The function of eye-movements in visual recognition memory

Faya Movchan

A thesis submitted for the degree of

Master of Philosophy

in

Cognitive Neuroscience

Supervised by:

Professor Neil Burgess, Institute of Cognitive Neuroscience

Institute of Cognitive Neuroscience

University College London

November 2, 2024

Declaration of Authorship

I, Faya Movchan, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Understanding the mechanisms of visual recognition memory is fundamental to exploring how we perceive and remember complex visual scenes. Traditionally, eye movements are viewed as responses to sensory input that enable detailed scene and object representation. However, recent studies suggest that some neural mechanisms may guide eye movements based on learned spatial relationships in visual recognition memory. This research aimed to investigate the role of eye movement patterns as a potential mechanism supporting recognition memory. To do this I focussed on sequential encoding-recollection similarity (SERS) across shape, direction, length, and position dimensions. Guided by a computational model proposing that grid-like coding in entorhinal cortex underlies spatial memory and eye movement, I developed a novel foveated vision paradigm, designed to produce distinct, measurable sequences of eye movements when looking for the missing object in an encoded array of five objects, without distraction from peripheral objects.

I hypothesised that participants would produce similar scan-paths between encoding and retrieval phases of the memory task, with these similarities correlating positively with memory performance. Using permutation analyses, mixed effects models and correlational analyses, we found significant within-trial SERS. This was particularly strong in the position dimension, which was most strongly associated with accurate recall, suggesting that precise re-fixations on the location of specific object features can enhance recognition and reduce error in locating missing objects. Although other dimensions did not consistently show this effect on memory performance, great variability in shape similarity across participants suggests two distinct individual encoding strategies, emphasising the need for further investigation. Despite the limitations due to the small sample size, our findings highlight the role of eye movement reinstatements in memory recall, and show that the new paradigm offers a robust framework for future exploration into the neural basis of these eye movement patterns.

Impact Statement

Visual recognition memory in its essence is the ability by the organism to determine whether a visual stimulus is novel or familiar. This type of memory is essential for humans and one can hardly imagine their life without recognizing their house or family. In this study, I aimed to look at visual recognition memory through a novel experimental paradigm, especially focusing on eye movements contributions and their connection to memory performance. However, our memory can decline in conditions such as Alzheimer's disease, which involves entorhinal cortical pathology in its early stages which correlates with path integration ability – a putative role of grid cells (Howett et al., 2019). The understanding of the computational processes underlying visual recognition memory might lead us to convenient ways to monitor the progression of, for example, Alzheimer's disease in those whose recognition memory for spatial-relational information is impaired. While this research cannot yet comment on medial temporal structures' involvement in recognition memory and the role of grid cells in cognitive function, it contributes valuable insight by linking scanpath similarities to memory performance, leading to the future research that will. Results from this study showed that precise re-fixations on specific object features in the position similarity dimension were strongly associated with accurate recall, suggesting that detailed focus on certain object features may enhance recognition memory. Additionally, the observed variability in shape similarity across participants pointed to possible individual encoding strategies, highlighting the need for further investigation into how these differences impact memory performance. The experimental paradigm developed in this study, with special focus on isolating the role of eye movements, provides a framework for future exploration, including studies utilising neuroimaging techniques like fMRI to investigate neural correlates such as grid-like representations in the brain. Beyond its theoretical contributions, the experimental paradigm developed in this study holds potential as a tool for assessing visual recognition memory in both healthy individuals and populations at risk for cognitive decline. By refining this task further, it has the potential to be applied to clinical research to help identify early markers of Alzheimer's disease, offering a non-invasive approach to monitor memory function over time. This could ultimately lead to improved methods for diagnosing and tracking memory impairments, supporting early intervention strategies in neurodegenerative diseases.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to Professor Neil Burgess for his full support and supervision throughout this project. Without his constant guidance, patience and the ability to show me the right direction, I could not have accomplished this work. I am incredibly grateful for him welcoming me into the Space & Memory Lab and for his understanding and empathy during challenging times, particularly in relation to the ongoing situation in Ukraine. His support extended beyond academic mentorship, and it was a great privilege to be working and learning from someone with such a great expertise and dedication to science.

