Exploring Factors Influencing Object Recall in Augmented Reality

Sangita Kunapuli
University of California, Santa Barbara
sangitakunapuli@ucsb.edu

Tobias Höllerer
University of California, Santa Barbara
holl@ucsb.edu

***Abstract*—Augmented Reality (AR) is an immersive technology that overlays virtual content onto the physical world, enabling users to seamlessly perceive and interact with a mixed reality environment. As AR technology advances, there is increasing potential for users to interface with virtual content in their daily lives, motivating the need to better understand AR's impact on core cognitive functions involving human attention, memory, and recall of their environment. This work aims to investigate the impacts of cueing (directing users to form a mental model of their surroundings) and cognitive load on a user's ability to remember objects within a mixed reality office scene. We examine object level attributes like size, semantic congruency, and virtuality to explore their impact on attention and recall. We discuss the implications of AR interfaces and how AR environments can be designed to better support human cognition in real-world tasks.**

***Index Terms*—human-centered computing, empirical studies in HCI, mixed/augmented reality, perception, memory, attention**

I. Introduction

Augmented reality (AR) enables the seamless integration of virtual content within real world environments, offering a wide range of applications including remote work [1], office productivity, contextual prototyping, medical modeling [2], and advertising [3]. AR headsets like the Magic Leap 2 use optical see-through technology, allowing users to perceive the physical environment through transparent lenses leveraging computer vision. This type of AR maintains realistic spatial depth and lighting, allowing for a more familiar mixed reality experience.

With the steady improvement of augmented reality technology and the increasing accessibility of interaction in mixed reality environments, it is important to understand how humans interact with these systems. Despite the growing use of AR, there is limited literature that specifically examines how AR interfaces and headsets influence cognitive functions such as attention and memory. While human attention is difficult to measure directly, we use object memory and recall as a proxy metric. If a user is able to remember an object, they must have attended to it during their task in the interface.

In this study, we contribute insights and data aligned with our goal of building a structural equation model (SEM) that captures the relative impact of different factors on object recall in AR. Specifically, we focus on variables related to user state (whether the user was alerted to the recall task), object semantic congruence with the environment, and the presence of a secondary audio task as a measure of cognitive load. We aim to quantify how these factors interact to affect user attention and memory performance in augmented reality systems.

II. Related Work

A. Augmented Reality and Memory

Prior work in AR has examined how interface design and augmentation strategies affect attention, navigation, and memory recall. Kumaran et al. (2023) demonstrated that navigation aids in wide-area AR environments significantly influenced both object search performance and recall [4]. The findings from this work suggest that guidance and navigation aids in AR can direct attention effectively and support memory encoding. This is especially notable for virtual targets embedded within physical spaces. In a subsequent study, Kim et al. (2025) explored how varying the density of AR augmentations in an indoor setting impacted attention to both virtual and physical targets during a search task in AR [5]. They found that higher augmentation density had the potential to increase cognitive load, making users less likely to remember nearby objects, especially when multitasking. They also explored semantic incongruence by including a life-sized Godzilla in the scene, and found that not all participants noticed the large incongruent object in light of augmentation density and multiple tasks.

These previous works highlight the role of attention in AR systems. They underscore the idea that memory in AR is largely dependent on the interface design, semantics of the environment, and object features. Our work builds on these prior AR studies by isolating cueing, object factors, and multitasking to evaluate memory and attention, while also simulating real world mixed reality task demands without the added stress of walking.

