving "spotlight" ring that indicates the area to stay within with a full 360° visual restriction that changes color when the user is not inside (but does not penalize them in any other way). We also deliberately chose to place the visual at approximately waist height rather than on the floor, so that participants would not need to look away from the surrounding environment in order to keep track of the ring while they were actively scanning for gems. We conducted several early pilots with the ring implementation to ensure that target (and other meaningful) objects were not obstructed by the ring.

In our pilot experiments, we tested five gliding spotlight ring speeds to determine the optimal speed for the gliding spotlight, aiming to approximate a "normal," non-dawdling, purposeful walking pace. Our findings indicate that the arrived-at speed, 0.92 meters per second, ended up being only slightly higher than the median speed CHI '25, April 26-May 01, 2025, Yokohama, Japan



Figure 7: Additional physical and virtual objects for the surprise recall test. The environment includes six objects of different sizes that were tested on the surprise memory task. Three virtual objects (top row): a coconut, charcoal grill, and hammock. Three physical objects (bottom row): a mini ice box, camping chair, and beach umbrella. Object sizes ranged from small to large, shown from left to right. After all eight trials of the gem search task were completed, participants were asked to recall selected objects and indicate whether each was physical or virtual. The size of each object was measured using its bounding box.

chosen by individuals when freely optimizing target detection accuracy and time. This suggests that the differences in participant behavior between the spotlight and free walking conditions were not solely due to the time taken to complete the task.

Although the spotlight speed is still relatively slow compared to typical brisk walking speeds, it closely matches the speed at which users walk while actively scanning the environment for targets.

3.4 Apparatus

The experiment was conducted on a Microsoft HoloLens 2 headset, and responses were recorded from a Bluetooth clicker, with two buttons used for the two types of gems (marked and unmarked). The physical gems were Philips Hue Go portable smart lamps, with color and brightness of all lamps in the scene controlled from the Philips Hue app.

We measured light levels in the environment and the HoloLens-2 display brightness for a representative object from the participant's eye position relative to the headset. Measurements were made avoiding direct light sources, looking parallel to the ground, and were collected at either ends of the hallway and at the middle. The luminance measurement for the physical environment reads (M = 336.00, SD = 105.78) while the AR objects emit (M = 487.33, SD = 166.54) lux. Measurements were made for the two types of target gems in the search task. Virtual gems emit (M = 1917.67, SD = 274.87) lux while physical gems emit (M = 2242.33, SD = 60.02) lux.

The three physical objects used in the memory task (Figure 7) were carefully placed in position before the participants began their trials. These three objects were deliberately chosen to be of different sizes (small, medium, and large), and the three virtual objects used

CHI '25, April 26-May 01, 2025, Yokohama, Japan



Figure 8: An aerial perspective of our user study location, with virtual objects marked in green, physical objects highlighted in pink, and Godzilla represented in yellow. From this top view, it can be observed that the distribution of both virtual and physical objects is spread out evenly across the entire location.

in the memory task were also specifically chosen to have the same range of sizes as the physical ones.

3.5 Participants

The study involved 24 adults aged 19 to 32 years (M = 23.88, SD = 4.22), comprising 13 females and 11 males. Nine participants had corrected vision and wore contact lenses during the study, and all were right-handed. Participants were compensated at a rate of \$20 per hour. Their familiarity and experience with VR/AR varied: 20.83% had no experience with VR, 54.16% had no experience with AR, 58.33% had used VR 1-10 times, and 41.66% had used AR 1-10 times. All procedures were approved by the University's Human Subjects Committee.

3.6 Procedure

Participants first signed the informed consent form and completed a pre-study questionnaire collecting demographic information, remembrance of objects, sense of direction (Santa Barbara Sense of Direction Scale; SBSOD [28]), and experience with AR/VR and gaming. Participants were then fitted with the AR headset and watched a video to orient them to the experiment and explain the search task. They then performed eye gaze calibration within the study headset. Following this, participants completed an eight-minute in-headset tutorial to practice gem classification and spotlight navigation. Only virtual gems were included in the tutorial, though participants were explicitly told they would be searching for both physical and virtual gems in the task.



Figure 9: Highly salient object manipulation. A virtual model of the monster "Godzilla", standing 190 cm tall, was positioned approximately halfway down the hallway. Godzilla was present in every trial, and we measured whether participants perceived it by asking if they saw anything "noteworthy" at the end of each trial.

The experimenters set up the experimental area out of sight from participants. When the scene was ready, the participant began the trial, which involved walking forwards through the experimental area. Participants were required to only move forwards so that they could not retrace their steps and revisit an area of the environment multiple times. Once the trial was complete, the participant was taken outside of the experimental area to complete a post-trial questionnaire while an experimenter set up the gems for the next trial. In the post-trial questionnaire, they were asked about the difficulty of the trial they just completed, and whether they saw anything noteworthy in the trial (later used to detect when Godzilla was first noticed). This was repeated for all eight trials. Once all the trials were complete, participants filled out the encountered objects questionnaire to assess memory of objects in the environment, followed by a post-experiment questionnaire for experience ratings. The experiment took approximately 90 minutes in total, including the tutorial and post-experiment survey.

3.7 Analysis

Three experimenters reviewed all 192 user trials using a custom playback software to ensure stability and accuracy in user categorization input, and check for misregistrations of responses. Any such misregistrations were corrected by the experimenters before data analysis. Common tactics used by participants included identifying the gem from a distance, categorizing it to the best of their ability, and then re-categorizing it when they had the opportunity to get closer. We did not penalize for re-categorizations, and when a gem had multiple responses, we only considered the latest response for analysis.

We recorded two behavioral metrics during each trial: the total *distance traveled* during the trial (in metres), and the total head rotation measured as the accumulated quaternion distance norm [30]. We also measured gem search performance as *detection accuracy* (proportion of gems in the trial that were classified, even if the classification response was incorrect) and *discrimination accuracy* (proportion of detected gems that were correctly classified). After each trial, participants filled out a post-trial questionnaire measuring cognitive load (as indexed by the *ease* and *perceived success* in each trial), as well as their interactions with the spotlight, lighting, and objects.

