The Effects of Visual Realism on Spatial Memory and Exploration Patterns in Virtual Reality

Conference Paper · November 2020

DOI: 10.1145/3385956.3418945

2 authors:

Jiawei Huang
Pennsylvania State University

Alexander Klippel
Wageningen University & Research

The Effects of Visual Realism on Spatial Memory and Exploration Patterns in Virtual Reality

Jiawei Huang*
jzh87@psu.edu
Department of Geography
Center for Immersive Experiences
Pennsylvania State University

Alexander Klippel
klippel@psu.edu
Department of Geography
Center for Immersive Experiences
Institute for Computational and Data Sciences
Pennsylvania State University

Figure 1: Left to right: Part of the living room to be memorized; Part of the living room with misplaced objects; A participant recalled the location of a guitar by putting it back (short introductory video: https://vimeo.com/454138851).

ABSTRACT

Understanding the effects of environmental features such as visual realism on spatial memory can inform a human-centered design of virtual environments. This paper investigates the effects of visual realism on object location memory in virtual reality, taking account of individual differences, gaze, and locomotion. Participants freely explored two environments which varied in visual realism, and then recalled the locations of objects by returning the misplaced objects back to original locations. Overall, we did not find a significant relationship between visual realism and object location memory. We found, however, that individual differences such as spatial ability and gender accounted for more variance than visual realism. Gaze and locomotion analysis suggest that participants exhibited longer gaze duration and more clustered movement patterns in the low realism condition. Preliminary inspection further found that locomotion hotspots coincided with objects that showed a significant gaze time difference between high and low visual realism levels. These results suggest that high visual realism still provides positive spatial learning affordances but the effects are more intricate.

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**; **Empirical studies in HCI**; **Human computer interaction (HCI)**.

VRST '20, November 1–4, 2020, Virtual Event, Canada

ACM ISBN 978-1-4503-7619-8/20/11...$15.00
https://doi.org/10.1145/3385956.3418945

KEYWORDS

Virtual reality, Spatial ability, Spatial memory, Visual fidelity, Visual realism

ACM Reference Format:
Jiawei Huang and Alexander Klippel. 2020. The Effects of Visual Realism on Spatial Memory and Exploration Patterns in Virtual Reality. In *26th ACM Symposium on Virtual Reality Software and Technology (VRST '20), November 1–4, 2020, Virtual Event, Canada.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3385956.3418945

1 INTRODUCTION

One of the ultimate goals of virtual reality (VR) is to be able to simulate the real world as convincingly as possible. Fidelity captures this aspect and is defined as the extent to which a virtual environment (VE) and interactions are indistinguishable from a real environment [47]. Considering a VE that depicts an imaginary world, the concept of fidelity can be extended to the plausibility of any virtual scene, which may or may not reflect the real world. Fidelity consists of display fidelity, interaction fidelity, and simulation fidelity [27]. Visual realism is an important component of display fidelity. There are a number of factors contributing to visual realism: lighting, geometry, shadow softness [21, 35], rendering techniques [24, 25], and texture quality. They can be roughly categorized as geometric realism and illumination realism [39].

With the advances in computer graphics, which result in ever-increasing visual realism, it is both theoretically and practically desirable to provide empirical evidence of the potential effects of visual realism. High visual realism and low visual realism provide different learning affordances. High visual realism facilitates perception of details, concretization of abstract concepts, and experiential learning in scenarios connecting to real life. Low realism promotes abstract thinking, high-level comprehension and generalization of concepts [17]. Further, the need for high visual realism and low visual realism often varies by fields. Education and training often benefit from high visual realism for knowledge transfer. Virtual surgery, for example, often uses highly-detailed representations of the body; natural environments [16][11] and historic architecture [15] visualizations may also benefit from high realism. Certain visual arts, games and scientific visualizations, on the other hand, may benefit from low visual realism taking the forms of stylized arts, symbolization and abstraction. To make things more complex, naive realism points out a potential mismatch between preference and performance [40]. In terms of spatial cognition, there is less clear evidence regarding the effects of visual realism [23], while such insights are crucial to guide the design of VE to present space to the audience and facilitate spatial learning. Theoretically, we find that spatial cognition and visual realism are linked: elements that are important for achieving visual realism such as texture, light, and shadows coincide with cues that are important for depth and spatial location: Texture can provide information about distances [38] and depth [30]; shadow can provide information of spatial depth and object location. In terms of the relationship between spatial memory and visual realism, Kettunen et al. [20] suggest that human spatial knowledge may be influenced by vantage point, the number of visible vertical features, and visual realism. Complexity is also thought of as a major determinant of attention. When people are not occupied with a task, spontaneous looking emerges which selects stimuli that are evolutionarily likely to be significant [19]. These stimuli are physical properties (such as presence of many contours), and collative properties, such as novelty and complexity [3, 4].

Object location memory is an important component of spatial cognition [44]. Research from neuroscience proposed three processing stages in object location memory: (a) object processing, (b) spatial-location processing, and (c) object to location binding [34]. What happens in the object processing stage is interesting to look at: for example, flowers of all kinds are called flowers, but are also recognized as having different visual features. Some researchers therefore proposed two dissociable subsystems in the brain: abstract-category recognition, which maps different inputs to the same output representation and categorizes objects in the absence of visual details. Second, the specific-exemplar recognition which maps different inputs, even fairly similar ones, to different outputs and preserves visual details of objects to distinguish specific exemplars. In the object processing stage, the abstract-category and the specific-exemplar object recognition neural subsystems work in parallel to recognize an object [26, 28]. Therefore, it is possible that by increasing visual realism, the specific-exemplar object recognition enhances, which could result in better object location memory.