B. Insights from Psychology For Memory and Cognitive Load

Early psychology research has examined how contextual expectations and mental schemas shape memory. Brewer and Treyens (1981) showed in their psychology user study that individuals are more likely to recall objects that do not adhere to their "schema" of a familiar environment, such as unexpected or incongruent items in a grad student's office [6]. This insight emphasizes that novelty and semantic inconsistency can enhance memory by capturing attention and prompting

Set A Objects

Physical	Virtual
Rubber band ball	Stapler
Calculator	Brown mug on coaster
Pack of ruled index cards	Blue paper tray
Protractor and compass	Wall clock
Water bottle	Stacked book set
Mesh pen holder	Potted desk plant
Composition notebooks	Monitor
Laundry detergent	Large potted plant
Inflatable beach ball	Globe
Standing fan	***Toy train***
Recycling bin	White chair
Swivel office chair	Brown shelves
Gray filing cabinet	Storage cabinet with blue drawers
Bookshelf with textbooks	White desk with drawers
Office desk	Personal computer tower
Round bench with tabletop	***Traffic light***

Fig. 1. The first set of objects (set A), separated by physical and virtual objects and ordered by increasing size. The bolded objects are incongruent in an office scene.

Set B Objects

Physical	Virtual
Sticky notes	Orange pencil holder
Binder clips	Small vase
Tape dispenser	Open book
Scissors	Charging phone with cable
Small desk clock	Desk light
Whiteboard markers	Electric kettle
Blue binder	Monitor
Stuffed camel toy	Large potted plant
Oven mitts	Printer
Standing floor lamp	***Pink framed mirror***
Trash can	Black chair
Swivel office chair	Water dispenser
Gray filing cabinet	White door cabinet
Bookshelf with textbooks	Folding table
Filing cabinet desk	Blue lounge chair
Gray storage basket with lid	***Treasure chest***

Fig. 2. The second set of objects (set B), separated by physical and virtual objects and ordered by increasing size. The bolded objects are incongruent in an office scene.

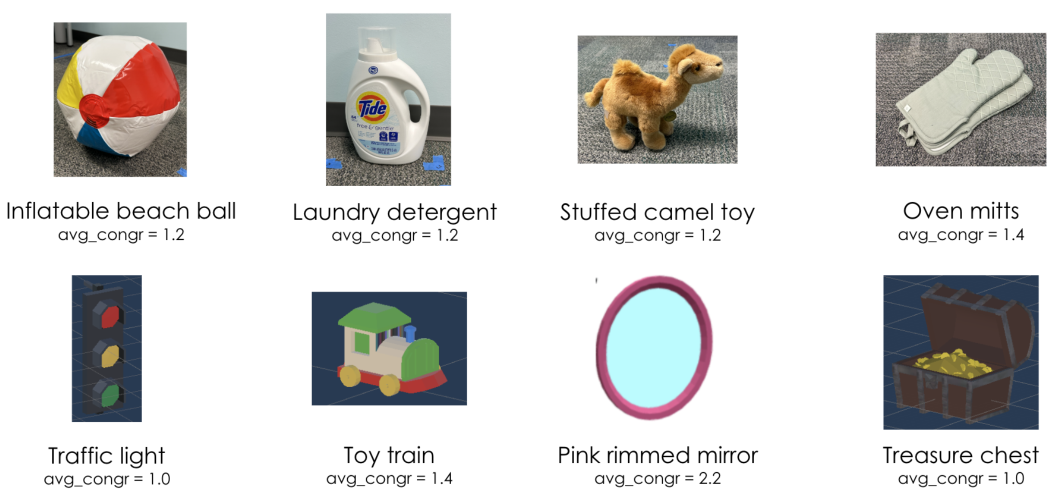

Fig. 3. The incongruent objects that appear across both sets of objects. Each of these objects received an average congruency ranking (avg_congr) based on our object congruency survey. Objects with average congruency ratings less than 2.5 were considered incongruent.

deeper processing. This insight directly informs our hypothesis about incongruent objects having a high likelihood of being remembered. In addition, cognitive load has been shown to have a negative effect on memory performance. Hanway et al. (2021) investigated how multitasking affected memory recall in an investigative interview task and found that high cognitive load diminished participants' ability to recall key information [7]. The results from this study imply that in scenarios with increased cognitive load and multitasking, working memory is constrained. Our study integrates these insights about human cognition with the interface of interactive AR. We investigate how memory and attention in augmented environments depends on interface specifications (like object congruence and other factors) and cognitive constraints (multitasking and cognitive load).