I am also deeply thankful to Professor Andrej Bicanski, my secondary supervisor. Though not physically present in the lab, he was consistently involved and supportive of my project, offering valuable insights and ideas. After all, the foundation of this thesis was inspired by his and Neil Burgess's work, and his personal enthusiasm for the project played a significant role in shaping my research. I am especially grateful for his hospitality when he welcomed me for a week at Newcastle University to collaborate on the project and for his understanding and solidarity during difficult times.

My heartfelt thanks also go to the members of the Space & Memory Lab, including Yan from the neighbouring Memory and Space Lab, for their warm welcome, ongoing support, and inspiring discussions during our lunchtime conversations. A special thank you goes to Dr. Andrea Castegnaro and Dr. Luke Emrich-Mills for their invaluable contributions to the design and structure of my experiment. Luke's willingness to let me train on his data and shadow his experiment was crucial in helping me understand the practical aspects of experimental work. Andrea's expertise and assistance with coding, specifically with the alpha mask, one of the most challenging components of my procedure, were indispensable.

I would like to acknowledge my department: the UCL Division of Psychology and Language Sciences and the UCL Institute of Cognitive Neuroscience for providing a supportive environment for my research. I am particularly grateful to Professor Benedetto De Martino for granting me access to his Eyelink equipment and for introducing me to his PhD students, who taught me how to operate it.

I owe a debt of gratitude to my family for their constant support and motivation throughout this journey. To my sister, Polina, who is also pursuing her PhD at UCL, for sharing this academic adventure with me. And to my best friend, Myron — for being there every single step of the way. From our phone conversations on the way to the lab to being a test subject for every iteration of my experiment, to sitting with me through the writing process, his presence kept me motivated and focused. I could not have done this without them.

Contents

1. Introduction

Visual recognition memory relies not only on the content of what we see but also on how we explore it through eye movements. This thesis examines the contribution of those eye movements to visual recognition memory, guided by the predictions of the computational model developed by Bicanski and Burgess (Bicanski & Burgess, 2019). This model suggests that grid cells in the medial temporal lobe (MTL) are not only crucial for spatial navigation but may also play a role in visual memory by encoding spatial relationships between elements of a scene and guiding eye movements back to them during recognition. The primary aim of this research is to develop a new experimental paradigm to test to what extent eye movements are guided back to scene elements during recognition and to explore how this relates to memory performance. A secondary aim is to make this paradigm useful in future studies, testing specific predictions about the role of grid cell firing in this process.

1.1 GRID CELLS

The medial temporal lobe (MTL) has long been recognized as crucial for spatial memory and navigation (Scoville & Milner, 1957; O'Keefe & Nadel, 1978; Bird & Burgess, 2008; Burgess et al., 2002). Research into the neural mechanisms underlying these cognitive functions has identified specific types of neurons, including place cells and grid cells, that exhibit spatially modulated firing patterns. These specialised neurons are critical for creating and maintaining an internal representation of the environment, often referred to as a "cognitive map" (Hartley et al., 2014). Place cells, which are primarily found in the hippocampus, become active when an organism is in a specific location, contributing to the encoding of spatial information (O'Keefe & Nadel, 1978). Grid cells are neurons that fire at regular, periodic locations as the organism navigates through the environment, which form a grid-like pattern of equilateral triangles. Discovered first in the medial entorhinal cortex of rodents in 2005 (Hafting et al., 2005), they provide the brain with a spatial coordinate system. Initially, grid cells were believed to function exclusively in spatial navigation, helping animals (including humans) to orient themselves and create a mental map of their surroundings. Over time, this spatial map allows an organism to recognize familiar environments and navigate more effectively.

Recent studies using animal electrophysiology and human functional neuroimaging (e.g., Killian et al., 2012; Constantinescu et al., 2016) suggests that grid cells can support a wide range of tasks beyond spatial navigation. Depending on the origin of their inputs these grid cells could support a variety of cognitive functions including spatial navigation, memory and planning. While the majority of research on grid cells has been done using electrophysiology in rodents, some recent human functional magnetic resonance imaging (fMRI) studies have allowed researchers to study neural mechanism for cognitive information processing by grid cells, non-invasively in human brain in both spatial and conceptual 2D tasks and in a 3D environment. Furthermore, in a groundbreaking study by Doeller, Barry & Burgess in 2010, grid-like fMRI activity was shown in humans while performing navigation tasks in a virtual environment, providing the first non-invasive evidence for grid cell-like activity in the human

memory network. This discovery has led to a new wave of research exploring whether grid cells could be involved in more general forms of cognition, including visual recognition memory.