There have only been a few studies of visual realism in the field of spatial cognition [49]. The literature also showed unclear relationships between visual realism and spatial task performance. In a 3D non-VR supermarket environment, participants in the photorealistic group showed higher accuracy and lower time used in a route reproduction and a scene recognition tasks compared to the nonrealistic group, but no effects were found in a map identification and a route drawing tasks [29]. In another study, participants in the high visual realism environment performed better in a wayfinding task, sketch-mapping task and picture-sorting task [49]. Lokka and Çöltekin [23], on the other hand, found the participants had the highest route recall accuracy in a middle-realism VE compared to an abstract VE and a realistic VE. The VEs were represented as videos. Another study suggests that individual differences such as gender, prior computer use, and cognitive ability accounted for more variance in performance on tasks requiring spatial knowledge acquisition from a desktop VE than visual realism of the VE [48].

Studies regarding visual realism and object location memory specifically have been even more sparse: Mania et al. [25] found no significance in object location memory in high and low visual realism conditions which varied in rendering techniques and colors (colored vs. grey). In their VR task, participants recalled the shape of 3D objects in each numbered position and reported their confidence level. Murcia-López and Steed [33] did not find overall significance either in an object location memory task where participants recalled the location of three objects. They also created heatmaps of navigational patterns under high and low realism levels but no quantitative analysis was conducted. Some earlier studies will not be discussed here due to the technological limitations resulting in lower visual realism overall, as noted by several authors (e.g. [9]).

Based on the above theories and empirical studies, we hypothesize that higher visual realism increases object location memory, and such an effect is mediated by individual characteristics such as spatial abilities. We conducted a study with two visual realism levels. The level of visual realism was controlled in two aspects: illumination realism and geometric realism. Geometric realism was further manipulated with polygon count and texture resolution [18]. Our study has the following characteristics: 1. We designed a more robust task where participants had to reconstruct an environment without any information, compared to some previous studies that matched objects with given locations. The environments were also relatively complex which possessed abundant visual information in the forms of objects and details (Fig. 2) 2. With VR, we were able to obtain and quantitatively analyze other objective metrics, such as gaze and locomotion which shed light on participants' attention and exploration patterns in VR that are important for spatial cognition and performance.

The rest of the article is organized as follows: We first describe the environment design, experiment procedure and data collected in Section 2; we then present the statistical analysis results in Section 3. In Section 4, we summarize the statistical implications; we discuss future research directions in spatial cognition, trajectory analysis, gaze analysis, and visual realism quantification in Section 5. Finally, we summarize the findings in Section 6.

2 METHOD

In order to analyze the effects of visual realism on spatial memory, we used a two-level mixed group design with counterbalancing. Participants were randomly placed into four groups, where each participant went through two different visual realism levels in two environments, a living room and a kitchen, respectively, in order to remove carry-over effects: (1) group 1: high visual realism living room followed by low visual realism kitchen. (2) group 2: high visual realism kitchen followed by low visual realism living room. (3) group 3: low visual realism living room followed by high visual realism kitchen. (4) group 4: low visual realism kitchen followed by high visual realism living room.

2.1 Participants

Twenty student participants (9 female, 10 male, and 1 transgender female, mean age: 21.3, age range: 18-29) were recruited. The participant inclusion criteria were (self-reported): 1. normal or near-normal vision. If not, participants should wear contact lenses. 2. healthy 3. not cognitively impaired 4. 18-30 years old.

2.2 Apparatus

We selected two daily environments to mimic the real-life object location memory scenarios. The environments were a living room and a kitchen, each with two different visual realism levels (Fig. 2). There were a total of 48 moveable objects and other immovable objects scattered in each room. Among them, 19 objects were intentionally misplaced in the living room, and 21 objects were intentionally misplaced in the kitchen for the participants to reconstruct. The study was conducted using the HTC Vive Pro. The environments were developed in the Unreal Engine 4.21.2 to achieve high graphics quality in the high visual realism level. The high visual realism was photorealistic and took advantage of the DirectX 11 pipeline that included deferred shading, global illumination, lit translucency, and various post-processing effects [10]. From a high realism environment to a low realism environment, polygon count was reduced by 95%; texture size was reduced to 1/256 of the original size; rendering quality was reduced with the engine scalability settings in the Unreal Engine 4. The reduction in polygon count only affected surface details, leaving the general shape of the objects intact, so that participants could complete the task with no difficulty in recognizing objects. We noticed the scene became darker with lighting and shadow calculation that resulted in darker places under high realism, whereas there was homogeneous lighting in the low realism level. But such difference did not cause difficulty to recognize objects. We aimed to control visual realism quantitatively and keep the changes the same for different parts of the environment. The VE was generated with the assets [1] from the Unreal Engine 4 marketplace. The teleporting and object grabbing interactions were implemented so that users could freely teleport in the space and used natural grabbing to intuitively interact with the objects. A moving area restriction was also implemented to prevent users from making physical movement, so that their locomotion can be precisely recorded using software tools.

The objects to be memorized were everyday small-sized objects such as cups, toys, laptops, or books. Efforts were made to ensure multiple possible locations to place the objects to lower the possibility that participants would use semantics to speculate on correct locations. In addition, objects were non-intrusive (i.e., in logical positions). There were three types of misplacement: (1) topological changes: objects were put into different surroundings, that is, the reference objects have changed; (2) displacement: objects were still on the same object, but the relative location has changed; (3) orientation changes: objects orientations were changed.

2.3 Procedure

The walk-through video of the experiment can be found online (https://vimeo.com/435231081). At the start of the experiment, participants were informed that their task was to complete a set of spatial memory tasks. They were not informed about the visual realism differences. Upon entering the experiment, participants were automatically placed in a virtual bedroom which was the training room. The self-paced training included learning the VR interactions through reading the instruction on the wall, and performing a task. The task was a mini-version of the actual experiment that was shorter in time and simpler in difficulty to prepare the participants. In this phase, the experimenter monitored the participants' actions on the screen, and offered help or correction, although the experimenter would not offer information during the main experiment. Participants were able to enter the real experiments after they got familiar with the VR interaction. The result from the pilot study showed that all participants were able to master VR interactions after this training phase.

