EXAMINING EYE MOVEMENTS IN RESPONSE TO REPEATED EXPOSURE TO BUILDING STIMULI

MAXYM SERGIYOVY YERKEYEV

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILLMENT OF OHE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

GRADUATE PROGRAM IN PSYCHOLOGY

YORK UNIVERSITY

TORONTO, ONTARIO

August 2024

Abstract

Research has consistently indicated a strong connection between eye movements and memory processes. This thesis explores this relationship by examining how repeated exposure to AI-generated images of buildings influences eye movements and memory. Twenty-four participants viewed 120 building images across four levels of repetition (novel, once, three, and five times), while their eye movements were recorded, and their memory assessed. The results showed significant repetition effects on both eye movements and memory. Eye movement measures revealed a decrease in fixation count and saccadic amplitude, and an increase in fixation duration, with increased repetitions. Memory measures revealed improved recognition and confidence, with increased repetitions. These repetition effects align with previous studies on faces and scenes, suggesting that the effects supersede differences in how specific object categories are processed. Overall, this thesis demonstrates that memory and oculomotor systems are associated in processing buildings, just as they are in processing faces and scenes.

Table of Contents

List of Figures

Chapter 1: Introduction

In a typical state, our eyes move at an average of three saccades per second (Tatler et al., 2014). These rapid movements accumulate visual information, enabling us to form and retrieve visual memories (Damiano & Walther, 2019). Visual memory holds great importance as it is vital to many essential functions, such as spatial navigation. We use visual memory to learn the geometry of new spatial environments and to recognize the visual appearance of familiar landmarks that can aid in navigation. For example, a visually salient building situated at an intersection of an unfamiliar street can act as a landmark, providing directional cues and preventing us from getting lost. The efficacy of visual memory critically depends on the interactions between the oculomotor and memory systems, which are anatomically and functionally linked, and collaborate in encoding and retrieval of visual information (Ryan & Shen, 2020; Shen et al., 2016). Despite the importance of visual memory, the exact nature of how visual memories are encoded, stored, and retrieved, is not well understood.

One method to bridge this gap in understanding is through eye tracking research involving visual memory tasks. Previous eye tracking studies have demonstrated that memory formation and retrieval are associated with distinct eye movement expressions (Hannula & Ranganath, 2009; Johansson et al., 2022; etc.). These studies have used eye tracking to probe memory processes, and their findings have provided some insight into how visual memory representations are formed, stored, and subsequently retrieved.

For example, multiple eye tracking studies involving repeated exposure to face, scene, and everyday object stimuli have revealed that, during initial viewing, a greater number of fixations and a more dispersed distribution of fixations throughout the image, are associated with stronger recognition memory upon subsequent viewing (Kafkas & Montaldi, 2011; Olsen et al., 2016; Ramey et al., 2020). However, when eye movements are restricted during initial viewing, recognition memory becomes weaker (Damiano & Walther, 2019). These findings suggest that visual stimuli across different visual categories are remembered better if they are explored more extensively (Mikhailova et al., 2021). Although it seems intuitive that the more we look at things, the better we remember them, these findings demonstrate how eye movements help construct visual representations in memory.

Further evidence that eye movements can be used to index memory comes from findings of a memory-dependent eye movement effect known as the repetition effect. This effect is characterized by a negative correlation between the frequency of stimulus viewings and the number of fixations made upon subsequent viewings (Mazloum-Farzaghi et al., 2023). The repetition effect was found for both face and scene stimuli, suggesting that the more we look at visual stimuli, the less we tend to explore them when we re-encounter them. Additionally, the repetition effect was present regardless of whether participants explicitly recognized the stimuli (Olsen et al., 2016; Smith & Squire, 2017). The presence of the repetition effect in the absence of explicit recollection suggests that implicit memory is involved in the formation and retrieval of visual representations, or that visual memory may be a form of implicit memory (Ryals et al., 2015; Olsen et al., 2016).

Other research has revealed that greater overlap between fixation positions made during initial viewing and repeated viewing is associated with memory strength (Nikolaev et al., 2023). This effect is known as the resampling effect. Ramey et al. (2020) found that a greater magnitude of exploration overlap is associated with enhanced familiarity strength for scenes, while others found the effect to be associated with improved recollection strength (Damiano & Walther, 2015; Johansson et al., 2022). Taken together, greater exploration overlap seems to be associated with better overall memory performance, but the interpretation of this effect remains a subject of ongoing debate. Some have proposed that resampling represents the reactivation of the motor memory traces that are stored along with perceptual memory traces made during initial viewing (Noton & Stark 1971), while others have argued that resampling reflects the process of reinstating the spatiotemporal context of scenes (Robin, 2018).

In contrast to scenes, the resampling effect has not been found for faces. Although an exploration overlap between encoding and retrieval of faces has been previously found, it was not linked to stronger memory (Chuk et al., 2017). The absence of the resampling effect in faces may be explained by the fact that eye movements tend to overlap in face exploration (Blais et al., 2008, Caldara et al., 2010) due to the similarity between the components (eyes, nose, mouth) and the configurations of all faces in general, which, in turn, diminishes the resampling effect. Additionally, the resampling effect may be absent in faces because faces and scenes are processed by distinct neural mechanisms; scenes activate the parahippocampal place area (PPA), while faces activate the face fusiform area (Epstein, 2008; James et al., 2010). Faces are also processed holistically as unified objects, whereas scenes typically consist of multiple objects, their spatial configurations, and background spatial context (Epstein & Baker, 2019). Lastly, resampling might only apply to scenes because scenes possess a greater amount of spatial information compared to faces. Overall, the absence of the resampling effect in faces indicates that there are stimulus category-related differences in eye movement behavior.