1.2 *MEMORY*

Memory is a complex cognitive process that relies on several regions within the brain, including the hippocampus, parahippocampal cortex, and other areas within the MTL. It is important to define for this study that relational memory (Olsen et al., 2015) refers to the ability to remember the relationships between various elements, such as objects, places, and people, within a particular context. This type of memory is critical for recognizing how different components are connected to form the whole picture. And it is believed that the hippocampus plays a crucial role in relational memory by binding these elements together into a unified representation (Hannula et al. 2007). When relational memory functions effectively, it enables individuals to recall how items are linked, such as, important for our experiment, remembering the location of a misplaced object. For instance, Olsen et al. (2015) demonstrated that individuals with developmental amnesia struggle with tasks requiring relational binding, specifically in item memory tasks like face recognition. The hippocampal damage in these cases disrupts the ability to bind features of a memory together. Similarly, Pertzov et al. (2013) identified binding deficits in patients with MTL damage due to limbic encephalitis, highlighting the role of the MTL in linking objects and their spatial relationships within memory. The interaction between spatially-modulated cells like grid and place cells and these broader memory processes highlights the MTL's integral role not only in spatial cognition but in memory as a whole. Understanding how these specialised cells contribute to both spatial and non-spatial memory systems is a key area of ongoing research.

1.3 EYE MOVEMENTS

Eye movements are considered to be key players in visual memory processes, and there is a lot of evidence supporting this. For example, eye movements are linked to relational memory, as the brain often guides saccades based on relational information. For example, research by Johansson and Nyström (2022), demonstrates that eye-movement replay plays a crucial role in episodic remembering. In their study, participants were first shown images to encode and later asked to recall these images while staring at a blank screen. Remarkably, the eye movements participants made during recall mirrored the scanpaths they had followed during encoding, even in the absence of any visual stimuli, suggesting that participants were replaying the same visuospatial sequences to reconstruct the previously stored information. This replay mechanism offers direct evidence that eye movements play an active role in retrieving spatio-temporal relationships from memory, but it is not yet clear how. Further extending this idea, Johansson and Nyström analysed not only the shapes of the paths created by eye movements but also positions, lengths, directions and durations, finding overlaps in both the spatial locations and the shapes. Their results support the hypothesis that the sequential order of eye fixations, or scanpaths, plays an active role in memory retrieval, similar to what we want to look at in this study. However, while they provided valuable

insights into how eye movements contribute to memory, it did not directly address how grid cells might be involved in encoding and recalling these spatial relationships.

Related studies, such as one from the University of Toronto (Wynn, Olsen, Binns, Buchsbaum, & Ryan, 2018), have shown that fixation reinstatements can help mitigate memory decline, particularly in older adults. This research highlights how the reinstatement of specific gaze patterns can compensate for cognitive deficits, further emphasising the link between eye movements and memory performance. Eye movements have also been shown to play a critical role in scene construction and memory retrieval processes. The other studies explore the role of eye-movements in memory even further, on a neural level, according to Ladyka-Wojcik et al. (2022), free viewing during memory tasks, where participants can move their eyes without restriction, improves the effective connectivity between the hippocampus (HPC) and the frontal eye fields (FEF). This enhanced connectivity suggests that the coordination between the hippocampus, which is responsible for memory encoding and retrieval, and the FEF, which controls eye movements, becomes more efficient as participants engage in free-viewing tasks. As a result, individuals are better able to reconstruct scenes from memory, further emphasising the interplay between eye movements and hippocampal activity in these tasks.

In the context of visual imagery and memory recall, eye movements have also been linked to the richness of the details recalled. Studies by Armson et al. (2020) and Barker & Armson (2024) explored memory for a gallery tour and found that individuals who are better at generating detailed visual imagery tend to make more eye movements during memory tasks. These high-images produced more detailed recollections, and their eye movements often preceded the production of those details. This finding suggests that the process of recalling details is closely tied to eye movements, which may help individuals retrieve and reconstruct visual and spatial information stored in memory. In other words, the more eye movements a person makes, the better they reinstate spatiotemporal context. These studies provide a foundation for exploring the potential role of grid cells in memory tasks, as they illustrate how patterns of eye movements can influence cognitive performance, not just on behavioural, but on neural level.