After the training phase, the main experiment started and participants entered the learning phase. In the learning phase, participants explored the room and memorized objects locations and rotations for the duration of four minutes. There was a timer on a table that counted down. The duration of this phase was determined through a pilot study. They were not informed on how many or which objects will be misplaced in the room. They could teleport freely in the room but not interact with the objects. After the four minutes, participants were automatically taken to an empty virtual space where the short-term memory formed into long-term memory during the duration of 30 seconds [1]. Next, participants entered the reconstruction phase. They were asked to reconstruct the room, i.e., to place objects into their original locations and rotations. Participants could not only freely teleport in the space, but also interact with the objects by picking them up and placing them down. Participants had the maximum of ten minutes in this task. They may exit early if they think they have completed the reconstruction. The duration was determined according to a pilot study where most participants finished in ten minutes or less. After completing the first environment, participants completed questionnaires on a desktop computer. The questionnaires were 1. Demographics questionnaire; 2. Presence questionnaire; 3. VR system evaluation; 4. Short answers questionnaire that asked what strategies they used to memorize the objects. Participants then completed the second half of the experiment where they repeated the learning phase and the reconstruction phase for the second environment. After participants went through both VR environments, they completed post-questionnaires on a desktop computer. The questionnaires were 1. short answers questionnaire that asked about whether they have noticed any differences between the two environments. 2. spatial working memory test 3. Object location memory test.

[1]Dream Apartments: https://www.unrealengine.com/marketplace/en-US/product/dream-apartments, HQ Residential House https://www.unrealengine.com/marketplace/en-US/product/hq-residential-house, Houseplant Pack: https://www.unrealengine.com/marketplace/en-US/product/houseplant-pack, Timers, Clocks and Counters Pack: https://www.unrealengine.com/marketplace/en-US/product/timers-clocks-and-counters-pack. VR Integrator radial and dockable menus: https://www.unrealengine.com/marketplace/en-US/product/vr-integrator-radial-and-dockable-menus

Figure 2: left to right, top to bottom: high realism living room, high realism kitchen, low realism living room, low realism kitchen.

2.4 Metric

2.4.1 Response variables. In order to analyze the performance quantitatively, we collected the x, y, z coordinates and the roll, pitch, yaw of each interactable object after the reconstruction phase.

2.4.2 Player variables. (1) Gaze data. We used a plug-in [2] in the Unreal Engine to deduct gaze position from ray tracing from the center of the eye. We also tracked the start and end time of the gaze to calculate gaze duration. (2) Player locomotion data. Player locomotion was tracked through the in-game player character location along with the time stamps.

2.4.3 Moderator variables. (1) Reported presence. Presence data was collected through the spatial presence experience scale (SPES) [14]. (2) VR system evaluation. We changed the qualitative VR heuristic evaluation [42] to a five point Likert scale questionnaire (Table 3 in the Appendix). (3) demographics (age, gender, major, VR experience, game experience) [6, 22]; (4) object location memory: We measured the individual difference with object location memory test using the online version [13] developed based on the original version [37]. The test is in 2D on a computer screen, in which participants study an array of objects for a designated period of time, and the array of objects will then disappear. Next, the array of objects will reappear, but some of the objects will have exchanged positions. Participants then select the objects that have moved. Participants have five trials to complete the task. (5) Spatial working memory. We also collected spatial working memory through the Corsi block-tapping test that we developed on Unity based on the original version [8].

[2]Virtual Reality Pawn and Components Plugin: https://www.unrealengine.com/marketplace/en-US/product/vr-pawn-components-plugin

3 RESULTS

3.1 Spatial memory

3.1.1 Placement error. We used two different measures to evaluate participants' performance and the first is the placement error. We calculated the Euclidean distance between the original position and the new location placed by participants for all 48 objects and calculated the average as the error metric for each participant. We derived the placement error for each participant pooling the living room and kitchen data, since there was no significant difference between the two rooms. A paired t-test was then conducted to compare the error between levels of visual realism. There was no significant difference ($t(19) = -.728, p = .476$) in error for high visual realism ($n = 10, M = 37.574, SD = 19.665$), and low visual realism ($n = 10, M = 41.015, SD = 19.677$). Units are centimeters.

3.1.2 Categorized error. We considered the possibility that multiple types of errors may incur during the reconstruction in deriving the placement error. In order to remove confounding errors, we restored the room as participants left and categorized the error semantically into five categories using the method adapted from the object location memory error categories by Silverman and Eel [37]. We added topological error which was not included in the original categories, since the nature of their task does not incur this type of error. These five categories are (1) correct: participants correctly identified that the object had been moved, and moved it to a close location; (2) topological error: correctly identified object had been moved, but moved it to a wrong location; (3) correct rejection: correctly identified that the object had not been moved and left the object at the original place; (4) false alarm: Incorrectly determined that the object had been moved; (5) miss: incorrectly identified that the object did not move. We considered both the original room and also the misplaced room in order to categorize errors. Fig. 3 shows the complexity of possible errors. We conducted a Chi-square analysis which showed no difference between high and low fidelity ($\chi^2(4, 20) = 1.668, p = .797$ for the living room, $\chi^2(4, 20) = 1.820, p = .769$ for the kitchen).

Figure 3: Left to right, top to bottom: original locations; messy room; the construction result of one player.

Based on the new error category, we also summed up (1) - (3) and subtracting (4) - (5) to derive the composite measure according to Silverman and Eel [37]. The composite measure and the placement error were also strongly correlated ($r = -.586$) which indicated that the two error metrics both reflected participants' performance. The placement error appeared to be a better metric because it was continuous from the plots; it also captured the performance difference better, as we observed different participants having the same categorized error but different placement errors. The result from 3.1.1 and 3.1.2 indicate that there was neither a difference in the overall error, nor in any of the error subcategories.