Overall trial *ease* was measured relative to the previous trial ("Was this round easier than the last round?"), and the 7-point Likert scale responses were converted to absolute scores for each trial and normalized over the entire dataset. We also measured the *ease of the spotlight* ("How easy was it to follow the spotlight?") for trials that had a spotlight, and *perceived success* in every trial ("In your opinion, how successful were you in identifying the gems?") on a 7-point Likert scale.

We analyzed the open-ended responses from this questionnaire by implementing grounded theory (GT) techniques, taking a qualitative and inductive approach to derive new conclusions from participant responses [15]. In this post-trial questionnaire, we asked participants what "noteworthy objects" they found in each trial and read through each response, recording the trends in user feedback. For participants who mentioned noticing the highly salient virtual "Godzilla", we recorded the conditions of the trial in which it was first noticed (low vs. high augmentation density, and path guidance presence). At the conclusion of the study, we verified this information by asking participants to indicate if they had seen the Godzilla, and if so, during which trial they noticed it. We made sure that this information matched the post-trial questionnaire before analyzing it for accuracy and consistency.

Responses to the object memory test were only analyzed for the six objects that were actually present in the environment, out of the twelve presented to participants in the test. The metrics of *recall* ("Was <object> present in the environment") and *precision* (for objects that the user marked present, "Was <object> physical or virtual?") were computed to estimate user memory and spatial awareness. We measured participants' accuracy of *recall* of objects in the environment as the proportion of objects they correctly identified as present/absent, and *precision* as the proportion of recalled (and present) objects that were correctly identified as physical/virtual.

We conducted factorial analyses within a linear model framework, using the lme4 package in R. For variables that were continuous and appropriately distributed, we used the general linear model. For those that were binomial (gem detection accuracy and gem discrimination accuracy), we used the generalized linear model. For the linear models we report β , standard errors, 95% CIs and p-values from Wald chi-square tests.

For the per-trial metrics (distance traveled and total head rotation), the fixed effects were path guidance and augmentation density. The participant ID was included as a random effect. Total head rotation was modeled as a Gaussian distribution, and distance traveled did not satisfy normality and was therefore modeled as a Gamma distribution.

The Godzilla observation data was analyzed using the Chi-square test of independence, and follow-up comparisons for each independent variable (augmentation density and path guidance presence) were performed using the Chi-square goodness of fit test. The cognitive load metrics (ease and perceived success) did not satisfy the normality assumptions of the ANOVA, and were hence analyzed



Figure 10: Distance traveled (left) and total head rotation (right), plotted as a function of augmentation density and path guidance. Participants traveled significantly further distances in the absence of the spotlight. Head rotation was significantly higher in trials when the spotlight was absent and was also significantly higher in high augmentation trials vs low augmentation ones. Gray dots represent data for individual participants and trials. Error bars = 95% CI.

using the ART-ANOVA [81] from the ARTool package in R, with results reported as Type III Wald F tests with Kenward-Roger adjusted degrees of freedom. The *spotlight ease* metric also did not satisfy the normality assumption, and was analyzed using the Wilcoxon signed-rank test with the Bonferroni correction for multiple comparisons. Precision and recall in the object memory test were also analyzed using the Wilcoxon signed-rank test from the rstatix package, with the Bonferroni correction for multiple comparisons.

4 Results

The results are organized into four sections. First, we examine participants' physical movement as they navigated through the environment, analyzing distance travelled and head-rotation patterns. Second, we evaluate performance on the gem search task, focusing on gem detection and discrimination accuracy. Third, we assess participants' recall for regular objects encountered in the environment, measured by the number of items recalled in the surprise memory probe at the end of the study. Fourth, we explore recall for the highly salient item, Godzilla, to determine the condition in which it was most likely to be first detected.

4.1 Movement Metrics

Distance travelled and total head rotation are plotted as a function of augmentation density (low, high) and path guidance (spotlight absent vs. present) in Figure 10.

Augmentation density did not impact distance traveled (β = -0.012, *SE* = 0.012, *p* > 0.05, *CI* = (-0.035, 0.011)), but the presence of the spotlight had a significant influence on the distance traveled (β = -0.076, *SE* = 0.012, *p* < 0.0001, *CI* = (-0.099, -0.053)), such that participants travelled further when the spotlight was not present. Participants were instructed to maintain their position within the spotlight on spotlight-present trials, but despite these instructions



Figure 11: Gem detection accuracy, plotted as a function of augmentation density and gem type for the two path guidance conditions. Overall, participants detected significantly more physical gems than virtual ones. Gray dots represent data for individual participants and trials. Error bars = 95% CI.



Figure 12: Gem discrimination accuracy, plotted as a function of augmentation density and gem type for the two path guidance conditions. Overall, participants accurately classified physical gems significantly more than virtual ones. Gray dots represent data for individual participants and trials. Error bars = 95% CI.

some participants still fell behind on specific trials. These participants and trials are depicted in the outlier data points in Figure 10. There was no significant interaction between augmentation density and path guidance ($\beta = 0.007$, SE = 0.017, p > 0.05, CI = (-0.025, 0.04)).

Total head rotation was influenced by both augmentation density (β = -2.166, *SE* = 1.05, *p* < 0.05, *CI* = (-4.303, -0.021)) and path guidance (β = -3.684, *SE* = 1.05, *p* < 0.001, *CI* = (-5.789, -1.569)). If we consider head rotation as a proxy for scanning of the environment, these data can be interpreted as increased scanning as a function of both conditions of higher augmentation density and also conditions where the spotlight was absent. There was no significant interaction

between augmentation density and path guidance (β = 1.213, *SE* = 1.485, *p* > 0.05, *CI* = (-1.79, 4.18)).

In summary, distance traveled was influenced by the presence of the spotlight, and head rotation was impacted by both the augmentation density and the presence of the spotlight.

4.2 Gem Search Performance

Gem detection accuracy is plotted as a function of augmentation density, gem type and path guidance in Figure 11.

Neither augmentation density (β = -0.946, *SE* = 0.593, *p* > 0.05, *CI* = (-2.11, 0.216)) nor path guidance (β = -0.228, *SE* = 0.67, *p* > 0.05, *CI* = (-1.541, 1.085)) impacted gem detection accuracy, but gem type was a significant predictor of detection accuracy (β = On the Go with AR



Figure 13: Overall ease (left) and perceived success (right) as reported in the post-trial questionnaire, higher scores indicate higher ease and success respectively. Gray dots represent data for individual participants. Error bars = 95% CI.