1.4 BICANSKI AND BURGESS' MODEL

The suggestions that eye movements can drive grid cell-like firing and recent publications on the role of grid cells in vector navigation have helped Bicanski & Burgess (2019) develop a computational model of visual recognition memory. As mentioned above, this model suggests that grid cells have an important role in recognition of the spatial relationships in visual stimuli, drawing a connection between grid cells' firing patterns and visual recognition memory. In "A computational model of visual recognition memory via grid cells", the authors propose a computational framework that links the role of grid cells, located in the entorhinal cortex, to visual recognition memory. The authors show how grid cells can theoretically facilitate memory processes through their spatial firing patterns. The model describes how recognition memory for objects is represented through grid cells encoding translation vectors

between salient stimulus features in eye-movements (saccades). These vectors help map the spatial relationships between features, providing a coordinate system that assists in recalling the layout of objects or a relative location of the objects in respect to one another. The proposed model of recognition involves sequential hypothesis testing for the visual feature to be found at the end of each vector from the previous feature. This process integrates both spatial and visual information to aid in memory-driven recognition.

Moreover, as visual perception is closely linked with eye movements, which occur in a necessary sequential order, the series of saccades can be interpreted as a complex path on a two-dimensional plane. This allows for parallels between spatial navigation in two-dimensional space and the movement of the eyes across the visual field. While in the previous research grid-like representations have already been linked to eye-movements, they hadn't been linked to behaviour until Bicanski and Burgess' model. In other words, the research emphasises the potential function of grid cells in encoding and recalling visual memories.

Both previous research on grid-like activity in visual paradigms as well as the model's ability to produce fine within-category judgments and also accommodate for a broader Bayesian interpretation of perception, through cell-tuning, suggest the external validity of the model and gives evidence for the grid cells' activity in visual processing. While the computational model of visual recognition memory described above creates a working quantitative theory for the role of grid cells in visual recognition, it requires further testing. The hypothesis would suggest that visual recognition memory engages grid cells to trace these vectors, aiding in recognizing familiar layouts. This theory shifts the focus from spatial memory (usually linked to the temporal lobe) to visual memory, proposing the link between visual perception and eye movements, suggesting that the brain engages grid cells to direct saccades in a top-down manner during visual memory tasks. In this view, rather than passively exploring the environment, the brain actively controls eye movements based on previously encoded spatial relationships. When recognizing familiar layouts, grid cells guide eye movements from one feature to the next, creating a series of vectors that ultimately lead to the recognition of the visual scene. If successful, this research could offer new insights into how grid cells contribute to both spatial and non-spatial cognitive tasks, further supporting the idea that these neurons function as part of a universal representational system in the brain.

1.5 NEW EXPERIMENTAL PARADIGM

To test the Bicanski & Burgess (2019) model of visual recognition memory, it was first important to establish a reliable experimental paradigm. A logical starting point was to adapt a similar technique used in the experiment by Johansson and Nyström's (2022), mentioned above, which also focuses on examining sequential reinstatement of eye movements during memory tasks, allowing us to refine the paradigm for testing the specific predictions of grid-like representations in visual memory. However, human memory is an exceptionally complex system and shaped by numerous factors, including the specific types of memory being used for different tasks, as well as the strategies that individuals unconsciously or

consciously employ to encode and recall information. Each of the memory types relies on different brain structures and mechanisms, often operating simultaneously, making it challenging to distinguish which areas of the brain are primarily responsible for the specific memory tasks in a given context (Squire, L.R., & Zola, S.M. 1996; Henderson et al. 2005; Eichenbaum et al. 2007). This variability makes it extremely hard to pinpoint which strategies are being employed by participants in Johansson and Nyström's (2022) study, especially since people may not always be consciously aware of the methods they are using.

Eye movements, which are crucial for this experiment, can vary significantly depending on whether someone is looking at one object at a time or perceiving the overall shape of all objects together to remember the spatial relationships. Given these complexities, constructing an experiment that controls for as many variables as possible is crucial for gaining clearer insights into memory mechanisms. One of the primary challenges in memory research is to limit factors that could influence performance, like bottom-up signals, where attention is driven by the salient features of the visual stimuli such as contrast or luminance. Or such strategies, like peripheral vision — the ability to perceive objects outside of the direct line of sight — that can unintentionally aid memory encoding by allowing participants to process multiple stimuli simultaneously, thus bypassing the primary focus of the experiment. To create more controlled conditions, experimental designs need to exclude such influences, ensuring that participants rely solely on the cognitive processes under investigation. This is why it was so important to create a completely new experimental paradigm, which would force participants to focus only on what is in their direct line of sight.