The stimulus-specific difference in eye movement expression of memory signals the importance of stimulus category in eye tracking research. While faces and scenes have received considerable attention in eye tracking studies, another category of stimuli, buildings, has never been investigated in this context. Building stimuli, which refer to images of architectural structures, can be considered a subset of scenes. However, the underlying neural mechanism involved in buildings is posterior to and separate from the more broadly focused scene-processing area, and the PPA's response to buildings is smaller than that of scenes (Haxby et al., 2000; Epstein, 2008). While buildings contain spatial contextual information, much like scenes, buildings also represent unified, context-free, holistic entities, more akin to faces. In this sense, buildings fall somewhere between scenes and faces. What makes buildings completely unique, however, is the fact that they are large, non-living, space-defining entities, with the potential to be recruited as navigational landmarks (Bastin et al., 2013; Cate et al., 2011; Mullally and Maguire, 2011; Troiani et al., 2014). In sum, despite having shared similarities and differences with faces and scenes, it remains uncertain whether buildings exhibit the same eye movement effects found for faces and scenes.

To understand if the relationship between memory and eye-movements in the processing of buildings is more similar to faces or scenes, it is necessary to first establish whether buildings elicit repetition effects similar to those observed in other visual categories. My thesis aims to explore this inquiry by investigating how eye movements in response to building stimuli are modulated by repeated exposure. To this end, I probed eye movements while novel and repeated building stimuli are presented, with every trial followed by recollection and familiarity memory judgments. By investigating changes in eye movement behavior, and evaluating their relationship with memory, I aim to provide insight into how visual memories are encoded, stored, and retrieved.

Chapter 2: Method

2.1 Participants

Twenty-five participants were recruited at York University. The sample size was selected to provide more than 80% power to detect the weakest effect of memory on eye movements, based on a prior study (Ramey et al., 2020). One participant was excluded due to technical difficulties with the eye tracking program at testing. Participants were screened to have normal or corrected-to-normal vision, and had no neurological or psychological conditions. The study was approved by the York University Human Participant Review Committee (HPRC).

2.2 Design

The experiment used a within-subjects design. The independent variable was the level of image familiarity (4 levels: novel, repeated once, repeated three times, repeated five times), manipulated by presenting images repeatedly. The dependent variables were the fixation count, fixation duration, saccadic amplitude, recognition accuracy, and confidence rating.

2.3 Materials and Apparatus

The experiment was conducted in a lab setting using a PC desktop computer and a 24" monitor, with a resolution of 1920x1080, and a refresh rate of 120 Hz. Participants were seated at a distance of 20" (50 cm) away from the screen, and a chin rest was used to reduce head movements. Eye movement behavior was recorded using an Eyelink 1000+ eye-tracker, which tracks ocular movements at 1000 Hz. SR Research software (SR Research, 2010) was used to display stimuli and record participants' ocular behavior.

The stimulus set consisted of 120 images of buildings generated using MidJourney (MidJourney, 2023), an AI image generative program. The program combines various combinations of words, or prompts, to generate images. The language prompts were produced by

ChatGPT 3.5 (OpenAI, 2023). First, ChatGPT was asked to "generate a list of 50 adjectives describing visual and architectural features of buildings," which produced a list of words such as 'Minimalist' and 'colorful'. ChatGPT was then asked to "generate a list of 50 building types," producing a list of words such as 'Recording Studio' and 'Arts and Culture Center'. The two lists were then used to generate 100 prompts by randomly combining three adjectives with one building type. The prompt used was: "Generate 100 prompts using the format /imagine [Photograph adjective, adjective, adjective building], using adjectives from List 1 and buildings from List 2 in random order." This produced prompts such as:

1. /imagine [Photograph Aesthetic, angular, tranquil Arts and Culture Center]
2. /imagine [Photograph Post-war, colorful, brick-clad duplex]
3. /imagine [Photograph Brick, Colonial, intertwined building]
4. /imagine [Photograph Minimalist, concrete, utilitarian power plant]
5. /imagine [Photograph Steep stepped, sleek Recording Studio]

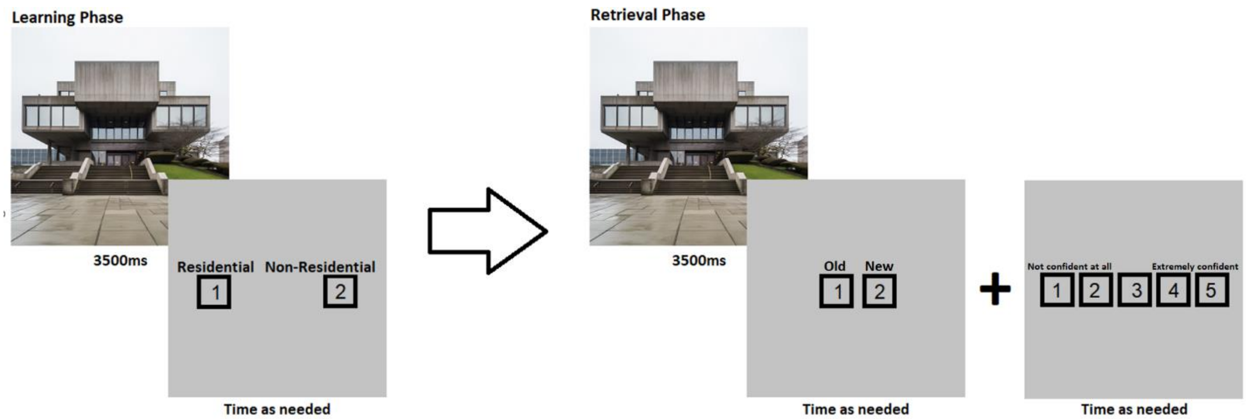
These language prompts were subsequently used to generate 400 images of buildings in MidJourney. Lastly, 120 building images were manually selected based on specific criteria, ensuring they depict a unified, whole building (e.g., an apartment building) that is photorealistic, fully visible, and unobstructed by other objects. All images maintained a resolution of 1024 x 1024 pixels.

2.4 Procedure

The experiment consisted of a study phase and a retrieval phase, separated by a one-minute break during which the eye-tracker was recalibrated. On each trial, a fixation cross was presented for 1000ms, followed by the presentation of a stimulus for 3500ms. After each stimulus presentation, participants were given as much time as they need to make building-function judgment (residential or non-residential?) in the study phase, and recollection and familiarity judgments in the test phase.

Figure 1

Example of an image presented in the learning and retrieval phases.