-2.024, SE = 0.539, p < 0.001, CI = (-3.08, -0.969)), with physical gems being detected significantly more than virtual gems. There were no significant two-way or three-way interactions (all β < 1.281, SE < 0.892, p > 0.05).

Gem discrimination accuracy is plotted as a function of augmentation density, gem type and path guidance in Figure 12. Gem discrimination results were similar to gem detection, such that neither augmentation density (β = -0.021, *SE* = 0.706, *p* > 0.05, *CI* = (-1.404, 1.362)) nor path guidance (β = 0.694, *SE* = 0.861, *p* > 0.05, *CI* = (-0.994, 2.381)) impacted gem discrimination accuracy, but gem type was a significant predictor of detection accuracy (β = -2.342, *SE* = 0.536, *p* < 0.0001, *CI* = (-3.393, -1.29)), with physical gems being correctly discriminated significantly more than virtual gems. There were no significant two-way or three-way interactions (all β < 1.373, *SE* < 1.168, *p* > 0.05).

In summary, both gem detection and discrimination were both only significantly influenced by the gem type, with higher accuracies for the physical gems as compared to virtual ones.

4.3 Cognitive Load

There was a main effect of path guidance on the self-reported ease of the trial, with lower ease scores when the path guidance was present (F(1, 165) = 18.165, p < 0.0001). There was no main effect of augmentation density (F(1, 165) = 0.203, p > 0.05) or interaction (F(1, 165) = 0.013, p > 0.05).

There was no effect of augmentation density on the ease of following the spotlight (V = 153, p > 0.05, r = .282, *small*).

There was a main effect of path guidance on the perceived success in the task, with lower perceived success when the path guidance was present (F(1, 165) = 35.006, p < 0.0001). There was no main effect of augmentation density (F(1, 165) = 0.54, p > 0.05) or interaction (F(1, 165) = 0.079, p > 0.05).

In summary, the presence of the path guidance led to lower ease and lower perceived success in the task. Augmentation density did not impact any of our metrics of cognitive load.



Figure 14: Recall (left) and precision (right) on the postexperiment object memory test, for objects that were present in the environment. 1 represents a perfect score for both metrics, and 0 the lowest score. Black dots represent data for individual participants. Error bars = 95% CI.

4.4 Memory for Goal-Irrelevant Objects in the Environment

There was no significant difference in the recall of physical vs. virtual objects (V = 69.5, p > 0.05, r = .288, *small*), as shown in Figure 14. However, recall precision (i.e. the proportion of recalled objects that were correctly classified as physical or virtual) was significantly higher for virtual objects as compared to physical objects (V = 2.5, p < 0.01, r = .605, *large*). Thus, although there does not appear to be a difference in recall of physical vs. virtual objects, participants experienced source confusion much more with physical objects as compared to virtual ones.

4.5 Awareness of Highly Salient Item (Godzilla)

Here, we examined whether the level of augmentation density and presence or absence of path guidance influenced participants' ability to detect and remember the virtual Godzilla figure. Given that Godzilla was constantly present in the environment, we were able to obtain two metrics. First, did participants detect Godzilla at all during the experiment? Second, if detected, under what conditions (augmentation density and path guidance) did detection first occur?

The task was relatively challenging but still feasible, as 16/24 (66.66%) of participants noticed Godzilla. A Chi-square test of independence indicated that the experimental condition had an impact on awareness of Godzilla ($\chi^2(1, N=24) = 15, p < 0.0005$). Participants noticed Godzilla significantly more in the absence of path guidance ($\chi^2(1, N=24) = 61.25, p < 0.0005$), with only one participant noticing it during a spotlight-present condition. The augmentation density also influenced observation of Godzilla, with participants noticing it significantly more in the low-augmentation condition ($\chi^2(1, N=24) = 5.33, p < 0.025$). Furthermore, a Mann-Whitney U test showed that head rotation was significantly higher in participants who noticed Godzilla ($U = 19, Z = -2.84, p < 0.005, \eta^2 = 1.17$), suggesting that Godzilla detection likelihood was greater in participants who scanned the environment more extensively.

4.6 Subjective Participant Feedback

In this section, we analyze and discuss the qualitative feedback provided by participants, highlighting key themes that emerged from their experiences during the trials.

Participants did notice the difference between lightly and heavily augmented trials, with one participant (P6) noting, "In heavily augmented trials, it was harder to identify gems and objects sometimes." Interestingly, participants also commented that "With background trial (high-augmentation), it was easier for me to find physical gems" (P13), indicating a potential divide between the physical and virtual environments which made it easier to focus on the physical environment when the virtual environment was cluttered.

This divide between the virtual and physical worlds was further evidenced by the results of the gem search task. Participants detected physical targets at a significantly higher rate than virtual targets across all conditions, even though both types of targets were designed to be as visually identical as possible, with similar relative brightness to the surrounding objects. This difference in detection could be attributed to an increased focus on the physical environment over the virtual, though the limited field of view of the virtual environment compared to the physical one may have also contributed to this effect.

Following the spotlight was designed to mimic regular walking, as the guided pace matched a natural walking speed during the search task. One participant (P11) mentioned, "Following the ring felt similar to regular walking, as the guided pace was consistent with a natural walking speed." While some participants felt the spotlight was too fast, others found it slower than their usual pace, which allowed them to spend more time on task. Participant (P24) stated, "The spotlight's slower pace in one of the trials allowed me to spend more time looking closely at the gems." Ultimately, each individual's comfort level with these spatial UI interfaces and their typical task performance speed played a significant role, though most participants managed to keep up with the designated timing. This precision is optimal for our goal of studying attention effects in a dynamic, "on the go" manner by enforcing controlled walking motions. Our approach involves balancing the benefits of a predetermined and enforced walking path with those of a predetermined but more flexibly followed walking path.

This feedback provided valuable insights that should inform the design considerations for AR applications intended for use on the go. Considerations include managing the balance between physical and virtual elements and designing guided navigation that aligns with natural walking speeds while accommodating individual user preferences. We elaborate on these in the following section.