By restricting peripheral vision, I could ensure that participants are consciously and actively engaging with the task, without relying on extraneous visual cues, helping us to confirm whether eye movements directly contribute to encoding and recalling spatial relationships between the scene elements themselves. Without such controls, it becomes nearly impossible to distinguish between the different strategies participants might be employing, leading to ambiguous results. In my experiment, we implement a novel paradigm, calling it: "The Object-Location Memory Experiment", which consists of two core stages: encoding and retrieval.

During the encoding phase, participants will observe a randomly generated arrangement of objects and form a visuospatial representation of their locations. In the retrieval phase, they will compare the arrangement to the stored visuospatial image and identify the missing object, focusing on the location where that object was previously seen. And in order to enforce controls, mentioned above, this experiment introduces a unique challenge: foveated vision. During both stages, only objects directly fixated on will be visible, thus preventing participants from relying on peripheral vision. This constraint forces them to make deliberate eye movements, relying on top-down control to navigate their gaze based on the spatial configuration of objects that is actively constructed in their memory.

1.6 RESEARCH AIMS AND HYPOTHESIS

The aim of this research project is to create a paradigm that can reliably produce measurable and distinct sequences of eye movements, investigating how memory performance correlates with eye movement behaviour, excluding other memory techniques that can interfere with the hypothesis testing. By establishing this foundation, we are building a framework that will be essential for future studies, where this paradigm can help investigate the presence and involvement of grid-like representations in visual recognition memory as a function of eye movements. And, as mentioned above, this study aims to build upon the findings from Johansson and Nyström's (2022) study, replicating their results, using the same MaltiMatch analysis algorithm (the description can be found in the Methods section). However, we plan to take it a step further by partially obscuring the visual scene to specifically investigate eye movements driven by mnemonic processing. Therefore, the rationale behind this particular experiment is that if participants exhibit more encoding-recollection similarity (SERS) in their scanpaths than would be expected by chance, and if these scores correlate with memory performance, it would suggest that memory representations are driving eye movements. In essence, this would provide evidence that the brain's memory systems influence the visual exploration patterns seen during retrieval in a top-down manner.

HYPOTHESIS ONE:

We hypothesise that participants performing a memory task will produce similar scanpaths between encoding and retrieval, reflecting the presence of a sequence of saccadic eye-movement vectors moving from one object to the expected location of the next.

HYPOTHESIS TWO:

We hypothesise that measures that capture eye-movement similarities — such as position, shape, length, and direction — will be positively correlated with memory performance. The higher the similarity between encoding and retrieval scanpaths, the better the memory performance is expected to be, supporting the theory that memory-driven eye movements are an essential component of visual recognition and relational memory.

2. Methods

Here, I will be outlining the process involved in designing the study paradigm. It includes analysing Luke Emrich-Mills' preliminary data, a description of the novel task implementation using EyeLink and details on how eye-tracking data was processed and analysed.

2.1 *Luke Emrich-Mills' data test*

At the beginning of my project, I had an opportunity to analyse eye-tracking data from Dr. Emrich-Mills' spatial memory task. The experiment was focused on designing a virtual reality task to study spatial memory in healthy ageing and the early stages of Alzheimer's disease. (Emrich-Mills, 2024). It involved a memory task where participants viewed an array of objects from a specific viewpoint and then, after a change or no change in viewpoint, were required to identify which object had changed position during retrieval. This task tested their ability to detect object movement and recall locations efficiently. The trials included variations in whether participants stayed in the same location or moved, and whether the table remained stationary or rotated.