2.4a *Learning Phase*

At the start of the learning phase, the eye tracker was calibrated, and participants were presented with instructions and 4 practice trials. Following, a total of 60 building images were presented over 180 trials: 20 buildings were presented once (Category 1), 20 were presented 3 times (Category 2), and 20 were presented 5 times (Category 3). All images were presented in random order and counterbalanced across participants. Each trial was succeeded by a building-function judgment asking participants to decide whether the building is "Residential" or "Non-residential". Participants were told that there are no right or wrong answers, and were given as much time as they needed to respond by pressing buttons on a keyboard.

2.4b *Retrieval Phase*

In the retrieval phase, 120 buildings images were presented, 20 of which were repeated images from Category 1, 20 repeated from Category 2, 20 repeated from Category 3, and 60 were completely novel. Each stimulus presentation was followed by a recognition judgment and a confidence rating. All stimuli were presented in random order and counterbalanced across participants.

Following each trial, participants made a recognition judgment indicating whether the building was "Old" (previously seen) or "New" (never before seen) using a keyboard. Immediately afterward, participants provided a confidence judgment by selecting a number on a scale from one to five, where one indicated "Not confident at all" and five indicated "Extremely confident”.

2.5 Statistical Analysis

Statistical analyses were performed using the JASP statistical analysis program (JASP Team, 2024). The eye movement data during the retrieval phase of the experiment were collected using SR Research Data Viewer (SR Research, 2010), which provides eye movement metrics such as fixation count, fixation duration, and saccadic amplitude. Recognition accuracy and confidence ratings during the retrieval phase were recorded using SR Research Experiment Builder (SR Research, 2010), which logged participants’ keyboard responses.

A repeated-measures ANOVA was conducted on each dependent variable across four levels of repetition to assess the repetition effect with increasing repetitions. Additionally, to ensure the consistency of the repetition effect across the entire image set, the mean slope of fixation counts across the images was calculated, and a one-sample t-test was performed to determine if the mean slope significantly differed from zero.

Chapter 3: Results

In this study, we aimed to uncover novel insights into the role of eye movements in visual memory processes, particularly within the context of building stimuli. To this end, we manipulated memory through repeated exposures to building stimuli while tracking eye movements. The independent variable was the number of repeated presentations (novel, 1, 3, 5) of building images, and the dependent variables included recognition accuracy, confidence rating, fixation count, duration of fixation, and saccadic amplitude. We applied repeated-measures ANOVA to each variable across the four levels of repetitions. Finally, we conducted an item-analysis to examine whether the reported effects were consistent across the different items in our stimulus set.

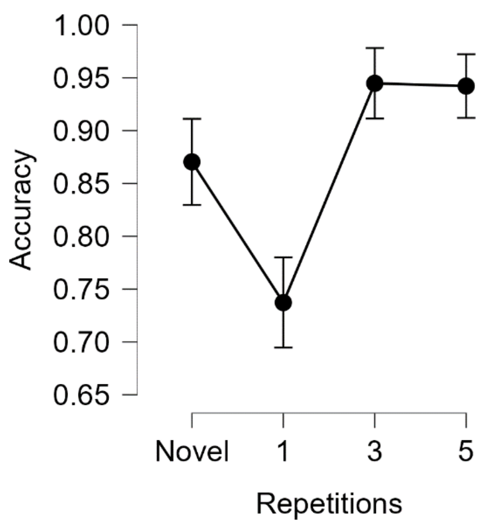
3.1 Recognition Accuracy

To measure recognition accuracy, participants were asked to indicate whether they recognized the stimulus as old (previously encountered) or new (not previously encountered). Figure 2 visualizes the differences in mean recognition accuracy across four levels of repetitions. Mean accuracy was highest for three and five repetitions ($M = 0.94$). In contrast, accuracy was lowest for one repetition ($M = 0.74$). For novel stimuli, accuracy was moderately high ($M = 0.87$). A repeated measures ANOVA revealed a highly significant main effect, $F(3, 69) = 29.427$, $p < .001$, $\eta^2 = 0.561$. These findings indicate that greater repetitions improve recognition accuracy.

To better understand the main effect, two contrasts were conducted. The first contrast evaluates whether recognition accuracy for the novel condition significantly differs from 1, 3 and 5 repetitions, combined. The second contrast evaluates whether the recognition accuracy for 1 repetition significantly differs from 5 repetitions, similar to linear regression analysis. Contrast results showed no significant difference for Comparison 1 (Estimate = 0.013, SE = 0.062, t = 0.207, p = 0.836), suggesting that the novel condition does not differ significantly from the repetition conditions combined. However, Comparison 2 showed a significant effect (Estimate = 0.205, SE = 0.025, t = 8.083, p < .001), indicating that recognition accuracy for 1 repetition is significantly lower than for 5 repetitions.

Figure 2

Mean accuracy across repetition levels (novel, 1, 3, 5). The error bars on the plot represent 95% confidence intervals (CI).



3.2 Confidence Rating

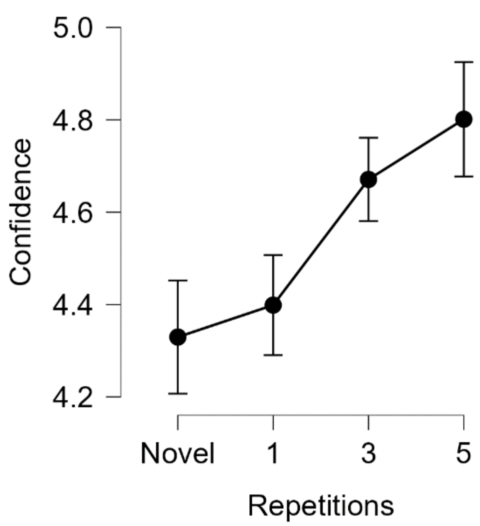
To measure memory confidence, participants were asked to indicate their level of confidence that they had seen the stimulus before on a scale from 1 to 5, with 1 indicating not confident at all, and 5 indicating extremely confident. Participants showed the highest mean confidence for 5 repetitions (M = 4.801), and the lowest mean confidence for novel stimuli (M = 4.33). Figure 3 visualizes the increase in confidence across the four levels of repetitions. Repeated-measures ANOVA revealed a statistically significant main effect of repetitions, F(3,

69) = 16.952, $p < .001$, $\eta^2 = 0.424$. Thus, as repetitions increased, participants showed higher confidence in their responses.