5 Discussion

AR technology allows us to integrate virtual content into the physical world and has potential to be a primary source of information consumption on the go. While these technologies hold much promise, it is crucial to address the potential issues associated with AR clutter. This visual overload has the potential to distract users from their primary tasks, and perhaps even jeopardize their safety by reducing awareness of the physical surroundings. To address these potential concerns, we designed a study to test the impact of varying levels of augmentation and path guidance on participants' performance and awareness of objects in their environment in a locomotion-based search task.

Our first goal was to examine the impact of varying levels of augmentation on user behavior and awareness of their environment. While there did not appear to be any change in search task performance (gem detection and discrimination accuracy) with augmentation density, participants did look around the environment significantly more (as evidenced by increased head rotation) in the high augmentation condition. This suggests that participants opted to scan the environment much more to find gems when augmentation density was high, which was supported by participant comments in the post-study interview. Similarly, participants spotted Godzilla significantly less in the high augmentation trials. These results suggest that increased levels of augmentation clutter in the environment can exacerbate inattention blindness and make exploring the environment more challenging.

There was no difference in the recall of physical and virtual objects from the post-study object recall test, though participants falsely remembered some physical objects as virtual but no virtual objects as physical, a source confusion phenomenon previously reported in the literature [22]. One contributing factor to this source confusion effect might be the novelty effect of the virtual augmentations. Notably, 19 out of 24 participants had prior VR headset experience, while 13 out of 24 had never tried head-mounted AR devices such as the HoloLens 2.

Previous research involving unannounced recall tasks [34, 41, 75] has shown that participants performing a search task with only virtual targets had a much better awareness of the virtual environment compared to the physical one. The inclusion of physical targets in our experiment appears to reduce this effect, suggesting that the nature of the search targets impacts awareness of different aspects of the environment. This is good news to mobile AR designers to some extent, as the previous results left open the possibility that use of AR *generally* diminishes recall of real-world objects and infrastructure if not carefully designed.

Our results also further support the hypothesis that there is a competing focus effect of the virtual and physical environments, and participants tend to focus on one or the other depending on the nature of the task. In some of the prior work[34, 41], the task only required attention to the virtual environment and attention to the physical environment was important primarily for safe navigation, whereas our experiment demanded attention to both virtual and physical spaces. And the perceived nature of the task at hand (whether one is operating in virtual space or in physical space) definitely matters. A study comparing three-week-long care for a virtual pet in AR and VR [59] for example led to vastly different feelings of ownership and emotional connection. Surely, task virtuality can influence focus and recall of objects as well, such as to result in higher recall of virtual task.

In our case, the lack of difference in recall for the two types of objects, while obviously not a finding in itself, is noteworthy when compared to the mentioned previous research and seems to indicate that our task is indeed perceived as taking place in both the virtual and physical realm. With our 50-50 split of virtual and physical gems, the search results suggest higher participant attention to the physical environment, as evidenced by the higher detection accuracy for physical gems. At the same time, the novelty of the virtual objects could have been a competing factor that led to the observed results in the recall test, with higher source confusion for physical objects.

Participants were also worse at discriminating surface markings for the virtual gems compared to physical ones, which was likely due to the HoloLens 2's display aliasing which made the markings on virtual gems harder to discern from a distance. This issue was not observed in our pilot studies, perhaps because pilot participants were slightly more expert with AR and more familiar with the optical see-through display technology.

The presence of the path guidance spotlight altered movement behavior, with reduced distance traveled, reduced head rotation, and strongly reduced likelihood of spotting the highly conspicuous Godzilla model. The spotlight also decreased both self-reported ease and perceived success in the task, although most participants managed to stay within the designated timing, demonstrating that the interface effectively guides users with precise timing. Additionally, the presence of the spotlight did not impact the task performance metrics (which were only affected by the gem type).

For appreciating the broader significance of our results, imagine a future where people routinely utilize lightweight AR glasses while walking, either in unfamiliar environments, where a use case might be navigation and wayfinding, or during daily common familiar walks, where a use case might be parallel consumption of AR for such purposes as entertainment, education, communication, corporate duties, or personal information management. Our results demonstrate that AR tools, even if effective at the function they are designed for, may have negative side effects. Application designers must carefully consider the benefits and costs of AR annotations and tools, and focus on minimizing adverse effects on user experience. For our path guidance tool, participants' subjective feedback suggested that the level of comfort with the spotlight varied depending on each users' general walking pace, indicating a potential advantage of personalization for guidance tools of this kind.

5.1 Design Considerations

Here, we share our guidance and recommendations for the design of on-the-go AR applications, based on insights gathered from our physical and virtual natural locomotion study.

We found that increased clutter in environments can exacerbate inattention blindness. Therefore, designers should strive to reduce unnecessary clutter and highlight or draw attention to important objects, both task-relevant and task-irrelevant, especially when these objects are crucial for user safety.

The nature of a user's task focus–whether primarily physical or virtual–affects their awareness of various environmental aspects. For tasks that require focus on the virtual environment, designers should find ways to emphasize important physical objects to ensure safe navigation in the real world.

Effective mechanisms to control walking paths can ensure safe navigation without impairing the user's focus on their main task. Personalizing these tools based on the user's walking behavior and familiarity with different spatial interfaces can enhance their effectiveness. Additionally, matching the task with the user's natural walking speed can provide a seamless experience, allowing users to concentrate on the task while being safely guided, without compromising performance.

5.2 Limitations

While all our efforts were focused on ensuring experimental rigor and controlling unwanted impact factors and dependencies, our study does have limitations. First, the HoloLens 2 has a relatively small field of view as compared to some newer AR HMDs. We are familiar with the existing literature on FOV limitations and their potential impacts, as discussed in Section 2. However, the HoloLens 2 was the most suitable device for this task due to the need for an optical-see-through headset with stable tracking and registration in larger environments (which we achieved with MRTK's world locking tools). When advancements in FOV and resolution are achieved with the next generation of AR devices, follow-up studies should be conducted to determine if these effects are still observed.