C. Study Design

In each trial, participants complete an indoor standing selection task where they are presented with a list of 16 objects and they must sequentially find and select as many objects as possible in the mixed reality office scene within 45 seconds. There are 8 trials total. Participants are instructed to complete a surprise memory recall task after the 4th trial and a cued memory recall task after the 8th trial.

1) Variables: The main independent variables we are concerned with in this study are cueing and cognitive load. We

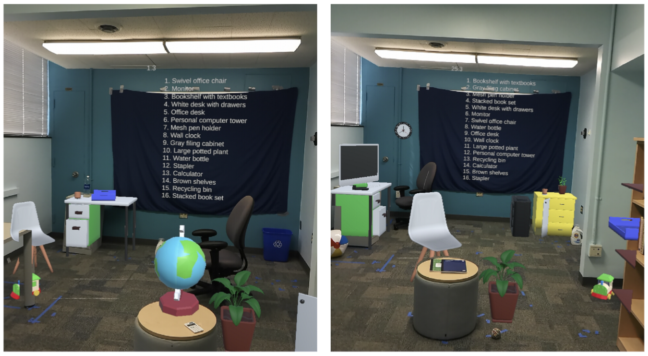

Fig. 4. The two layouts that correspond to set A objects. Layout 1 is pictured to the left, and layout 2 is pictured to the right. The objects of set A show up in different places in both layouts.

Fig. 5. The two layouts that correspond to set B objects. Layout 3 is pictured to the left, and layout 4 is pictured to the right. The objects of set B show up in different places in both layouts.

also investigate a number of ancillary object-level factors, including virtuality, congruency, and size. We define cueing as a verbal cue directing participants to remember their scene. Participants complete the task in the first 4 trials without knowledge of a memory recall task (no cue condition) first. Participants are then verbally cued (cue condition) to remember the objects in the scene, and complete another 4 trials. We use a simultaneous secondary audio task as a measure of cognitive load. Trials without the audio task are considered low cognitive load and trials with the audio task are considered high cognitive load. The cue condition is kept within subjects, so all participants experience the no cue and cued conditions in that exact order. For the audio task, participants either experience the audio task in the first block or second block of 4 trials.

2) Objects: We use two different sets (A, B) of objects in this study (Figure 1 and Figure 2). A set of objects consists of 32 objects total that appear in the scene: 16 physical objects and 16 virtual objects. The sets of objects are balanced in terms of virtuality (physical vs. virtual) and size (very small, small, medium, large) across virtuality. Each set contains 32 objects which are all of the objects we consider to be within the room where our study takes place. A set of objects contains 4 incongruent objects (2 real and 2 virtual). In this context, congruency determines whether an object is thematically congruent with an office space. For example, a beach ball would be considered incongruent in an office space, while a printer would be considered congruent. In order to determine our set of incongruent objects, we conducted an object congruency survey with 5 participants and had each participant rank how congruent they felt each object was with an office scene. The results from this survey are shown in Figure 3, and we use a threshold of $avg\_congruency < 2.5$ to determine incongruity. We also employ ChatGPT and the object congruency survey to determine the most universally understood descriptor names for the objects, and we use these consistently throughout the tasks.

3) Selection Task: In the selection task, participants are given 45 seconds to sequentially select up to 16 objects from a list provided to them. Participants are told that objects in the scene can show up anywhere within a 180 degree radius and on surfaces and the floor. The participants must search around the room, pointing at and selecting both physical and virtual objects with the controller. They see visual feedback of the bounding box around the object turning green on correct selection. Of the 32 objects that are part of the scene, 16 of these objects are part of the selection task.