To better understand the main effect, two contrasts were conducted. The first contrast assessed whether confidence ratings for the novel condition differed from 1, 3, and 5 repetitions, combined. The second contrast evaluated whether confidence ratings for 1 repetition significantly differed from 5 repetitions. Contrast results showed a significant difference for Comparison 1 (Estimate = 0.882, SE = 0.188, t = 4.702, $p < .001$) and Comparison 2 (Estimate = 0.403, SE = 0.077, t = 5.255, $p < .001$).

Figure 3

Mean confidence across repetition levels (novel, 1, 3, 5). The error bars on the plot represent 95% confidence intervals (CI).



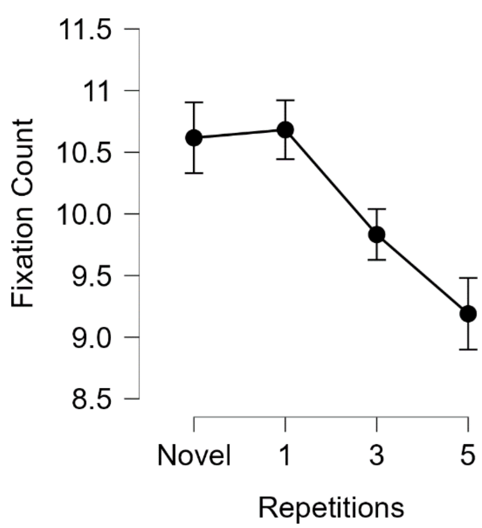
3.3 Fixation Count

The fixation count represents the number of fixations made during the 3500ms viewing of each stimulus. Figure 4 visually depicts a gradual decrease in mean fixation count with increasing repetitions. The fixation count for stimuli repeated once was the highest (M = 10.683), and the fixation count for novel stimuli was similar (M = 10.617). However, as repetitions increased to three, the fixation count decreased (M = 9.833), and decreased further with five repetitions (M = 9.19). A repeated-measures ANOVA revealed a significant main effect, $F(3, 69) = 32.252$, $p < .001$, $\eta^2 = 0.584$, signifying the presence of a repetition effect.

To better understand the main effect, two contrasts were conducted. The first contrast assessed whether fixation counts for the novel condition differed from 1, 3, and 5 repetitions, combined. The second contrast evaluated whether fixation counts for 1 repetition significantly differed from 5 repetitions. The results revealed a significant difference for Comparison 1 (Estimate = -2.146, SE = 0.432, t = -4.966, $p < .001$) and Comparison 2 (Estimate = -1.493, SE = 0.176, t = -8.463, $p < .001$).

Figure 4

Fixation count across repetition levels (novel, 1, 3, 5). The error bars on the plot represent 95% confidence intervals (CI).



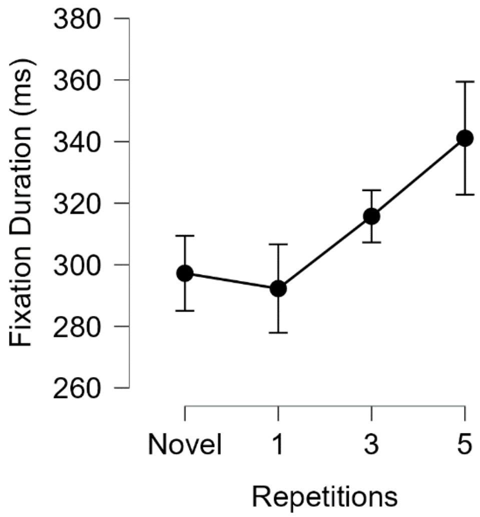
3.4 Fixation Duration

The fixation duration represents the mean amount of time (ms) spent on each fixation, throughout the 3500ms viewing of each stimulus. Given the constant exposure duration, this effect is somewhat predicted and negatively correlated with the measurement of fixation count. Accordingly, the mean fixation duration was the lowest for stimuli repeated once ($M = 292.26$), and the highest for stimuli repeated five times ($M = 341.116$). A repeated measures ANOVA revealed a statistically significant main effect, $F(3, 69) = 10.988$, $p < .001$, $\eta^2 = 0.323$. These results demonstrate that, as repetitions increase, the fixation duration also increases (see Figure 5).

Two contrasts were conducted to better understand the main effect. The first contrast assessed whether fixation duration for the novel condition differed from 1, 3, and 5 repetitions, combined. The second contrast evaluated whether fixation duration for 1 repetition significantly differed from 5 repetitions. Contrast results showed a significant difference for Comparison 1 (Estimate = 57.431, SE = 23.125, $t = 2.484$, $p = 0.015$) and Comparison 2 (Estimate = 48.856, SE = 9.441, $t = 5.175$, $p < .001$).

Figure 5

Fixation duration (ms) across repetition levels (novel, 1, 3, 5). The error bars on the plot represent 95% confidence intervals (CI).



3.5 Saccadic Amplitude

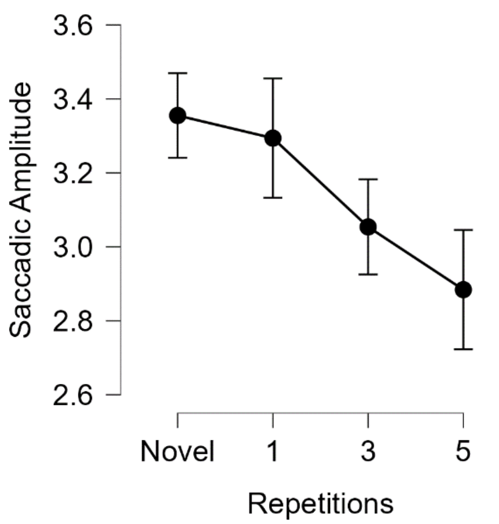
Saccadic amplitude is the measure of the range of eye movement expressed in degrees of visual angle. Larger values indicate a greater distance covered by the eyes during a saccade, implying increased exploration. Figure 6 visually depicts a decline in saccadic amplitude with increasing repetitions. A repeated measures ANOVA revealed a statistically significant main effect of repetitions, $F(3, 69) = 9.961$, $p < .001$, $\eta^2 = 0.302$. These findings show that participants made gradually shorter saccades as repetitions increased, suggesting reduced visual exploration.