Second, we put a lot of effort into trying to ensure a similar appearance of physical and virtual gems, but some challenges, such as uneven color appearance at different visual angles in the HoloLens 2 display could not be addressed by our modeling and rendering efforts. Positive comments and reactions from some participants on their first encounter of both physical and virtual targets give us confidence that our efforts were worthwhile, but in the end, OST-AR technology does not currently support a truly seamless integration of real and virtual space, and it remains to be seen if some effects would come out slightly different with more accurate visual coherence.

Third, while our intention for this study was to inform realistic future AR-on-the-go scenarios, the need to run a controlled user study meant that we needed to pick a particular representative task. Our choice fell on target search as a well-studied attention task that is a likely element of any AR scenario involving the physical space around a mobile user, but there will clearly be some other possible AR-on-the-go tasks for which this choice represents but limited ecological validity. Still, we believe that a sizable set of future AR efforts can benefit from the initial results presented here.

Fourth, there are still alternative hypotheses regarding the lack of difference in recall between physical and virtual objects, when compared to previous work that found diminished recall of physical objects. We believe that the introduction of physical targets in our study, balanced with virtual targets, is the main factor behind the absence of differing recall accuracy, but our experiment had also some other key differences with the work we compared results to, namely a smaller search environment and being indoors vs. outdoors. It is possible that these differences may have played a role too, and future experiments will be needed to examine these additional factors.

6 Conclusion

Through this project, we examined how varying levels of augmentation affect user behavior and awareness of their environment. Our findings show that high augmentation density leads to increased head rotation and environmental scanning, suggesting that participants needed to search more thoroughly to find virtual gems. User comments supported this, noting difficulties in identifying gems amidst heavy augmentations and the relative ease of finding physical gems in cluttered virtual settings (an indication that there is clearly still a divided-worlds effect – clutter in virtual annotation space does not necessarily impact the physical world to the same extent). The higher detection rate of physical targets over virtual ones, despite their visual similarities and focus on virtual targets in the tutorial, further emphasizes divide between the physical and virtual world. The limited virtual-environment field of view of the AR headset most likely plays a role here.

In a post-study recall test, there was no significant difference in recalling physical versus virtual objects, although participants sometimes misremembered physical objects as virtual, a previously demonstrated source-confusion effect likely influenced by the novelty of virtual augmentations. Previous studies showed higher recall of virtual objects when tasks are focused solely on the virtual environment. The nature of the task at hand – is it a task in the virtual realm, in the physical realm, or in both? – appears to make a big difference concerning scene and object awareness.

This study highlights the complex interplay between physical and virtual environments in augmented reality, emphasizing the need for better AR design to integrate virtual elements seamlessly into the physical world and enhance user engagement while walking. It provides insights about the impact of AR content on walking users' behavior and cognitive processes, which is relevant for such tasks as navigation and wayfinding in unfamiliar environments, or for personal applications (entertainment, education, personal information management) that make use of referencing the physical world during common daily walks.

Acknowledgments

This work was funded by NSF Award IIS-2211784, the Army Research Office under contract W911NF-19-D-0001, the Office of Naval Research with grant N00014-23-1-2118, and DARPA under contract HR001124C0409.

References

- Haley Adams, Jeanine Stefanucci, Sarah Creem-Regehr, and Bobby Bodenheimer. 2022. Depth perception in augmented reality: The effects of display, shadow, and position. In 2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (Christchurch, New Zealand). IEEE.
- [2] Edward Awh, Artem V. Belopolsky, and Jan Theeuwes. 2012. Top-down versus Bottom-up Attentional Control: A Failed Theoretical Dichotomy. *Trends in Cognitive Sciences* 16, 8 (Aug. 2012), 437–443. https://doi.org/10.1016/j.tics.2012. 06.010
- [3] Kassandra Bauerfeind, Julia Drüke, Jens Schneider, Adrian Haar, Lennart Bendewald, and Martin Baumann. 2021. Navigating with augmented reality-how does it affect drivers' mental load? *Applied ergonomics* 94 (2021), 103398.
- [4] Kai Biermeier, Ingrid Scharlau, and Enes Yigitbas. 2024. Measuring visual attention capacity across xReality. In Proceedings of the 17th International Conference on PErvasive Technologies Related to Assistive Environments (Crete Greece). ACM, New York, NY, USA.
- [5] Mark Billinghurst, Hirokazu Kato, and Ivan Poupyrev. 2001. The MagicBook: a transitional AR interface. Comput. Graph. 25, 5 (Oct. 2001), 745–753.
- [6] Frank Biocca, Arthur Tang, Charles Owen, and Fan Xiao. 2006. Attention funnel: omnidirectional 3D cursor for mobile augmented reality platforms. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Montréal, Québec, Canada) (CHI '66). Association for Computing Machinery, New York, NY, USA, 1115–1122. https://doi.org/10.1145/1124772.1124939
- [7] Elise Bonnail, Julian Frommel, Eric Lecolinet, Samuel Huron, and Jan Gugenheimer. 2024. Was it Real or Virtual? Confirming the Occurrence and Explaining Causes of Memory Source Confusion between Reality and Virtual Reality. In Proceedings of the CHI Conference on Human Factors in Computing Systems. 1–17.
- [8] Yu-Cheng Chang, Yen-Pu Wang, Chiao-Ju Chang, Wei Tian Mireille Tan, Yu Lun Hsu, Yu Chen, and Mike Y Chen. 2024. Experience from Designing Augmented

Kim et al.

Reality Browsing Interfaces for Real-world Walking Scenarios. Proceedings of the ACM on Human-Computer Interaction 8, MHCI (2024), 1–26.