4) Audio Task: For the audio task, participants are expected to listen to a randomized stream of 5 words from the NATO phonetic alphabet (alpha, bravo, charlie, delta, echo) that play at intervals of 3 seconds. Whenever they hear the word *charlie*, participants respond by pressing the menu button on top of the Magic Leap 2 controller.

5) Memory Recall Task: Participants are asked about 32 objects on the memory recall task. 16 of these objects were present in the scene (but not part of the selection task), and 16 objects were not part of the scene (absent). The intention is that participants are asked about peripheral objects in the scene that they did not interact with. For each object on the task, participants are asked if that object was present or absent, physical or virtual, and their associated confidence rating about their answers.

D. Methodology

We use the Magic Leap 2 for this study to leverage optical see through AR and depth preservation cues. Participants are primed with a presentation containing instructions for the selection task, but prior to completing the trials they are not informed that they will be completing a memory recall assessment. After fitting the headset comfortably, participants complete a short eye tracking calibration and are able to practice selection with the controller as well as the audio response task. Each trial is completed standing up within a designated box at the front of the room. Participants view the mixed reality office environment as a "museum display" and are not allowed to walk around the space to minimize the extra effort of walking.

The first 4 trials consist of the same set of objects but in two different layouts (Figures 4 and 5). After the 2nd trial,

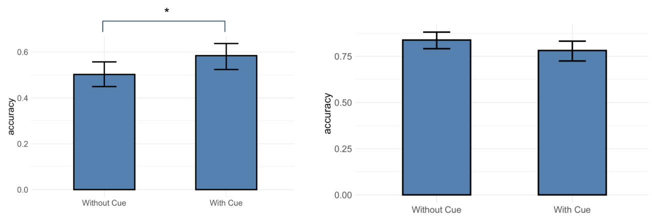

Fig. 6. Plots indicating the impact of the cue on recall accuracy filtered on present objects (left) and absent objects (right). When filtered on present objects, the difference before and after the cue is significant.

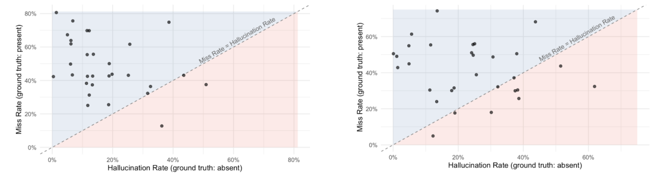

Fig. 7. Each participant's miss rate compared with their hallucination rate before the cue (left) and after the cue (right). Participants generally tended to have higher miss rates across both cue conditions: selecting absent when the object was present.

the layout changes: the same objects show up in different locations. We shuffle the positions of the objects randomly so that objects that are further away move closer and vice versa to minimize bias based on location of certain objects. The incongruent objects also move around the room to ensure they have an equally likely chance of being seen. Participants complete each layout twice with 2 different lists to mitigate bias from list difficulty

and to maximize visibility of objects in the scene. After the 4th trial, the set of objects changes while participants complete their first surprise memory recall assessment. They are asked about objects they did not directly interact with through selection.

The verbal cue is administered after the 4th trial once participants have completed the first memory recall assessment. Participants are allowed to view their performance on the first memory recall assessment to incentive performance for the second set of trials. The verbal cue is: *"In this next set of trials, while completing the selection task to the best of your ability within the time constraint, please focus on creating a mental model of the objects in the room."*

E. Participants

The study involved 32 participants aged 18-30 years, consisting of 18 female and 14 male. Of the 32 participants, 29 were right handed and 3 were left handed. 7 participants wore contacts for the study to have corrected-to-normal vision. Participants had varying familiarity with VR/AR, over half of the participants had not used augmented reality previously (56.3%). In comparison, 87.5% of participants had used virtual reality 1 or more times. Participants were compensated at $15 per hour. All procedures were approved by the University's Human Subjects Committee.

III. Results

In accordance with our main research questions, we divide our results into a few main areas: impact of the cue on recall, influence of the specific object-level factors on recall, and effect of cognitive load on selection and recall task performance.