Lastly, two contrasts were conducted to better understand the main effect. The first contrast assessed whether saccadic amplitude for the novel condition differed from 1, 3, and 5 repetitions, combined. The second contrast evaluated whether saccadic amplitude for 1 repetition significantly differed from 5 repetitions. Contrast results showed a significant difference for

Comparison 1 (Estimate = -0.833, SE = 0.239, t = -3.482, p < .001) and Comparison 2 (Estimate = -0.410, SE = 0.098, t = -4.194, p < .001).

Figure 6

Saccadic amplitude across repetition levels (novel, 1, 3, 5). The error bars on the plot represent 95% confidence intervals (CI).

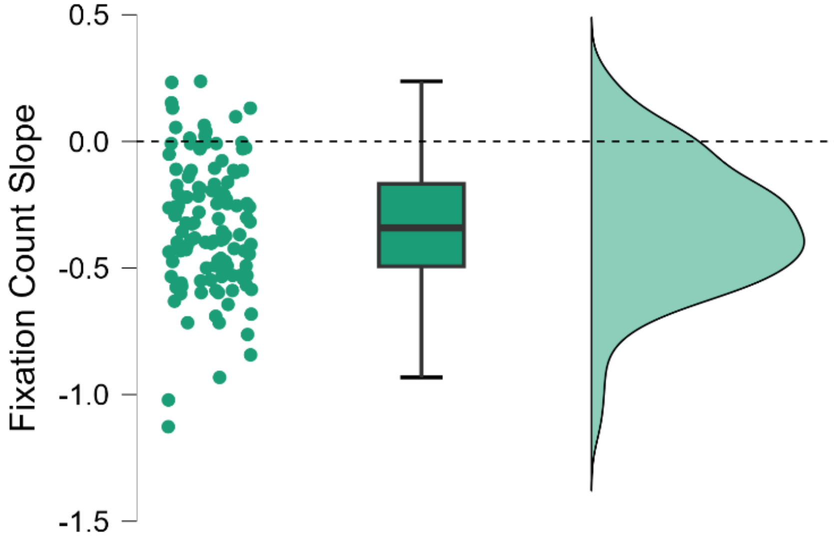


3.6 Item Analysis

An item analysis was conducted to ensure the consistency of the repetition effect across the entire image set. For each image, we calculated the slope of fixation counts across the four levels of repetition. The rationale behind measuring the slope of fixation counts was to assess how they changed across repeated exposures, to determine the presence of the repetition effect for each stimulus. Values below zero indicate that the fixation count decreased across four levels of repetitions. The mean slope of fixation counts was found to be -0.330, signifying an overall decrease in fixation counts over repetitions. A one-sample t-test showed a significant difference from zero (t(119) = -14.205, p < .001), indicating a highly consistent repetition effect across the stimuli. Out of 120 images, only 9 (7.5%) exhibited a slope of 0 or greater (see Figure 7), but these images were not noticeably distinct from the other images. This analysis confirms the consistency of the repetition effect across the AI-generated set, and validates that the repetition effect was not mediated by particular exemplars.

Figure 7

Plot of the fixation count slope across 120 stimuli. Slopes smaller than 0 indicate that the fixation count decreased as the number of repetitions increased.



Chapter 4: Discussion

The main objective of this study was to explore how repeated exposure to images of buildings influences episodic memory and eye movements, with a focus on understanding how memory and oculomotor systems interact in processing building stimuli. While faces and scenes have received considerable attention in eye tracking studies, buildings, a visual category that shares properties with both faces and scenes, have never been investigated in this context. The results revealed a repetition effect for buildings, characterized by increased memory and decreased fixations with repeated exposure, similar to findings on faces and scenes (Olsen et al., 2016; Smith & Squire, 2017; Mazloum-Farzaghi et al., 2023). These results suggest that the repetition effect supersedes differences in how specific object categories are processed. Moreover, they indicate that the memory and oculomotor systems are associated in processing buildings, akin to their association in processing faces and scenes.

4.1 Memory Performance

In the present study, we assessed recognition accuracy and confidence ratings. Recognition accuracy measures episodic memory for specific events, such as encountering an image, while confidence rating is a general measure of episodic memory strength that may or may not involve explicit recall (Yonelinas et al., 2010). It has been suggested that the two measures reflect two distinct underlying cognitive processes: recollection and familiarity (Yonelinas et al., 2010). The results indicated improved recognition accuracy with repeated presentation of building images in participants, increasing from one to three repetitions and plateauing due to a ceiling effect. Participants also showed incremental confidence ratings with increasing repetitions. Interestingly, recognition accuracy was lower for images repeated only once compared to novel images, likely due to a weaker recognition signal for images repeated once compared to those repeated three or five times, which led participants to misclassify them as novel.

The results show that repeated exposure to building images enhances participants' ability to recognize images, and boosts their confidence in recognition, although previous research has shown that recognition and confidence do not always correspond (Busey et al., 2000; Reinitz, 2012). For instance, Busey et al. (2000) found that increasing the luminance of an image may increase confidence, but reduce recognition accuracy, suggesting a dissociation between subjective confidence and objective accuracy. In the present study, we only manipulated repeated exposure as the independent variable. However, to control for extraneous variables such as image luminance, all stimuli were counterbalanced in six sequences to ensure equal exposure, and the experiment occurred in a controlled laboratory setting to minimize the impact of external influences. Thus, our findings serve as validation that repeated exposure to building stimuli effectively improves recognition and familiarity.