3.1.3 Relationship between task performance and different explanatory variables. There was a negative relationship ($r = .694$ for high visual realism, and $r = .623$) between placement error and 2D object location memory (tested with the same apparatus from Silverman and Eel [37]), indicating that people with higher 2D object location memory also performed better in 3D (Fig. 4). Participants' performance did not appear to differ in high and low visual realism (Fig. 4). The 2D object location memory score was calculated using the discrimination index proposed by Banks [2].

Figure 4: Placement error and 2D object location memory, showing kitchen and living room performance separately for each participant.

In order to investigate the overall relationship between multiple variables, we conducted multiple linear regression with the dependent variable of placement error, and independent variables of age, gender, major, game experience, object-location memory, spatial working memory, and visual realism level, for living room and kitchen separately. Then we used backward selection to remove variables with the highest p-value, one at a time, to find a model with both high r^2 and low overall p-value. The final regression model after variable selection is summarized in Table 1 and Table 2 (see the appendix). The results again indicated that people with higher 2D memory test scores performed better as above. It is also shown in the table that women outperformed men in the kitchen.

3.2 Gaze

Current gaze analysis and visualization methods do not adequately support eye tracking in VEs [41]. We did preliminary analysis with the gaze data focusing on gaze duration. We aggregated the gaze duration for all interactable objects for each participant to calculate the total gaze duration of objects. An unpaired t-test revealed a marginally significant higher gaze duration ($t(18) = -1.980, p = .063$, Fig. 5) in the living room for the low visual realism ($n = 10, M = 48.938, SD = 13.429$) compared to the high visual realism ($n = 10, M = 39.180, SD = 7.915$). Note that this was the gaze during the learning phase instead of the reconstruction phase, since we are more interested in how participants explored the place during memorization. Specifically, in the living room, participants gazed on the following objects significantly longer in the low visual realism version: cups, family photo, golden bowl, phone, white pillow, white vase, and wine. However, we did not find such a relationship ($t(18) = -.616, p = .546$) in the kitchen for high visual realism ($n = 10, M = 30.232, SD = 8.247$) and low visual realism ($n = 10, M = 33.768, SD = 16.167$) (Fig. 5).

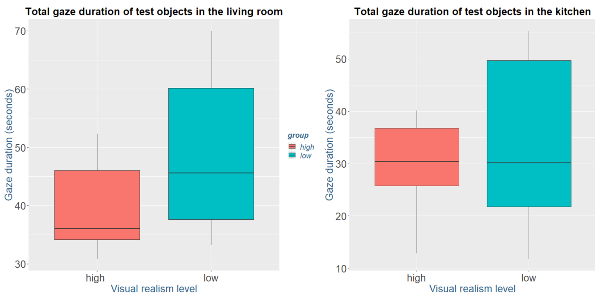

Figure 5: Gaze duration and visual realism levels in the living room (left) and in the kitchen (right).

3.3 Locomotion

Locomotion hotspots can be analyzed using cluster analysis of point locations which resulted from teleportation. We performed Chi-Square Significance Test and variance-to-mean Ratio (VMR) calculations over the locomotion data (Fig. 8 and 9 in the appendix). The VMR score is a numerical indicator of the spatial distribution of a given point pattern. In the living room, we found the point pattern to be slightly clustered in the high visual realism condition based on a (χ^2= 199.67, df = 155, p = .018; VMR score = 1.374) and moderately clustered in the low visual realism condition (χ^2 = 260.6, df = 155, p = 4.585$e-07$; VMR score = 1.795). In the kitchen, we also saw a more clustered locomotion pattern in the low visual realism condition: the points were moderately clustered in the high visual realism condition (χ^2= 248.07, df = 173, p = 3.171$e-4$; VMR score = 1.736) and were significantly clustered in the low visual realism condition (χ^2 = 287.67, df = 173, p = 1.984$e-07$; VMR score = 2.003). The result is consistent with the gaze data. The clusters emerged in the place where there were significant differences in gaze between high and low visual realism, that is, where cups, family photo, golden bowl, phone, white pillow, white vase, and wine were.

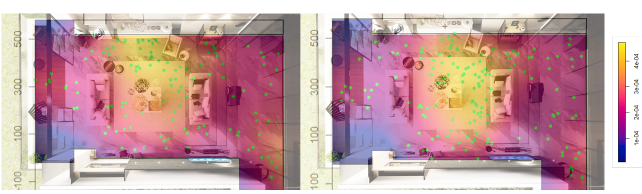

Figure 6: Participants' movement in the living room by aggregating the data from all participants. Left: high realism; right: low realism.

4 DISCUSSION

We designed a VR reconstruction task to evaluate spatial memory in high and low visual realism environments. Participants memorized object locations and moved misplaced objects back to their original location. Visual realism, according to our results, is not a significant factor influencing object location memory. However, individual differences such as 2D object location memory and gender accounted for more variance in performance than the fidelity differences. This

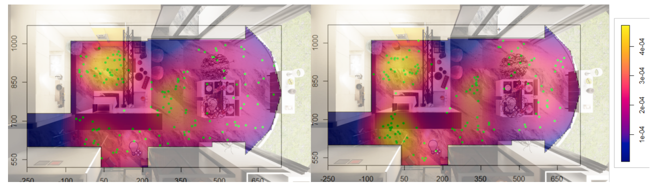

Figure 7: Participants' movement in the kitchen by aggregating the data from all participants. Left: high realism; right: low realism.

finding is interesting as there are several differences between the original object location memory test [37] and our reconstruction task: 1) the former was in 2D and the latter was in VR 2) in the object location memory test, participants only need to point out the objects that are misplaced; in our reconstruction task, they also need to remember where they were and place them at the correct location. 3) By letting participants put the object anywhere they remember, participants may make topological errors as mentioned in 3.1.2. Despite these differences, the results indicate that people perform better in the object reconstruction task if they have higher scores in 2D object location memory independent of environmental features such as the visual realism level. This finding aligns with other studies that found individual differences played a more critical role than visual fidelity in navigation tasks [48], and visual realism did not affect identifying object locations [25].