To practice my analysis techniques and prepare for the creation of my own experimental paradigm, I used a subset of Dr. Emrich-Mills' trials. Specifically, I focused on trials where neither the table nor the viewpoint changed, allowing me to directly compare participants' eye movements between encoding and retrieval phases. I then analysed whether participants produced identical saccades, such as moving their gaze from one object (e.g., an apple) to another (e.g., a tomato), across phases. This enabled me to quantify the number of repeated vector pairs and test whether these repetitions occurred more frequently than chance would predict. To statistically validate this, I performed a sign test to test the hypothesis that participants produce more identical vector pairs between encoding and retrieval than would be expected randomly. And to find the probability that a specific number of similar vector pairs are repeated between the encoding and retrieval the following formula was used, derived on the basis of the binomial coefficient formula:

$$C_n^N = \frac{N!}{n!(N-n)!} \qquad P(R) = \frac{n_e C_R \cdot (n_p - n_e) C_{n_r - R}}{n_p C_{n_r}}$$

;

Where:

- n_e is the number of pairs fixated during encoding.
- n_r is the number of pairs fixated during retrieval.
- n_p is the total number of possible pairs (calculated as $N(N-1)$ for N objects).
- R is the specific number of pairs that are repeated between encoding and retrieval.

The significant p-value for overall participants in the sign test (p-value = 0.031) indicated that eye movement patterns (vector pairs) observed during retrieval are not random but rather reflect the reinstatement of the encoding-phase scanpaths. Additionally, I ran the same test separating the trials into correct and incorrect, finding significance as well. The significantly lower p-value for correct trials (p = 0.008), compared to that of incorrect trials (p = 0.0228), highlights the connection between scanpath reinstatement and memory performance. While I achieved promising results, such as identifying more repeated vector pairs than expected by chance across participants, the small sample size (n=5) and limited trial conditions made it difficult to draw any definitive conclusions about my hypothesis. However, this exercise proved to be invaluable for two key reasons: it provided me with hands-on experience in analysing eye movement data, and it highlighted areas where my own task design could be improved to better address my research questions. Furthermore, despite all limitations, the data still suggested that eye movements during encoding and retrieval are closely linked, supporting the broader literature on the role of saccades in memory tasks.

2.2 *Materials and Set-up*

My experiment used the EyeLink 1000 Plus eye-tracking system, chosen for its low noise levels and high spatial resolution, which minimises the need for data filtering. The manufacturer reported that the system operates at a sampling rate of 1000 Hz, offering high temporal resolution with 1-millisecond precision. During each saccade or fixation, data is collected on gaze velocity, position, and pupil size. This data is then used to calculate starting, ending, and average positions, pupil size, velocity, and peak velocity. Velocity was also converted into degrees per second using real-time resolution information, which is essential for creating saccade and fixation events that are recorded in an EDF file.

A velocity threshold of 22 degrees per second is typically employed to detect small saccades (as little as 0.3°), which is recommended for “smooth pursuit” experiments. However, for this experiment, a conservative threshold of 30°/sec was used, which shortens saccades and lengthens fixation durations—parameters suitable for cognitive research. The following parameters were employed for data recording, as recommended in prior research for memory and related tasks (SR Research, 2018, Parker & Tao, 2023):

- Recording parse type = GAZE
- Saccade velocity threshold = 30
- Saccade acceleration threshold = 8000
- Saccade motion threshold = 0.1
- Saccade pursuit fixup = 60
- Fixation update interval = 0

Both the participant and the eye tracker are set up via the Camera Setup screen, shown in Figure 1.

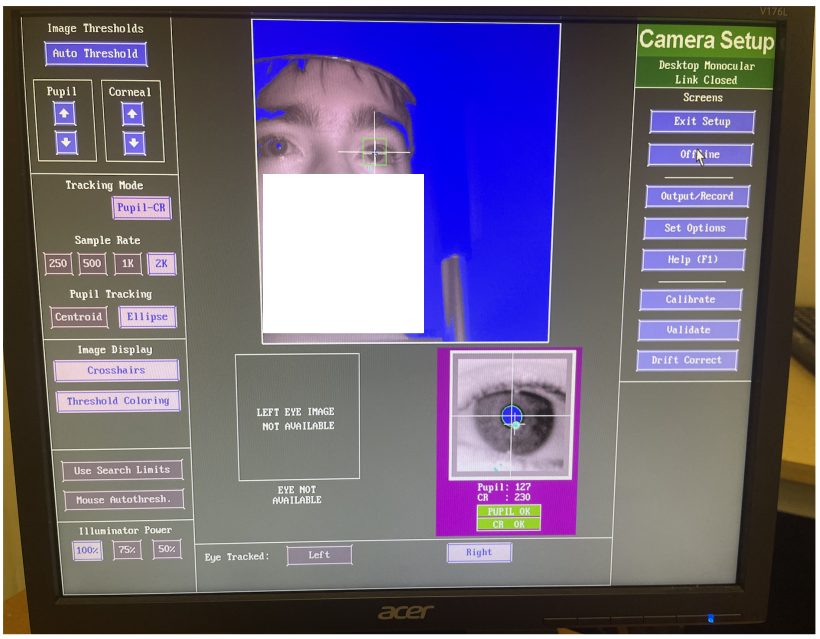

Figure 1: Example Camera Setup Screen.