4.2 Eye Movement Metrics

Changes observed in eye movement metrics such as fixation count, fixation duration, and saccadic amplitude presumably uncover the cognitive processes underlying visual memory of buildings. In the current study, a significant repetition effect for building stimuli was found, characterized by a decreasing fixation count with increasing repetitions. The repetition effect showed a linear increase from three to five repetitions, compared to novel or once-repeated stimuli. Additionally, longer fixation durations, which refers to the mean length of time participants spent on each fixation, were strongly correlated with lower fixation counts. However, this correlation is predictable because fewer fixations leads to longer fixation duration, given the constant exposure duration of 3500ms for each stimulus. Finally, saccadic amplitude significantly decreased with more repetitions, implying reduced visual exploration. In sum, eye movement metrics indicate the presence of a repetition effect, previously found for faces and scenes, which is presumed to reflect implicit processing of previously encountered information by the brain (Olsen et al., 2016; Smith & Squire, 2017; Mazloum-Farzaghi et al., 2023).

4.3 Implicit vs. Explicit Memory

There is an ongoing debate as to whether the eye movement repetition effect reflects implicit or explicit memory processing. Prior studies examining the association between eye movement measures and memory measures have found support for both perspectives. Some results suggest the repetition effect may be implicit and hippocampus-independent, while others suggest it correlates with explicit memory and is hippocampus-dependent (Olsen et al., 2016; Smith & Squire, 2017; Mazloum-Farzaghi et al., 2023). The key difference between studies that showed associations with implicit memory, and those that did not, was the methodology used.

For instance, in a study by Olsen et al. (2016), healthy controls showed a repetition effect for fixed viewpoint faces and varied viewpoint faces, while, in contrast, a participant (H.C.) with compromised hippocampal function exhibited a repetition effect only for fixed viewpoint faces. The absence of a repetition effect for varied viewpoint faces indicates that the repetition effect exhibited by H.C. reflects implicit, hippocampal-independent, processing (Olsen et al., 2016). Consequently, if the task involves more complex visual memory processing, such as binding multiple face representations from variable viewpoints, then an intact hippocampal system is necessary (Ryan et al., 2000; Olsen et al., 2016). These results suggest that that the repetition effect may reflect both implicit and explicit processing, depending on the methodology used in the study.

Another study where methodology played a key role in determining whether the repetition effect was associated with implicit or explicit memory was conducted by Smith and Squire (2008; 2017). In an earlier study, Smith and Squire (2008) reported that the presence of a repetition effect was associated with explicit recognition, suggesting that the effect reflects explicit processing. However, in a later study conducted by the same researchers, Smith and Squire (2017) reported conflicting results, finding that the repetition effect was not associated with explicit recognition. The key difference was that the findings in the former study pointed towards an association with explicit memory, while the latter study indicated that the repetition effect may also reflect implicit memory, but only if memory testing is not expected (Smith & Squire 2017). In line with these findings, other studies found no association between explicit recognition memory and eye movement repetition effects when they used a distractor task, such as judging the gender of the presented faces, followed by a surprise memory test (Olsen et al., 2016; Mazloum-Farzaghi et al., 2023). In summary, awareness of upcoming memory testing has a significant impact on the repetition effect while viewing scenes and faces, leaving uncertainty as to whether the effect represents implicit or explicit memory.

Our experiment's findings point to both implicit and explicit memory involvement. In our experiment's study phase, participants judged whether buildings were residential or non-residential, intended to facilitate incidental encoding, which arguably reflects implicit memory (Olsen et al., 2016; Smith & Squire, 2017; Mazloum-Farzaghi et al., 2023). However, during the test phase, explicit memory measures showed repetition effects, while changes in eye movements suggested implicit memory engagement (Ryals et al., 2015; Ryan & Shen, 2020). Therefore, while our task likely involved implicit encoding, our behavioral results indicate the presence of both implicit and explicit memory in retrieval.

Additionally, despite implementing a distractor task involving judging building function, participants were informed via recruitment emails that the study aimed to explore how eye movements can uncover stored memories. This introduced uncertainty about whether encoding was genuinely incidental and whether participants had expectations of a memory test, or not. Consequently, although we observed a repetition effect, it remains unclear whether this effect reflects implicit or explicit memory, highlighting the challenge in interpreting this effect.

4.4 Stimulus Analysis

Another factor that influences eye movements is object category. Building images represent a distinct category of objects that are large, non-living, space-defining entities that can serve as landmarks with navigational value (Bastin et al., 2013; Cate et al., 2011; Mullally and Maguire, 2011; Troiani et al., 2014). Buildings can be considered a subset of scenes which are processed spatially and contextually (James et al., 2010; Epstein & Baker, 2019). However, buildings also represent unified entities similar to faces, but not entirely holistic, as evidenced by the absence of an inversion effect that is present for faces (Epstein et al., 2006). Repetition effects observed in our study suggest that buildings undergo more efficient processing by the brain with repeated exposure, much like scenes and faces (Ryals et al., 2015; Olsen et al., 2016; Smith & Squire, 2017; Mazloum-Farzaghi et al., 2023), indicating that the effect is not limited to specific categories of objects.

4.5 Limitations

Despite the new evidence provided by this study, some limitations should be acknowledged. In the present study, we utilized AI to generate novel building images. To the best of our knowledge, this is the first eye-tracking study to use AI-generated images of buildings as stimuli. Our item analysis revealed that the repetition effect was present for nearly all of the images, thus validating the quality of our stimulus set, and highlighting the promising potential of AI-generated stimuli in eye-tracking research. While this approach allows for manual selection and modification of images based on specific criteria, such as ensuring that the buildings are unified as a whole, fully visible, and free from obstruction by other objects, it also poses important limitations. In particular, there is uncertainty about the computations in the latent space used to generate AI images. The latent space serves as a mathematical coordinate system that links text and image features (Byrne, 2023). However, it is challenging for humans to understand and verify these relationships, which affects how reliable AI-generated images are for scientific experiments. In addition, AI-generated building stimuli may appear artificial and not generalizable to the real world, thus lacking ecological validity.

Another potential limitation relates to the fact that the incidental nature of the encoding phase could not be verified. Participants were informed via the recruitment email that the research aims to explore how eye movements can uncover stored memories. However, during the encoding phase of the experiment, participants were instructed to judge the function of buildings to encourage them to explore the images, which likely led to incidental encoding of the stimuli. Despite participants' awareness of the involvement of memory in the experiment, significant repetition effects were observed in our study.