- [9] Lung-Pan Cheng, Eyal Ofek, Christian Holz, and Andrew D Wilson. 2019. Vroamer: generating on-the-fly VR experiences while walking inside large, unknown real-world building environments. In 2019 IEEE conference on virtual reality and 3D user interfaces (VR). IEEE, 359–366.
- [10] Yifei Cheng, Yukang Yan, Xin Yi, Yuanchun Shi, and David Lindlbauer. 2021. Semanticadapt: Optimization-based adaptation of mixed reality layouts leveraging virtual-physical semantic connections. In *The 34th Annual ACM Symposium on* User Interface Software and Technology. 282–297.
- [11] FRANCESCO Chiossi, U Gruenefeld, BAOSHENG JAMES Hou, JOSHUA Newn, CHANGKUN Ou, RULU Liao, ROBIN Welsch, and SVEN Mayer. 2024. Understanding the impact of the realityvirtuality continuum on visual search using fixation-related potentials and eye tracking features. In Proceedings of MobileHCI.
- [12] Hyunsung Cho, Yukang Yan, Kashyap Todi, Mark Parent, Missie Smith, Tanya Jonker, Hrvoje Benko, and David Lindlbauer. 2024. MineXR: Mining Personalized Extended Reality Interfaces (CHI'24). Association for Computing Machinery, New York, NY, USA.
- [13] Michael A Cohen, George A Alvarez, and Ken Nakayama. 2011. Natural-scene perception requires attention. *Psychol. Sci.* 22, 9 (Sept. 2011), 1165–1172.
- [14] Robbe Cools, Augusto Esteves, and Adalberto L Simeone. 2022. Blending spaces: Cross-reality interaction techniques for object transitions between distinct virtual and augmented realities. In 2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (Singapore, Singapore). IEEE.
- [15] Juliet Corbin and Anselm Strauss. 2014. Basics of qualitative research: Techniques and procedures for developing grounded theory (fourth ed.). Sage publications.
- [16] Juan David Hincapie-Ramos, Levko Ivanchuk, Srikanth Kirshnamachari Sridharan, and Pourang Irani. 2014. SmartColor: Real-time color correction and contrast for optical see-through head-mounted displays. In 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (Munich, Germany). IEEE.
- [17] Benjamin J Dixon, Michael J Daly, Harley HL Chan, Allan Vescan, Ian J Witterick, and Jonathan C Irish. 2014. Inattentional blindness increased with augmented reality surgical navigation. *American journal of rhinology & allergy* 28, 5 (2014), 433–437.
- [18] Roman Egger and Larissa Neuburger. 2020. Augmented, virtual, and mixed reality in tourism. *Handbook of e-Tourism* (2020), 1–25.
- [19] Stephen R Ellis and Brian M Menges. 1998. Localization of virtual objects in the near visual field. *Human factors* 40, 3 (1998), 415–431.
- [20] Austin Erickson, Kangsoo Kim, Gerd Bruder, and Gregory F Welch. 2020. Exploring the limitations of environment lighting on optical see-through head-mounted displays. In Symposium on Spatial User Interaction (Virtual Event Canada). ACM, New York, NY, USA.
- [21] Robin Eyraud, Elisabetta Zibetti, and Thierry Baccino. 2015. Allocation of visual attention while driving with simulated augmented reality. Transp. Res. Part F Traffic Psychol. Behav. 32 (July 2015), 46–55.
- [22] Ajoy S Fernandes, Ranxiao Frances Wang, and Daniel J Simons. 2015. Remembering the physical as virtual: Source confusion and physical interaction in augmented reality. In Proceedings of the ACM SIGGRAPH Symposium on Applied Perception. 127–130.
- [23] Joseph L Gabbard, Divya Gupta Mehra, and J Edward Swan. 2019. Effects of AR display context switching and focal distance switching on human performance. *IEEE Trans. Vis. Comput. Graph.* 25, 6 (June 2019), 2228–2241.
- [24] Joseph L Gabbard, J Edward Swan, Jason Zedlitz, and Woodrow W Winchester. 2010. More than meets the eye: An engineering study to empirically examine the blending of real and virtual color spaces. In 2010 IEEE Virtual Reality Conference (VR) (Boston, MA, USA). IEEE.
- [25] Juan Garzón and Juan Acevedo. 2019. Meta-analysis of the impact of Augmented Reality on students' learning gains. *Educational Research Review* 27 (2019), 244– 260.
- [26] Google Maps Live View 2020. Google Maps Live View. https://blog.google/products/maps/new-sense-direction-live-view/, accessed Sept. 2024.
- [27] Jens Emil Sloth Grønbæk, Juan Sánchez Esquivel, Germán Leiva, Eduardo Velloso, Hans Gellersen, and Ken Pfeuffer. 2024. Blended Whiteboard: Physicality and Reconfigurability in Remote Mixed Reality Collaboration. In Proceedings of the CHI Conference on Human Factors in Computing Systems. 1–16.
- [28] Mary Hegarty, Anthony E Richardson, Daniel R Montello, Kristin Lovelace, and Ilavanil Subbiah. 2002. Development of a self-report measure of environmental spatial ability. *Intelligence* 30, 5 (2002), 425–447.
- [29] Tobias Höllerer and Steve Feiner. 2004. Mobile augmented reality. Telegeoinformatics: Location-based computing and services 21 (2004), 221–260.
- [30] Du Q Huynh. 2009. Metrics for 3D rotations: Comparison and analysis. Journal of Mathematical Imaging and Vision 35 (2009), 155-164.
- [31] Adam Ibrahim, Brandon Huynh, Jonathan Downey, Tobias Höllerer, Dorothy Chun, and John O'donovan. 2018. Arbis pictus: A study of vocabulary learning with augmented reality. *IEEE transactions on visualization and computer graphics* 24, 11 (2018), 2867–2874.