A. All Factors and Interactions

We run an overall Repeated Measures ANOVA to test the main effects and their interactions with each other. The set of within subjects variables we test are: cue, $virtuality\_ground\_truth$, $congruence$, $size$, and $audio\ task$. From an overall ANOVA filtered on present objects from the memory recall assessment, we see that cue ($p < 0.05$), $size$ ($p < 0.01$), and $virtuality\_ground\_truth$ ($p < 0.05$) had a significant impact on recall individually. We also observe that the interaction effect of $virtuality\_ground\_truth$ and $size$ ($p < 0.01$) and $congruence$ and $size$ ($p < 0.001$) were highly significant, indicating that participants remembered objects differently based on the combination of the object level factors. However, we did not see any impact of the cue or audio task on these object level factors. The subsequent sections will dive deeper into each object level factor and their results.

B. Impact of Cue on Recall Accuracy

We notice that if we filter by present objects, the difference before and after the cue is significant with $p = 0.0353$, ($p < 0.05$), as shown in Figure 6. Another interesting observation was that when comparing each participant's miss rate against their hallucination rate, we see that participants are more likely to select absent when items were present (erring on the side of caution). Miss rate in this context is when an object is actually present but the participant thought it was absent. Hallucination rate is when the participant "hallucinated an object" or said present when the ground truth was absent. In general, most participants had a higher miss rate compared to hallucination rate. This inflates the average score on absent objects (if participants are selecting absent more often). Even after the cue, as shown in Figure 8, we see that more participants had a higher miss rate and selected absent more frequently.

When we filter by present objects on the memory recall assessment, we see that participants' accuracy without the cue is lower than with the cue ($p < 0.05$). This indicates that the cue did have an impact on accuracy and their memory of their scene, and participants remembered more objects in the scene after they were told to remember. On absent objects, it visually looks like after the cue the accuracy on absent objects decreased, however this change is not significant and could be attributed to random error.

For the majority of this analysis, we filter by present objects participants were asked about on the memory recall assessment. We analyze whether participants were able to

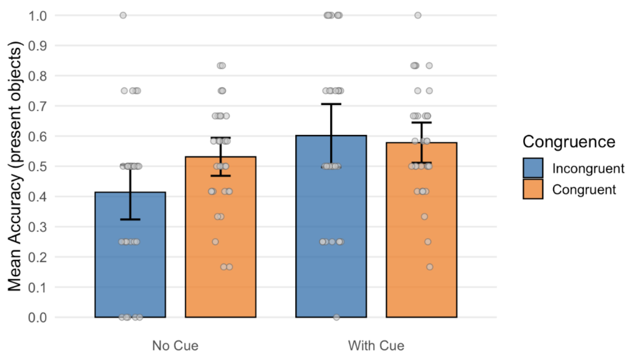

Fig. 8. The figure above shows the impact of cue on recall accuracy by **congruence**. We see that incongruent objects appear to have a visually larger increase in accuracy between cue conditions.

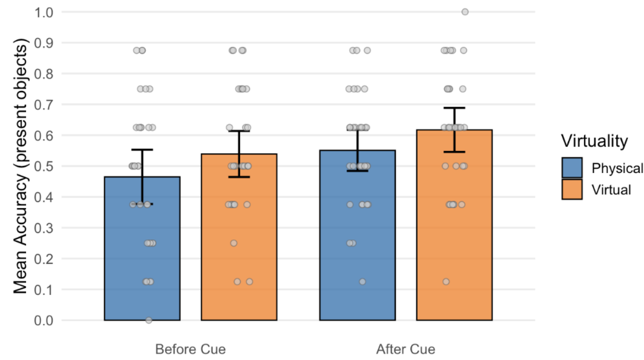

Fig. 9. This figure shows the impact of the cue on recall accuracy by **virtuality**. We can visually see that the virtual objects generally had a higher accuracy before and after the cue.