One final limitation found in our study was the presence of a ceiling effect. Most participants achieved very high recognition accuracy, with six out of twenty-four participants achieving near-perfect accuracy. Similarly, confidence ratings were very high across all levels of repetition. Participants consistently performed well on both memory measures, even without any repetitions, suggesting that the task may have been too easy behaviorally. Since our participants achieved such high accuracy and confidence rates, we were unable to compare eye movements between participants with high accuracy and those with low accuracy, as there were virtually no participants with low accuracy scores. The ceiling effect also prevented us from performing an effective receiver operating characteristics (ROC) analysis between recognition accuracy and confidence ratings due to a narrow range of scores. Additionally, our study had a low number of participants (N=24) which allowed minimal power to uncover the repetition effect, but was inadequate for exploring individual differences. The presence of a ceiling effect, as well as a low number of participants, limited our ability to find meaningful correlations between memory and eye movement measures.

To counter the ceiling effect, future studies should reduce the number of repetitions to one, two, and three to avoid plateauing above three repetitions. Additionally, task difficulty may be increased by introducing a longer delay period between study and test phases, accompanied by a secondary task like engaging with auditory stimuli. This secondary task would increase the cognitive load by engaging auditory processing without directly competing with visual memory, thereby making it more challenging to retain details of previously presented images.

4.6 Conclusion and Future Directions

Our study revealed significant repetition effects in both memory and eye movements, akin to those previously observed with faces and scenes, suggesting that the repetition effect supersedes differences in how specific object categories are processed. Our findings highlight the interconnectedness of the oculomotor and memory systems in encoding and retrieving of visual building stimuli, which is crucial for spatial navigation and landmark recognition.

While these findings are significant, there is considerable potential for further exploration. Notably, our study did not investigate the resampling effect (Damiano & Walther, 2015; Ramey et al., 2020; Johansson et al., 2022, Nikolaev et al., 2023). By investigating the resampling effect, future studies may uncover stimulus-specific differences between buildings and other object categories. This exploration could shed light on whether buildings are processed more like faces or scenes, or in an entirely unique way, which may provide further understanding of how buildings are visually processed and remembered.

Future studies should also compare AI-generated buildings to real-world buildings. Additionally, different subcategories of buildings can also be compared by sorting them based on their semantic properties (a hospital vs. a house). Furthermore, using the approach in the study by Olsen et al. (2016), buildings viewed from fixed viewpoints can be compared with those viewed from varied viewpoints. AI tools like Midjourney (Midjourney, 2023) can be used to generate stimulus sets of buildings with both varied and fixed viewpoints. Lastly, to investigate the impact of memory test awareness, comparisons can be made between participants who viewed buildings with the expectation of a subsequent memory test, and those who were not expecting a memory test.

In our study, we sampled from a population of young adults. In the long term, future research on buildings should include older adults and individuals with selective neurological impairments. For instance, Mazloum-Farzaghi et al. (2023) found that older adults exhibit a weaker repetition effect for faces compared to young adults, likely due to age-related memory decline. Investigating how age affects the repetition effect for buildings could provide deeper insights into the processes responsible for this effect, and into how visual memory changes with age. Moreover, including individuals with compromised hippocampal or neocortical systems may help determine whether the repetition effect for buildings is mediated by specific brain systems, and whether it reflects implicit or explicit processing (Ryan et al., 2000; Olsen et al., 2016).

References

Bastin, J., Vidal, J. R., Bouvier, S., Perrone-Bertolotti, M., Bénis, D., Kahane, P., David, O., Lachaux, J. P., & Epstein, R. A. (2013). Temporal components in the parahippocampal place area revealed by human intracerebral recordings. *Journal of Neuroscience, 33*(24), 10123–10131. https://doi.org/10.1523/JNEUROSCI.4646-12.2013

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PloS One, 3*(8), e3022. https://doi.org/10.1371/journal.pone.0003022

Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. F. (2000). Accounts of the confidence-accuracy relation in recognition memory. *Psychonomic bulletin & review, 7,* 26–48. https://doi.org/10.3758/BF03210724

Byrne, U. (2023). A Parochial Comment on Midjourney. *International Journal of Architectural Computing, 21(*2), 374–379. https://doi.org/10.1177/14780771231170271

Caldara, R., Zhou, X., & Miellet, S. (2010). Putting culture under the 'spotlight' reveals universal information use for face recognition. *PLoS One, 5*(3), e9708. https://doi.org/10.1371/journal.pone.0009708

Cate, A. D., Goodale, M. A., & Köhler, S. (2011). The role of apparent size in building- and object-specific regions of ventral visual cortex. *Brain research, 1388*, 109–122. https://doi.org/10.1016/j.brainres.2011.02.022

Chuk, T., Chan, A. B., & Hsiao, J. H. (2017). Is having similar eye movement patterns during face learning and recognition beneficial for recognition performance? Evidence from hidden Markov modeling. *Vision research, 141*, 204–216. https://doi.org/10.1016/j.visres.2017.03.010

Damiano, C., & Walther, D. B. (2015). Content, not context, facilitates memory for real-world scenes. *Visual Cognition, 23*(7), 852–855. https://doi.org/10.1080/13506285.2015.1093241

Damiano, C., & Walther, D. B. (2019). Distinct roles of eye movements during memory encoding and retrieval. *Cognition, 184*, 119–129. https://doi.org/10.1016/j.cognition.2018.12.014

Epstein, R. A., Higgins, J. S., Parker, W., Aguirre, G. K., & Cooperman, S. (2006). Cortical correlates of face and scene inversion: a comparison. *Neuropsychologia*, *44*(7), 1145–1158. https://doi.org/10.1016/j.neuropsychologia.2005.10.009

Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation*. Trends in cognitive sciences, 12*(10), 388–396. https://doi.org/10.1016/j.tics.2008.07.004

Epstein, R. A., & Baker, C. I. (2019). Scene perception in the human brain. *Annual review of vision science*, *5*, 373–397. https://doi.org/10.1146/annurev-vision-091718-014809

Hannula, D. E., & Ranganath, C. (2009). The eyes have it: hippocampal activity predicts expression of memory in eye movements. *Neuron, 63*(5), 592–599. https://doi.org/10.1016/j.neuron.2009.08.025