We also found that women outperformed men in the kitchen environment, which aligns with the general finding that women outperform men in object location memory task, as noted in a meta-analysis studies [46], despite that men outperform women in almost all other spatial tasks [6].

The result that there was no significant difference across conditions does not necessarily exclude the possibility that visual realism impacts object location memory. Given that participants in the high visual realism group used less effort in both gazing and locomotion, but did not perform worse than the low visual realism group, we speculate that high visual realism still helps people to memorize objects. The correspondence between the gaze data and the locomotion data could indicate that people tend to stay at a location and gaze-memorize. This could be related to their memorizing strategy. When asking about memorizing strategies, 43% of participants mentioned about counting. The hotspot of locomotion could be their "counting place".

The small number of participants could obscure the performance results. Further, given the objects used in the study were relatively simple in geometry and texture, participants might mainly utilize a viewpoint-abstract, category-tuned object-recognition system to memorize objects instead of attending to details. Although the objects were more complex than previous experiments that used primitive geometries such as cylinders and spheres, they still might not be complex enough for visual realism to show an effect. Verbal memory triggered by familiarity is another impacting factor in visual recognition and visual memory. Object location memory is confounded by verbal memory, particularly when everyday objects are used [5]. The literature has also discussed that semantic knowledge affects the visual processing of objects [7]. When asked about memorization strategies, 57% of participants mentioned heavy usage of tying semantics to object location memory, such as: "(I) tried making personal anecdotes to remember where things went," "(I) imagine why things are put in certain locations." As much as we tried to decrease the reliance on semantics by creating multiple logical locations to place objects, the results show how difficult it is to decrease semantics' use when participants entered familiar environments (kitchen and living room), and memorized daily objects.

Also, sensitivity to realism is quite different from person to person. Some participants commented that visual realism was drastically different from high to low. However, about 40% of participants finished the entire experiment without noticing any difference. Sensitivity to visual realism may be reflected in cognitive style, which is the tendency to perceive details from the surrounding environment. Cognitive style influences spatial memory potentially. For example, a study showed that participants who were better at perceiving the environment as a whole outperformed participants who focused their attention on single environmental features in an object location memory task [43]. Another possibility is that the difference in high and low visual realism may be too subtle. Further, the complex lighting created shadows and made the high realism environments generally darker (Fig. 2) that may cause lower performance.

5 FUTURE WORK

VR studies have the advantage of tracking and generating substantial behavior data. As a future directive, using machine learning methods could generate new insights analyzing user behavior and performance in VR. The locomotion and gaze results presented here are preliminary. Locomotion trajectory, in addition to point pattern, could also be analyzed. Trajectory analysis is an active research topic in spatial-temporal data analysis, although not frequently applied to VEs. However, there have been methods proposed to analyze locomotion trajectories in VEs [32]. Other than traditional statistical methods such as K-means, dynamic time warping (DTW), machine learning methods such as neural networks have also been used [36, 51]. Identifying linkages between spatial ability and locomotion pattern is a critical research question for the future that might provide additional insights [36]. Time, too, as an additional dimension, can support in-depth analyses. Future analysis could also consider 3D visualization and hotspot analysis of the gaze data [41].

Given the possible interference of verbal memory discussed in Section 4, future studies in spatial cognition may consider using complex, novel, or abstract objects to decrease semantic influences.

In terms of VR applications design, visual realism is one of the most essential and exciting environmental features. In previous empirical studies, different methods were used to control visual realism. In some studies, visual realism was manipulated in art style, such as from cartoonish to a realistic style [45]. Other studies addressed the presence or absence of scene objects (e.g., presence of items on grocery shelves, [29]). Whether artistic style is part of visual realism is an open question. For example, a cartoonish style can be of high visual realism, as seen in recent animated movies which used state-of-art facial motion capture; Removing scene objects may remove specific visual cues that lead to a decrease in information input. Some other studies changed one aspect of visual realism, such as comparing specific rendering models (e.g., ray casting v.s. real-time recursive ray tracing [39]; global v.s. local illumination [52]; flat-shaded v.s. radiosity rendering [25], or geometric realism [18]. Others changed the extent of visual realism. For example, textures might be removed entirely in the low visual realism condition [9].

Quantifying and measuring visual realism has been an active topic in computer graphics following two approaches: automated computational prediction/computation-based, and subjective human judgment/perception-based. Automated computational prediction quantitatively calculates the score induced by global illumination and artifacts given an ideal reference image [31]. Subjective human judgment typically involves conducting lab experiments to measure realism perception vs. rendered images/scenes [12, 35, 50]. In future studies, we argue that it is essential to standardize visual realism definitions and measures. Standardization is necessary to control precisely and draw broadly-applicable conclusions, whether it is to study the effects of visual realism on spatial cognition, or on other dependent variables that are commonly associated with visual realism such as presence, empathy, or emotions. Multidisciplinary knowledge is needed to fully understand the relationship between vision, locomotion patterns, individual differences, object location memory, and VR environmental features. We hope that the findings of this study will open up research avenues for further explorations.

6 CONCLUSION

Summarizing the study findings, we did not see a significant relationship between visual realism and object location memory. On the other hand, spatial ability showed significance in both environments as a predictor of performance; and, gender showed significance in the kitchen when females performed better. The correspondence between the gaze data and the locomotion data showed that participants' locomotion synchronized with their observation changes. We also detected that participants under low visual realism gazed at objects significantly longer than participants in the high visual realism conditions in the living room. Considering that their performance was the same across conditions, participants may still benefit from high visual realism, which might provide learning affordance that helped them memorize quickly. In terms of locomotion, lower visual realism led to a higher clustering effect, suggesting a tendency to restrain locomotion and active exploration under low realism.