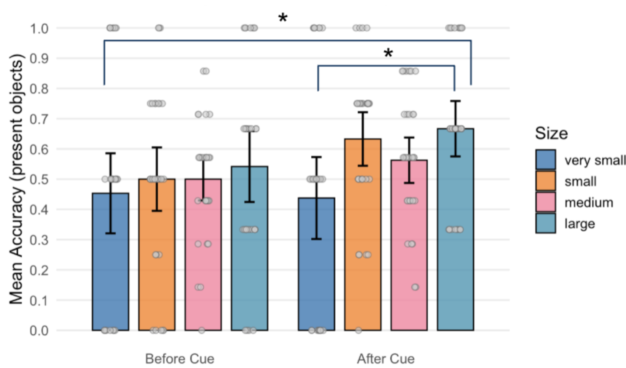

Fig. 10. This figure shows the impact of cue by **size**. There are 4 levels of size (very small, small, medium, and large). The difference between very small and large is significant.

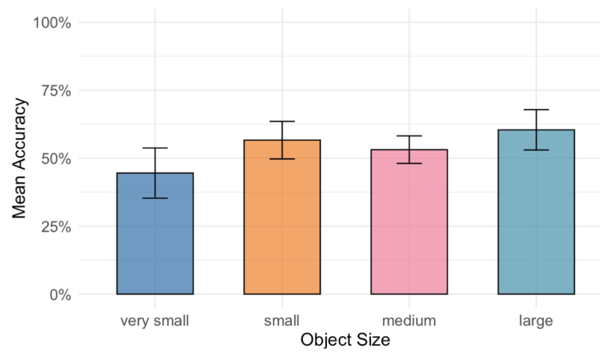

Fig. 11. Accuracy of recall averaged across all four size levels.

recall objects they were exposed to in the scene as opposed to absent objects (identifying what they did not see).

C. Influence of Object-Level Factors

1) Congruency: To further examine congruence, we run a 2-way repeated measures ANOVA on the cueing condition and object congruency, since both variables being tested were within subjects. We group each participant by cue and congruence level and (incongruent and congruent) and filtered by "Present" objects on the recall test. The results from our ANOVA indicate that the cue was significant (Figure 8), and the congruence across cue conditions was trending towards significance. There is an increasing trend in accuracy, specifically for incongruent objects after the cue.

2) Virtuality: Our analysis on virtuality examines differences between physical and virtual objects as part of the scene (Figure 9). A 2-way repeated measures ANOVA and post-hoc Tukey tests reveals that there was a significant difference between recall on physical and virtual objects. We can also visually see that virtual objects tended to have higher accuracy, indicating that participants paid less attention to their real-world surroundings in the AR interface. However, when they are cued, there is more attention to virtual objects in their scene. There is also no significant interaction between the two factors of cue and virtuality.

3) Size: Based on our ANOVA test and as shown in Figure 10, there are significant differences between sizes ($p < 0.01$). Running a post-hoc Tukey test on the size, we can see that these differences exist between the very small and large groups. The cue interaction is not significant but we can say that participants remember very small objects differently compared to large objects in this interface; it follows that large objects are easier to remember (Figure 11).

D. Impact of Cognitive Load

1) Selection Task Performance: We examine the impacts of the audio task on the primary selection task (Figure 12). We see that in trials without audio, the mean selection efficiency is visually higher than trials with audio. Based on our ANOVA test, we observe a trend towards significance with p = 0.0982. This trend indicates that the audio task did influence the cognitive load of participants, making the primary task efficiency lower. Participants were not able to find as many objects from the list of 16 when they had to multitask with the audio stimulus. Figure 13 also shows us that in each trial, the average selection efficiency with the audio task tended to