Haxby, J. V., Ishai, A., Chao, L. L., Ungerleider, L. G., & Martin, A. (2000). Object-form topology in the ventral temporal lobe: Response to I. Gauthier (2000). *Trends in cognitive sciences, 4*(1), 3–4. https://doi.org/10.1016/S1364-6613(99)01423-0

JASP Team (2024). *JASP* (Version 0.18.3) [Computer software]. https://jasp-stats.org/

James, T. W., Huh, E., & Kim, S. (2010). Temporal and spatial integration of face, object, and scene features in occipito-temporal cortex. *Brain and Cognition, 74*(2), 112-122. https://doi.org/10.1016/j.bandc.2010.07.007

Johansson, R., Nyström, M., Dewhurst, R., & Johansson, M. (2022). Eye-movement replay supports episodic remembering. *Proceedings of the Royal Society B*, *289*(1977), 20220964. https://doi.org/10.1098/rspb.2022.0964

Kafkas, A., & Montaldi, D. (2011). Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection from familiarity. *Quarterly journal of experimental psychology*, *64*(10), 1971–1989. https://doi.org/10.1080/17470218.2011.588335

Mazloum-Farzaghi, N., Shing, N., Mendoza, L., Barense, M. D., Ryan, J. D., & Olsen, R. K. (2023). The impact of aging and repetition on eye movements and recognition memory. *Aging, Neuropsychology, and Cognition*, *30*(3), 402–428. https://doi.org/10.1080/13825585.2022.2039587

MidJourney. (2023). *Midjourney*. https://www.midjourney.com

Mikhailova, A., Raposo, A., Sala, S. D., & Coco, M. I. (2021). Eye-movements reveal semantic interference effects during the encoding of naturalistic scenes in long-term memory. *Psychonomic Bulletin & Review*, *28*(5), 1601–1614. https://doi.org/10.3758/s13423-021-01920-1

Mullally, S. L., & Maguire, E. A. (2011). A new role for the parahippocampal cortex in representing space. *Journal of Neuroscience*, *31*(20), 7441–7449. https://doi.org/10.1523/JNEUROSCI.0267-11.2011

Nikolaev, A. R., Bramao, I., Johansson, R., & Johansson, M. (2023). Episodic memory formation in unrestricted viewing. *NeuroImage*, *266*, 119821. https://doi.org/10.1016/j.neuroimage.2022.119821

Noton, D., & Stark, L. (1971). Scanpaths in eye movements during pattern perception. *Science, 171*(3968), 308–311. https://doi.org/10.1126/science.171.3968.308

Olsen, R. K., Sebanayagam, V., Lee, Y., Moscovitch, M., Grady, C. L., Rosenbaum, R. S., & Ryan, J. D. (2016). The relationship between eye movements and subsequent recognition: Evidence from individual differences and amnesia. *Cortex, 85*, 182–193. https://doi.org/10.1016/j.cortex.2016.10.007

OpenAI. (2023). *ChatGPT*. https://chat.openai.com

Ramey, M. M., Henderson, J. M., & Yonelinas, A. P. (2020). The spatial distribution of attention predicts familiarity strength during encoding and retrieval. *Journal of Experimental Psychology: General, 149*(11), 2046. https://doi.org/10.1037%2Fxge0000758

Reinitz, M. T., Séguin, J. A., Peria, W., & Loftus, G. R. (2012). Confidence–accuracy relations for faces and scenes: Roles of features and familiarity. *Psychonomic Bulletin & Review, 19*, 1085–1093. https://doi.org/10.3758/s13423-012-0308-9

Research, S. R. (2010). *SR Research Ltd.* https://www.sr-research.com/

Robin, J. (2018). Spatial scaffold effects in event memory and imagination. *Wiley Interdisciplinary Reviews: Cognitive Science, 9*(4), e1462. https://doi.org/10.1002/wcs.1462

Ryals, A. J., Wang, J. X., Polnaszek, K. L., & Voss, J. L. (2015). Hippocampal contribution to implicit configuration memory expressed via eye movements during scene exploration. *Hippocampus*, *25*(9), 1028–1041. https://doi.org/10.1002/hipo.22425

Ryan, J. D., Althoff, R. R., Whitlow, S., & Cohen, N. J. (2000). Amnesia is a deficit in relational memory. *Psychological science, 11*(6), 454–461. https://doi.org/10.1111/1467-9280.00288

Ryan, J. D., & Shen, K. (2020). The eyes are a window into memory. *Current Opinion in Behavioral Sciences*, *32,* 1–6. https://doi.org/10.1016/j.cobeha.2019.12.014

Shen, K., Bezgin, G., Selvam, R., McIntosh, A. R., & Ryan, J. D. (2016). An anatomical interface between memory and oculomotor systems. *Journal of Cognitive Neuroscience, 28*(11), 1772–1783. https://doi.org/10.1162/jocn_a_01007

Smith, C. N., & Squire, L. R. (2008). Experience-dependent eye movements reflect hippocampus-dependent (aware) memory. *Journal of Neuroscience, 28*(48), 12825–12833. https://doi.org/10.1523/JNEUROSCI.4542-08.2008

Smith, C. N., & Squire, L. R. (2017). When eye movements express memory for old and new scenes in the absence of awareness and independent of hippocampus. *Learning & Memory, 24*(2), 95–103. https://doi.org/10.1101/lm.043851.116

Tatler, B. W., Kirtley, C., Macdonald, R. G., Mitchell, K. M. A., & Savage, S. W. (2014). The Active Eye: Perspectives on Eye Movement Research. *Current Trends in Eye Tracking Research* (pp. 3–16). Springer International Publishing. https://doi.org/10.1007/978-3-319-02868-2_1

Troiani, V., Stigliani, A., Smith, M. E., & Epstein, R. A. (2014). Multiple object properties drive scene-selective regions. *Cerebral cortex*, *24*(4), 883–897. https://doi.org/10.1093/cercor/bhs364

Yonelinas, A. P., Aly, M., Wang, W. C., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, *20*(11), 1178–1194. https://doi.org/10.1002/hipo.20864