ACKNOWLEDGMENTS

We wish to thank undergraduate student Joe Nadoley who helped with the gaze and motion data processing, and Jianglan Shi who helped with the gaze data preprocessing. We also wish to thank Dr. Jack (Shen-Kuen) Chang for providing feedback in the early development stage of this project.

REFERENCES

[1] Richard C Atkinson and Richard M Shiffrin. 1968. Human memory: A proposed system and its control processes. *Psychology of learning and motivation* 2, 4 (1968), 89–195.

[2] William P. Banks. 1970. Signal detection theory and human memory. *Psychological Bulletin* 74, 2 (1970), 81.

[3] D. E. Berlyne. 1960. *Conflict, arousal and curiosity.* McGraw-Hill., New York.

[4] D. E. Berlyne. 1966. Curiosity and Exploration. *Science* 153, 3731 (1966), 25–33.

[5] Jean Choi and Natasha L'Hirondelle. 2005. Object location memory: A direct test of the verbal memory hypothesis. *Learning and Individual Differences* 15, 3 (2005), 237–245. https://doi.org/10.1016/j.lindif.2005.02.001

[6] M. L. Collaer and M. Hines. 1995. Human behavioral sex differences: A role for gonadal hormones during early development? *Psychological Bulletin* (1995), 55–107.

[7] Jessica A. Collins and Ingrid R. Olson. 2014. Knowledge is power: how conceptual knowledge transforms visual cognition. *Psychonomic Bulletin & Review* 21, 4 (2014), 843–860. https://doi.org/10.3758/s13423-013-0564-3

[8] Philip M. Corsi. 1973. *Human memory and the medial temporal region of the brain.* Doctoral dissertation.

[9] H. Q. Dinh, N. Walker, L. F. Hodges, Chang Song, and A. Kobayashi. 1999. Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Proceedings IEEE Virtual Reality.* 222–228. https://doi.org/10.1109/VR.1999.756955

[10] Epic Games. 2020. Rendering Overview. https://docs.unrealengine.com/en-US/Engine/Rendering/Overview/index.html

[11] Marek Fabrika, Peter Valent, and Ľubomír Scheer. 2018. Thinning trainer based on forest-growth model, virtual reality and computer-aided virtual environment. *Environmental Modelling & Software* 100 (2018), 11–23. https://doi.org/10.1016/j.envsoft.2017.11.015

[12] Shaojing Fan, Tian-Tsong Ng, Bryan Lee Koenig, Jonathan Samuel Herberg, Ming Jiang, Zhiqi Shen, and Qi Zhao. 2018. Image Visual Realism: From Human Perception to Machine Computation. *IEEE transactions on pattern analysis and machine intelligence* 40, 9 (2018), 2180–2193. https://doi.org/10.1109/TPAMI.2017.2747150

[13] Kathleen A. Flannery and Marianna Eddy. 2020. Object Location Memory. https://opl.apa.org/experiments/index.html?#/Experiments

[14] Tilo Hartmann, Werner Wirth, Holger Schramm, Christoph Klimmt, Peter Vorderer, André Gysbers, Saskia Böcking, Niklas Ravaja, Jari Laarni, Timo Saari, Feliz Gouveia, and Ana Maria Sacau. 2016. The Spatial Presence Experience Scale (SPES). *Journal of Media Psychology* 28, 1 (2016), 1–15. https://doi.org/10.1027/1864-1105/a000137

[15] Jiawei Huang, Mahda M. Bagher, Heather Dohn Ross, Nathan Piekielek, Jan Oliver Wallgrün, Jiayan Zhao, and Alexander Klippel. 2018. From Archive, to Access, to Experience−−Historical Documents as a Basis for Immersive Experiences. *Journal of Map & Geography Libraries* 14, 1 (2018), 40–63. https://doi.org/10.1080/15420353.2018.1498427 arXiv:https://doi.org/10.1080/15420353.2018.1498427

[16] Jiawei Huang, Melissa. S. Lucash, Robert. M. Scheller, and Alexander Klippel. 2019. Visualizing Ecological Data in Virtual Reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR).* IEEE, Piscataway, NJ, 1311–1312. https://doi.org/10.1109/VR.2019.8797771

[17] Jiawei Huang, Melissa S. Lucash, Mark B. Simpson, Casey Helgeson, and Alexander Klippel. 2019. Visualizing Natural Environments from Data in Virtual Reality: Combining Realism and Uncertainty. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR).* IEEE, Piscataway, NJ, 1485–1488. https://doi.org/10.1109/VR.2019.8797996

[18] J. Hvass, O. Larsen, K. Vendelbo, N. Nilsson, R. Nordahl, and S. & Serafin. 2017. Visual realism and presence in a virtual reality game.. In *2017 3DTV Conference: the True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON).* 1–4.

[19] Daniel Kahneman. 1973. *Attention and Effort.*

[20] Pyry Kettunen, Katja Irvankoski, Christina M. Krause, Tapani Sarjakoski, and L. Tiina Sarjakoski. 2012. Geospatial images in the acquisition of spatial knowledge for wayfinding. *Journal of Spatial Information Science* 5 (2012). https://doi.org/10.5311/JOSIS.2012.5.85

[21] Cha Lee, Gustavo A. Rincon, Greg Meyer, Tobias Höllerer, and Doug A. Bowman. 2013. The effects of visual realism on search tasks in mixed reality simulation. *IEEE transactions on visualization and computer graphics* 19, 4 (2013), 547–556. https://doi.org/10.1109/TVCG.2013.41

[22] Eric L. G. Legge, Christopher R. Madan, Enoch T. Ng, and Jeremy B. Caplan. 2012. Building a memory palace in minutes: equivalent memory performance using virtual versus conventional environments with the Method of Loci. *Acta Psychologica* 141, 3 (2012), 380–390. https://doi.org/10.1016/j.actpsy.2012.09.002

[23] Ismini E. Lokka and Arzu Çöltekin. 2019. Toward optimizing the design of virtual environments for route learning: empirically assessing the effects of changing levels of realism on memory. *International Journal of Digital Earth* 12, 2 (2019), 137–155. https://doi.org/10.1080/17538947.2017.1349842

[24] K. Mania and A. Robinson. 2004. The Effect of Quality of Rendering on User Lighting Impressions and Presence in Virtual Environments. In *Proceedings of the 2004 ACM SIGGRAPH international conference on Virtual Reality continuum and its applications in industry.* 200–205.