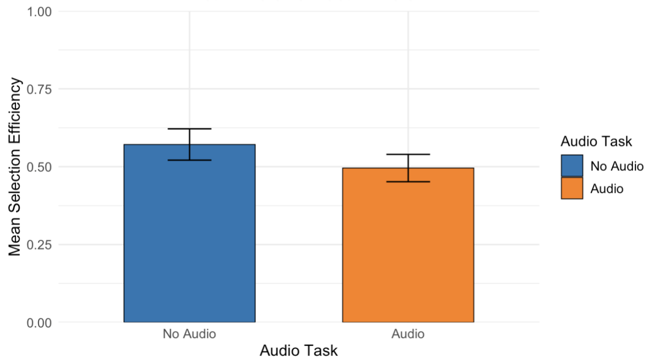

Fig. 12. Plot showing difference between selection efficiencies (number of objects correctly selected out of 16) by audio task condition. We can see that with the audio task, the selection task efficiency was trending lower with $p = 0.0982$ compared to without the audio task.

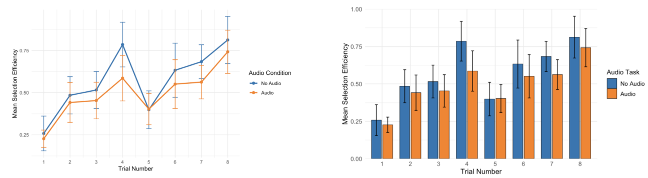

Fig. 13. Mean selection efficiency by trial. We can see that across trials the audio task visually tended to result in a lower selection efficiency. We also see a drop in efficiency after trial 4, when the set of object changes.

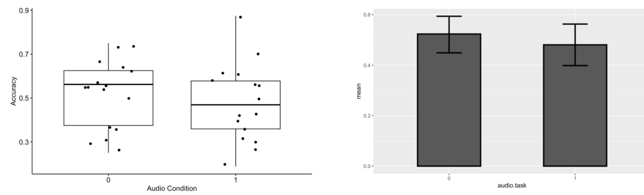

Fig. 14. Plots of the impact of the audio task on recall accuracy. Increased cognitive load with audio task has visual trend of decrease in recall accuracy.

be lower. We also notice a drop in selection efficiency after trial 4, and this can be attributed to the change in the set of objects.

2) Recall Accuracy: We run a simple t-test with two groups (the first group had audio in the first 4 trials and the second group did not have audio in the first 4 trials), to do a preliminary analysis of audio task on recall accuracy. We find that the results were not significant, but we can see a visual trend towards lower accuracy on the box plot in Figure 14 when the audio task is present. This indicates that with more data and a clear delineation of cognitive load, cognitive load could have an impact on recall.

IV. Discussion

Our findings demonstrate that memory recall of office objects in augmented reality (AR) is influenced by multiple object-level and task-level factors. Specifically, participants were more likely to recall objects that were large in size or virtual, which suggests that saliency has a clear impact on memory encoding. To a user unfamiliar with augmented reality, the novelty effect causes virtual objects themselves to feel incongruent, and thus showed an impact on participant's attention in the interface. Incongruent objects (items that stood out thematically relative to an office scene) were remembered more frequently after the cue, supporting our hypothesis that semantic incongruence influences attention and improves memory encoding. This aligns with prior research on schema and encoding in spatial memory.

Cueing had a measurable and selective effect on recall, significantly improving memory for present objects in the scene. However, this benefit did not extend to absent objects, indicating that cues primarily helped participants consolidate or retrieve accurate visual information rather than promoting guessing or biasing their responses. These results highlight the utility of guided attention mechanisms in AR environments, particularly when users are required to remember or revisit contextual details.

We also observed that the addition of an audio secondary distractor task increased cognitive load, resulting in significantly worse performance on the object selection task. While the effect on memory recall was less pronounced, there is a noticeable trend toward reduced recall accuracy under higher cognitive load. This suggests that multitasking in AR environments could impact a users' ability to retain spatial or object specific information, which becomes relevant when designing experiences that rely on memory and spatial navigation.