The participants were positioned in a setup as shown in Figure 2. The EyeLink 1000 illuminator was positioned 65 cm in front of the participant to provide the IR illumination necessary for accurate eye-tracking. The participant's gaze was recorded through the EyeLink 1000 camera, and the chin-and-forehead rest, positioned 87 cm from the computer screen, ensured minimal head movement, thus allowing for accurate and stable tracking throughout the experiment.

Figure 2: Example Experiment Setup

Gaze position data reports the participant's exact (x, y) coordinates on the display, adjustmenting for the distance between the participant and the computer screen. The data were recorded in pixels, representing the participant's actual gaze location on the 1920 x 1080 pixels display. Pupil size data (measured as a percent change relative to a baseline) was also collected but was not used in this particular experiment. For this study, gaze data were recorded from the dominant eye, which happened to be the right eye for all of the participants. The dominant eye was identified using a classic test where participants focused on a distant object through a triangular opening made with their hands and closed one eye at a time to determine which eye kept the object centred (Heiting, 2020). The calibration procedure followed the 9-dot routine as outlined in the EyeLink® 1000 Plus User Manual (SR Research, 2018). Calibration was deemed successful if no calibration errors exceeded 1°, or if errors were consistent in size and direction and could be corrected by drift correction.

The raw data output consisted of x, y gaze positions along with corresponding frames. To facilitate further analysis, we also recorded different timestamps during the experiment to match the frames to the corresponding time and convert it to seconds, so that each frame is matched with the relevant experimental block, trial number, and trial type. The final data table included the following variables: frame number, time in seconds, x and y position, block number, trial number, and trial type.

2.3 *New Object-Location Memory Paradigm Using Eye-Tracking*

As described in the Introduction, as well as fixing the limitations of Dr. Emrich-Mills' experiment, we had to come up with a paradigm that would eliminate the distraction of peripheral vision and force participants to move from one object to another, viewing only one at a time. And even if participants wanted to create a visuospatial map they would have to do it inside their head.

The Object-Location Memory Experiment consists of two primary stages: encoding and retrieval. During the encoding phase, participants are presented with a randomly generated arrangement of objects and instructed to form a visuospatial map by observing the locations of these objects. In the subsequent retrieval phase, participants are shown the arrangement again and must compare it to the visuospatial imagery they previously stored in memory, identifying any differences. Specifically, one object will be missing in each retrieval trial, and participants are tasked with locating where that object was and focusing their attention on the corresponding area.

Each testing block is composed of three encoding-retrieval pairs. This means that each block contains three consecutive encoding trials, followed by three retrieval trials. Alternating the encoding and retrieval trials directly would have simplified the task too much, so this structure was chosen to increase memory load and ensure varying levels of performance.

To further control the difficulty of the task, participants' vision is constrained to foveated vision during both encoding and retrieval stages. Objects are only visible when participants directly focus on their location, preventing them from relying on peripheral vision to

memorise the arrangement. This ensures that participants must actively foveate on each object to encode its location into memory. This was the most challenging part of the setup, as coding from scratch in MATLAB was required to achieve this obscured visibility. We adapted the alpha mask functionality from commonly used softwares like Photoshop and Polypop, but needed custom implementation since MATLAB does not have this feature. The alpha mask works by applying a mask layer (in this case, a cutout circle) on the layer containing the objects, making only the area inside the circle visible. The mask follows the participant's gaze, feeding from the eyetracker feedback, but smoothing the data by averaging the last 10 positions, filtering out noise like blinks and microsaccades for a more stable viewing experience. This results in a more stable and reliable visual area for the participant, ensuring that only meaningful and intentional eye movements adjust the mask position.