[25] Katerina Mania, Dave Wooldridge, Matthew Coxon, and Andrew Robinson. 2006. The effect of visual and interaction fidelity on spatial cognition in immersive virtual environments. *IEEE transactions on visualization and computer graphics* 12, 3 (2006), 396–404. https://doi.org/10.1109/TVCG.2006.55

[26] Chad J. Marsolek. 1999. Dissociable Neural Subsystems Underlie Abstract and Specific Object Recognition. *Psychological science* 10, 2 (1999).

[27] Ryan Patrick McMahan. 2011. *Exploring the Effects of Higher Fidelity Display and Interaction for Virtual Reality Games.* Doctor of Philosophy. Virginia Polytechnic Institute and State University.

[28] Brenton W. McMenamin, Rebecca G. Deason, Vaughn R. Steele, Wilma Koutstaal, and Chad J. Marsolek. 2015. Separability of abstract-category and specific-exemplar visual object subsystems: evidence from fMRI pattern analysis. https://doi.org/10.1016/j.bandc.2014.11.007

[29] Frank Meijer, Branko L. Geudeke, and Broek, Egon L. van den. 2009. Navigating through Virtual Environments: Visual Realism Improves Spatial Cognition. *Cyberpsychology & behavior* 12, 5 (2009).

[30] Lou Michel. 1996. *Light: The shape of space: designing with space and light.* Van Nostrand Reinhold, New York.

[31] Bochang Moon, Nathan Carr, and Sung-Eui Yoon. 2014. Adaptive Rendering Based on Weighted Local Regression. *ACM Transactions on Graphics* 33, 5 (2014), 1–14. https://doi.org/10.1145/2641762

[32] Hazarath Munaga, Sree, M. D. R. Mounica, and J. V. R. Murthy. 2012. DenTrac: A Density based Trajectory Clustering Tool. *International Journal of Computer Applications* 41, 10 (2012).

[33] María Murcia-López and Anthony Steed. 2016. The Effect of Environmental Features, Self-Avatar, and Immersion on Object Location Memory in Virtual Environments. *Frontiers in ICT* 3 (2016), 1. https://doi.org/10.3389/fict.2016.00024

[34] Albert Postma, Roy P. C. Kessels, and Marieke van Asselen. 2008. How the brain remembers and forgets where things are: the neurocognition of object-location memory. *Neuroscience and biobehavioral reviews* 32, 8 (2008), 1339–1345. https://doi.org/10.1016/j.neubiorev.2008.05.001

[35] Paul Rademacher, Jed Lengyel, Edward Cutrell, and Turner Whitted. 2001. Measuring the Perception of Visual Realism in Images. In *Rendering Techniques 2001.* 235–247. https://doi.org/10.1007/978-3-7091-6242-2{_}22

[36] Corina Sas. 2004. *Individual Differences in Navigating and Experiencing Presence in Virtual Environments.* PhD. National University of Ireland, Dublin.

[37] I. Silverman and M. Eals. 1992. Sex differences in spatial abilities: Evolutionary theory and data. In *The adapted mind: Evolutionary psychology and the generation of culture.* Oxford University Press, New York, 533–549.

[38] M. J. Sinai, W. K. Krebs, R. P. Darken, J. H. Rowland, and J. S. McCarley. 1999. Egocentric Distance Perception in a Virutal Environment Using a Perceptual Matching Task. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting.* 1256–1260.

[39] Mel Slater, Pankaj Khanna, Jesper Mortensen, and Insu Yu. 2009. Visual Realism Enhances Realistic Response in an Immersive Virtual Environment. *IEEE Computer Graphics and Applications* 29, 3 (2009), 76–84. https://doi.org/10.1109/MCG.2009.55

[40] Harvey S. Smallman and Mark St. John. 2005. Naive Realism: Misplaced Faith in Realistic Displays. (2005).

[41] S. Stellmach, L. Nacke, and R. & Dachselt. 2010. Advanced Gaze Visualizations for Three-dimensional Virtual Environments: Symposium on Eye Tracking Research & Applications ; Austin, TX, USA, March 22 - 24, 2010. In *Proceedings of the 2010 symposium on eye-tracking research & Applications.*

[42] Alistair Sutcliffe and Brian Gault. 2004. Heuristic evaluation of virtual reality applications. *Interacting with Computers* 16, 4 (2004), 831–849. https://doi.org/10.1016/j.intcom.2004.05.001

[43] Laura Tascón, Maddalena Boccia, Laura Piccardi, and José M. Cimadevilla. 2017. Differences in Spatial Memory Recognition Due to Cognitive Style. *Frontiers in pharmacology* 8 (2017), 550. https://doi.org/10.3389/fphar.2017.00550

[44] Marieke van Asselen, Roy P. C. Kessels, L. Jaap Kappelle, and Albert Postma. 2008. Categorical and coordinate spatial representations within object-location memory. *Cortex; a journal devoted to the study of the nervous system and behavior* 44, 3 (2008), 249–256. https://doi.org/10.1016/j.cortex.2006.05.005

[45] V. Vinayagamoorthy, A. Brogni, M. Gillies, M. Slater, and A. and Steed. 2004. An Investigation of Presence Response across Variations in Visual Realism. In *Presence.* 148–155.