From a design perspective, these results carry important implications for AR interface development. Attention related cues, including visual and semantic cues, can provide support for users where encoding and recalling environmental details are involved. However, passive use of AR interfaces without intentional guidance may result in low recall, especially for large but congruent physical items that blend into the background (such as the floor lamp and standing fan in Figures 15 and 16). This underscores the need for AR systems to actively direct user attention toward critical elements rather than relying on visual presence alone.

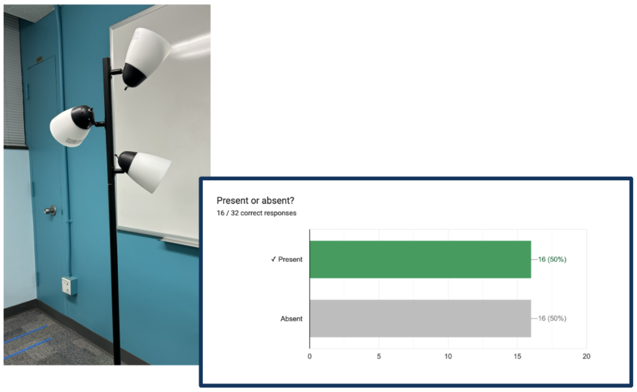

Fig. 15. Standing lamp: a large physical incongruent object. Exactly half of the participants were not able to recall the standing lamp.

Fig. 16. Standing fan: a large physical incongruent object. 71.9% of participants were not able to recall the standing fan.

Our findings necessitate the importance of designing AR systems with attentional guidance and minimal cognitive load. Such systems can not only enhance user memory and task performance but also mitigate risks associated with divided attention in mixed reality environments.

V. Conclusion

This study demonstrates that memory for office objects in augmented reality is influenced by object level features, cueing strategies, and cognitive load. Larger and virtual objects were generally easier to recall, and incongruent items were better remembered when cues were provided, which emphasizes the importance of attention and intentional cues. While cueing improved memory for present objects, absent objects did not benefit similarly, suggesting cues reinforce attention to objects within the scene. Furthermore, introducing an audio task imposed a measurable cognitive burden, reducing selection efficiency performance and trending toward a negative impact on recall accuracy.

These findings have implications for AR interface design: systems should incorporate attention informed elements while minimizing cognitive distractions to support user memory and task efficiency.

To probe cognitive load more rigorously, future iterations of the study could increase task difficulty, potentially leveraging a harder task like an N-back task, and examine effects across a range of cognitive load. Investigating both implicit (non-verbal) and explicit (verbal) cueing can reveal how different attention strategies can impact encoding. Additionally, controlling for potential learning effects across repeated selection trials will improve interpretation of performance trends. Incorporating physical walking during the task could serve as both an added cognitive load in a future study, which could help determine how movement affects memory and object interaction in AR contexts. Future research can further refine our understanding of human cognitive processes in AR and inform the design of adaptive, cognitively informed AR systems.

References

[1] Li, Keqin, et al. ”The application of augmented reality (ar) in remote work and education.” arXiv preprint arXiv:2404.10579 (2024).
[2] Yeung, Andy Wai Kan, et al. ”Virtual and augmented reality applications in medicine: analysis of the scientific literature.” Journal of medical internet research 23.2 (2021): e25499.
[3] Du, Zhao, Jun Liu, and Tianjiao Wang. ”Augmented reality marketing: A systematic literature review and an agenda for future inquiry.” Frontiers in psychology 13 (2022): 925963.
[4] Radha Kumaran*, You-Jin Kim*, Anne 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, 17 pages
[5] 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 Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, 16 pages
[6] Brewer, William F., and James C. Treyens. ”Role of schemata in memory for places.” Cognitive psychology 13.2 (1981): 207-230.
[7] Hanway, Pamela, et al. ”The effects of cognitive load during an investigative interviewing task on mock interviewers’ recall of information.” Legal and Criminological Psychology 26.1 (2021): 25-41.