- [32] Mohamed Kari, Reinhard Schütte, and Raj Sodhi. 2023. Scene Responsiveness for Visuotactile Illusions in Mixed Reality. In Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology. 1–15.
- [33] H Kato, M Billinghurst, I Poupyrev, K Imamoto, and K Tachibana. 2002. Virtual object manipulation on a table-top AR environment. In Proceedings IEEE and ACM International Symposium on Augmented Reality (ISAR 2000) (Munich, Germany). IEEE.
- [34] You-Jin Kim, Radha Kumaran, Ehsan Sayyad, Anne Milner, Tom Bullock, Barry Giesbrecht, and Tobias Höllerer. 2022. Investigating search among physical and virtual objects under different lighting conditions. *IEEE Transactions on Visualization and Computer Graphics* 28, 11 (2022), 3788–3798.
- [35] You-Jin Kim, Myungin Lee, Marko Peljhan, JoAnn Kuchera-Morin, and Tobias Höllerer. 2024. Spatial Orchestra: Locomotion Music Instruments through Spatial Exploration. In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI EA '24). Association for Computing Machinery, New York, NY, USA, Article 420, 5 pages. https://doi.org/10.1145/ 3613905.3648659
- [36] You-Jin Kim, Misha Sra, and Tobias Höllerer. 2024. Audience Amplified: Virtual Audiences in Asynchronously Performed AR Theater. In 2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). 475–484. https://doi.org/ 10.1109/ISMAR62088.2024.00062
- [37] You-Jin Kim, Andrew D Wilson, Jennifer Jacobs, and Tobias Höllerer. 2023. Reality Distortion Room: A Study of User Locomotion Responses to Spatial Augmented Reality Effects. In 2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). IEEE, 1201–1210.
- [38] Naohiro Kishishita, Kiyoshi Kiyokawa, Jason Orlosky, Tomohiro Mashita, Haruo Takemura, and Ernst Kruijff. 2014. Analysing the effects of a wide field of view augmented reality display on search performance in divided attention tasks. In 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). IEEE, 177–186.
- [39] Ernst Kruijff, Jason Orlosky, Naohiro Kishishita, Christina Trepkowski, and Kiyoshi Kiyokawa. 2018. The influence of label design on search performance and noticeability in wide field of view augmented reality displays. *IEEE transactions* on visualization and computer graphics 25, 9 (2018), 2821–2837.
- [40] Stas Krupenia and Penelope M Sanderson. 2006. Does a head-mounted display worsen inattentional blindness?. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 50. SAGE Publications Sage CA: Los Angeles, CA, 1638–1642.
- [41] Radha Kumaran, You-Jin Kim, Anne E Milner, Tom Bullock, Barry Giesbrecht, and Tobias Höllerer. 2023. The impact of navigation aids on search performance and object recall in wide-area augmented reality. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. 1–17.
- [42] Wallace S Lages and Doug A Bowman. 2019. Walking with adaptive augmented reality workspaces: design and usage patterns. In Proceedings of the 24th International Conference on Intelligent User Interfaces. 356–366.
- [43] Eric Laurier, Barry Brown, and Moira McGregor. 2017. Mediated pedestrian mobility: Walking and the map app. In *Traces of a Mobile Field*. Routledge, 117–134.
- [44] Nilli Lavie. 1995. Perceptual Load as a Necessary Condition for Selective Attention. Journal of Experimental Psychology- 21, 3 (1995), 451–468.
- [45] Nilli Lavie and Sally Cox. 1997. On the Efficiency of Visual Selective Attention: Efficient Visual Search Leads to Inefficient Distractor Rejection. *Psychological Science* 8, 5 (1997), 395–398.
- [46] GA Lee and W Hoff. 2020. Enhancing first-person view task instruction videos with augmented reality cues (pp. 666–676).
- [47] Hyunjin Lee, Sunyoung Bang, and Woontack Woo. 2020. Effects of background complexity and viewing distance on an ar visual search task. In 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). IEEE, 189–194.
- [48] Klemen Lilija, Henning Pohl, and Kasper Hornbæk. 2020. Who put that there? temporal navigation of spatial recordings by direct manipulation. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 1–11.
- [49] Ming-I Brandon Lin and Yu-Ping Huang. 2017. The impact of walking while using a smartphone on pedestrians' awareness of roadside events. Accident Analysis & Prevention 101 (2017), 87–96.
- [50] David Lindlbauer and Andy D Wilson. 2018. Remixed reality. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (Montreal QC Canada). ACM, New York, NY, USA.
- [51] Mark A Livingston, Joseph L Gabbard, J Edward Swan, II, Ciara M Sibley, and Jane H Barrow. 2013. Basic perception in head-worn augmented reality displays. In Human Factors in Augmented Reality Environments.
- [52] John Luksas, Kelsey Quinn, Joseph L Gabbard, Mariam Hasan, Janet He, Neha Surana, Moustafa Tabbarah, and Nishant Kishan Teckchandani. 2022. Search and Rescue AR Visualization Environment (SAVE): Designing an AR Application for Use with Search and Rescue Personnel. In 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) (Christchurch, New Zealand). IEEE, 488–492.