[46] D. Voyer, A. Postma, B. Brake, and J. & Imperato-McGinley. 2007. Gender differences in object location memory: A meta-analysis. *Psychonomic Bulletin & Review* 14, 1 (2007), 23–38.

[47] David Waller, Earl Hunt, and David Knapp. 1998. The Transfer of Spatial Knowledge in Virtual Environment Training. *Presence: Teleoperators and Virtual Environments* 7, 2 (1998), 129–143. https://doi.org/10.1162/105474698565631

[48] D. Waller, D. Knapp, and E. & Hunt. 2001. Spatial Representations of Virtual Mazes: The Role of Visual Fidelity and Individual Differences. *Human factors* 43, 1 (2001), 147–158.

[49] Grégory Wallet, Hélène Sauzéon, Prashant Arvind Pala, Florian Larrue, Xia Zheng, and Bernard N'Kaoua. 2011. Virtual/real transfer of spatial knowledge: benefit from visual fidelity provided in a virtual environment and impact of active navigation. *Cyberpsychology, behavior and social networking* 14, 7-8 (2011), 417–423. https://doi.org/10.1089/cyber.2009.0187

[50] Norman Wang and Wendy Doube. 2011. How Real is Really? A Perceptually Motivated System for Quantifying Visual Realism in Digital Images. In *2011 International Conference on Multimedia and Signal Processing*. 141–149. https://doi.org/10.1109/CMSP.2011.172

[51] Di Yao, Chao Zhang, Zhihua Zhu, Jianhui Huang, and Jingping Bi. 2017. Trajectory clustering via deep representation learning. In *2017 international joint conference on neural networks (IJCNN)*. 3880–3887. https://doi.org/10.1109/IJCNN.2017.7966345

[52] Insu Yu, Jesper Mortensen, Pankaj Khanna, and Mel Slater. 2011. Visual realism enhances realistic response in an immersive virtual environment – Part 2. *IEEE Computer Graphics and Applications* (2011). https://doi.org/10.1109/MCG.2011.69

A ADDITIONAL FIGURES AND TABLES

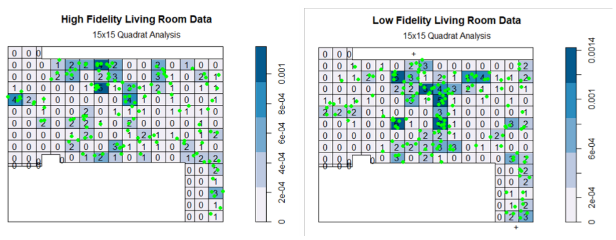

Figure 8: VMR analysis showing clusters of locomotion in the living room.

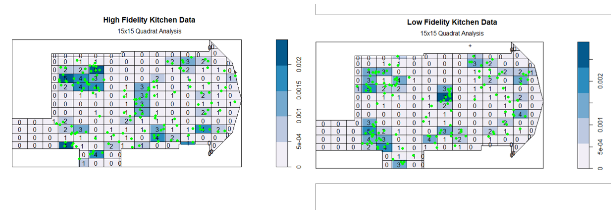

Figure 9: VMR analysis showing clusters of locomotion in the kitchen.

Table 1: Multiple linear regression model for the living room after backward selection

Living room					
Terms	Estimate	Error	t value	P value	Signif. Codes
(Intercept)	110.063	12.383	8.888	8.46e-08	***
T4ADI	-23.940	5.563	-4.303	4.81e-04	***
T3ADI	-37.089	7.743	-4.790	1.7e-04	***

Residual standard error: 13.56 on 17 degree of freedom
Multiple R-squared: .662, Adjusted R-squared: .622
F-statistic: 16.63 on 2 and 17 DF, p-value: 9.968e-05
T1ADI – T5ADI were spatial ability measures; high realism level was coded as 1 in is_high_realism; Similar with is_male and is_STEM

Table 2: Multiple linear regression model for the kitchen after backward selection

Kitchen					
Terms	Estimate	Error	t value	P value	Signif. Codes
(Intercept)	52.796	13.78	3.831	.002	**
Is_male	17.911	5.623	3.185	.006	**
Is_STEM	8.194	5.391	1.520	.149	
T2ADI	-30.331	13.737	-2.208	.043	*

Residual standard error: 11.49 on 15 degrees of freedom
Multiple R-squared: .654, Adjusted R-squared: .562
F-statistic: 7.097 on 4 and 15 DF, p-value: .002
T1ADI – T5ADI were spatial ability measures; male is coded as 1 in is_male; Similar with is_STEM

Table 3: VR system evaluation adapted from VR heuristic evaluation[42]

Heuristics	Questionnaire item
Natural engagement	• The overall graphics look good (for example, aesthetics, color schemes, lightings, 3D models details, textures, no serious lag or flickering). • I think the interactions with the menus were intuitive. • I think the interactions with the environment and objects were intuitive.
Compatibility with the user's task and domain	• The virtual environment is close to my expectation of real world environment. There was no unexpected objects/events that contradict my knowledge and/or physics laws (for example, floating objects, etc.). • I feel the interactions were compatible with the tasks needed to be performed.
Natural expression of action	• My body representation allowed me to act and explore in a natural manner. • I had no issue with the hardware (e.g., headset, earphones, straps, controllers).
Close coordination of action and representation	• The system responded to my actions smoothly and without delay.
Realistic feedback	• The effects of my actions were immediately visible and conform to the laws of physics and my perceptual expectations.
Faithful viewpoints	• The visual representation of the virtual world mapped to my normal perception. • The viewpoint change by head movement was rendered without delays that could impact my overall experience.
Navigation and orientation support	• I was able to know where I was in the virtual environment and was able to navigate from place to place.
Support for learning	• I feel the system provided support for learning of the virtual environment.
Clear turn-taking	• N/A