- [53] Shinnosuke Manabe, Sei Ikeda, Asako Kimura, and Fumihisa Shibata. 2019. Shadow inducers: Inconspicuous highlights for casting virtual shadows on osthmds. In 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). IEEE, 1331–1332.
- [54] Nadine M Moacdieh and Nadine B Sarter. 2012. Eye tracking metrics: A toolbox for assessing the effects of clutter on attention allocation. In *Proceedings of the Human Factors and Ergonomics Society annual meeting*, Vol. 56. SAGE Publications Sage CA: Los Angeles, CA, 1366–1370.
- [55] Gabrielle Naïmé Mourra, Sylvain Senecal, Marc Fredette, Franco Lepore, Jocelyn Faubert, François Bellavance, Ann-Frances Cameron, Elise Labonté-LeMoyne, and Pierre-Majorique Léger. 2020. Using a smartphone while walking: The cost of smartphone-addiction proneness. *Addictive behaviors* 106 (2020), 106346.
- [56] Youngmin Park and Woontack Woo. 2006. The ARTable: An AR-based tangible user interface system. In *Technologies for E-Learning and Digital Entertainment*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1198–1207.
- [57] D H Paul. 1960. The membrane potential of the anterior byssus retractor muscle of Mytilus edulis. *Nature* 187, 4743 (Sept. 1960), 1121.
- [58] Tomislav Pejsa, Julian Kantor, Hrvoje Benko, Eyal Ofek, and Andrew Wilson. 2016. Room2room: Enabling life-size telepresence in a projected augmented reality environment. In Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing. 1716–1725.
- [59] Lev Poretski, Ofer Arazy, Joel Lanir, Shalev Shahar, and Oded Nov. 2019. Virtual objects in the physical world: Relatedness and psychological ownership in augmented reality. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 1–13.
- [60] Aditya Raikwar, Domenick Mifsud, Christopher D Wickens, Anil Ufuk Batmaz, Amelia C Warden, Brendan Kelley, Benjamin A Clegg, and Francisco R Ortega. 2024. Beyond the Wizard of Oz: Negative Effects of Imperfect Machine Learning to Examine the Impact of Reliability of Augmented Reality Cues on Visual Search Performance. IEEE Transactions on Visualization and Computer Graphics (2024).
- [61] Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs. 1998. The office of the future: A unified approach to image-based modeling and spatially immersive displays. In Proceedings of the 25th annual conference on Computer graphics and interactive techniques. 179–188.
- [62] Donghao Ren, Tibor Goldschwendt, YunSuk Chang, and Tobias Höllerer. 2016. Evaluating wide-field-of-view augmented reality with mixed reality simulation. In 2016 IEEE Virtual Reality (VR). 93-102. https://doi.org/10.1109/VR.2016.7504692
- [63] Patrick Renner and Thies Pfeiffer. 2017. Attention guiding techniques using peripheral vision and eye tracking for feedback in augmented-reality-based assistance systems. In 2017 IEEE Symposium on 3D User Interfaces (3DUI) (Los Angeles, CA, USA). IEEE.
- [64] Damien Constantine Rompapas, Christian Sandor, Alexander Plopski, Daniel Saakes, Joongi Shin, Takafumi Taketomi, and Hirokazu Kato. 2019. Towards large scale high fidelity collaborative augmented reality. *Computers & Graphics* 84 (2019), 24–41.
- [65] Marc Satkowski and Raimund Dachselt. 2021. Investigating the impact of realworld environments on the perception of 2d visualizations in augmented reality. In Proceedings of the 2021 CHI conference on human factors in computing systems. 1–15.
- [66] Bastian Schildbach and Enrico Rukzio. 2010. Investigating selection and reading performance on a mobile phone while walking. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*. 93–102.
- [67] Benjamin Schöne, Marlene Wessels, and Thomas Gruber. 2019. Experiences in virtual reality: A window to autobiographical memory. *Current Psychology* 38, 3 (2019), 715–719.
- [68] Kathryn Y Segovia and Jeremy N Bailenson. 2009. Virtually true: Children's acquisition of false memories in virtual reality. *Media Psychology* 12, 4 (2009), 371–393.
- [69] Sarah Shomstein, Xiaoli Zhang, and Dick Dubbelde. 2023. Attention and Platypuses. WIREs Cognitive Science 14, 1 (Jan. 2023). https://doi.org/10.1002/wcs.1600
- [70] Daniel J Simons and Christopher F Chabris. 1999. Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events. *Perception* 28, 9 (1999), 1059–1074. https://doi.org/10.1068/p2952
- [71] Missie Smith, Nadejda Doutcheva, Joseph L Gabbard, and Gary Burnett. 2015. Optical see-through HUDs effect on depth judgments of real world objects. In 2015 IEEE Virtual Reality (VR) (Arles, Camargue, Provence, France). IEEE.
- [72] J Edward Swan, 2nd, Adam Jones, Eric Kolstad, Mark A Livingston, and Harvey S Smallman. 2007. Egocentric depth judgments in optical, see-through augmented reality. *IEEE Trans. Vis. Comput. Graph.* 13, 3 (May 2007), 429–442.
- [73] Brandon Victor Syiem, Ryan M Kelly, Jorge Goncalves, Eduardo Velloso, and Tilman Dingler. 2021. Impact of task on attentional tunneling in handheld augmented reality. In Proceedings of the 2021 CHI conference on human factors in computing systems. 1–14.
- [74] Markus Tatzgern, Raphael Grasset, Denis Kalkofen, and Dieter Schmalstieg. 2014. Transitional Augmented Reality navigation for live captured scenes. In 2014 IEEE Virtual Reality (VR) (Minneapolis, MN, USA). IEEE.

CHI '25, April 26-May 01, 2025, Yokohama, Japan

- [75] Markus Tatzgern, Valeria Orso, Denis Kalkofen, Giulio Jacucci, Luciano Gamberini, and Dieter Schmalstieg. 2016. Adaptive information density for augmented reality displays. In 2016 IEEE Virtual Reality (VR). IEEE, 83–92.
- [76] Bruce Thomas, Ben Close, John Donoghue, John Squires, Phillip De Bondi, and Wayne Piekarski. 2002. First person indoor/outdoor augmented reality application: ARQuake. *Personal and Ubiquitous Computing* 6 (2002), 75–86.
- [77] Christina Trepkowski, Alexander Marquardt, Tom David Eibich, Yusuke Shikanai, Jens Maiero, Kiyoshi Kiyokawa, Ernst Kruijff, Johannes Schöning, and Peter König. 2021. Multisensory proximity and transition cues for improving target awareness in narrow field of view augmented reality displays. *IEEE transactions* on visualization and computer graphics 28, 2 (2021), 1342–1362.
- [78] Franziska Westermeier, Larissa Brubach, Marc Erich Latoschik, and Carolin Wienrich. 2023. Exploring plausibility and presence in mixed Reality experiences. *IEEE Trans. Vis. Comput. Graph.* PP, 5 (Feb. 2023), 2680–2689.
- [79] Franziska Westermeier, Larissa Brübach, Carolin Wienrich, and Marc Erich Latoschik. 2023. A virtualized augmented reality simulation for exploring perceptual incongruencies. In 29th ACM Symposium on Virtual Reality Software and

Technology (Christchurch New Zealand). ACM, New York, NY, USA.

- [80] Christopher D Wickens and Amy L Alexander. 2009. Attentional tunneling and task management in synthetic vision displays. *The international journal of aviation psychology* 19, 2 (2009), 182–199.
- [81] Jacob O Wobbrock, Leah Findlater, Darren Gergle, and James J Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In Proceedings of the SIGCHI conference on human factors in computing systems. 143–146.
- [82] Jeremy M Wolfe and Todd S Horowitz. 2017. Five factors that guide attention in visual search. Nat. Hum. Behav. 1, 3 (March 2017), 0058.
- [83] Michelle Yeh and Christopher D Wickens. 2001. Display signaling in augmented reality: Effects of cue reliability and image realism on attention allocation and trust calibration. *Human Factors* 43, 3 (2001), 355–365.
- [84] Yiqing Zhu and Nan Li. 2021. Virtual and augmented reality technologies for emergency management in the built environments: A state-of-the-art review. *Journal of Safety Science and Resilience* 2, 1 (2021), 1–10.