Each encoding or retrieval trial has a 30-second time limit. However, in retrieval trials, if a participant focuses on the same area for 5 seconds continuously, indicating they've likely identified the missing object's location, the program automatically moves to the next trial. Participants are informed of these time constraints in the instructions provided prior to the start of the experiment. Participants are encouraged to take short breaks between blocks and are reminded of the task instructions before each set of encoding or retrieval trials. This includes an instruction during each encoding trial to locate 5 hidden objects by foveating on their positions and remembering those locations for future retrieval. And to focus on identifying the missing object in the retrieval phase. A practice session precedes the main trials, allowing participants to ask questions if needed and familiarise themselves with the task, as well as identify any calibration problems. A step-by-step process is provided in the Supplementary materials.

For this experiment, 90 distinct objects were used, grouped into 18 sets of 5 objects each. The objects were theme-based to facilitate the formation of distinct visuospatial maps. For example, one set could consist of 5 vegetables, another of 5 animals, and yet another of 5 stationary items. An example list of the 15 objects used can be seen in Figure 3. To ensure that the objects were easily recognizable, they were first pre-tested on a convenience sample of UCL students, who were asked to identify each object. Only objects with an identification rate of above 95% were selected for use in the experiment.

Each object was used only once, with no repetitions. The size of each object on the screen was standardised at 50 pixels. This size was chosen to ensure that the object would be fully visible when foveated upon, but small enough to prevent multiple objects from being viewed simultaneously. The locations of the objects were determined by a pseudo-random algorithm that generated random positions on the screen while adhering to specific constraints. These constraints ensured that objects were spaced at least 100 pixels apart from each other, none of the objects shared the same x or y coordinates (within a ±15 pixel range), and no regular geometric shapes were formed by the objects.

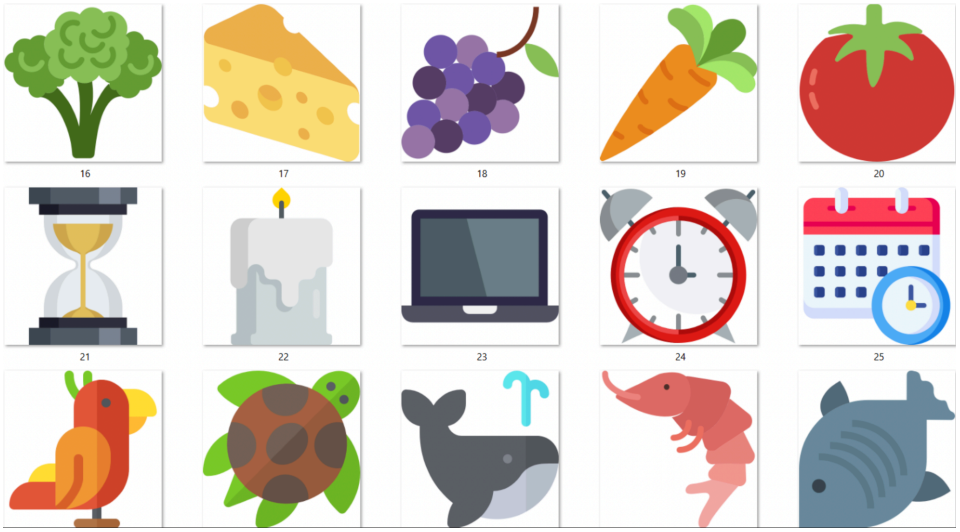

Figure 3: 15 Objects used in the experiment.

Each object was used only once, with no repetitions. The size of each object on the screen was standardised at 50 pixels. This size was chosen to ensure that the object would be fully visible when foveated upon, but small enough to prevent multiple objects from being viewed simultaneously. The locations of the objects were determined by a pseudo-random algorithm that generated random positions on the screen while adhering to specific constraints. These constraints ensured that objects were spaced at least 100 pixels apart from each other, none of the objects shared the same x or y coordinates (within a ±15 pixel range), and no regular geometric shapes were formed by the objects.

This process initially produced 30 random arrangements, of which the most similar ones were discarded. Ultimately, 23 unique arrangements were selected, and 18 of these were used in the actual experiment. Additionally, the object positions file contained the normalised coordinates of the objects presented during each trial. This was done, so that positions could be converted into any screen coordinates based on the resolution of the testing screen (1920x1080 pixels in our case), ensuring accurate representation of object locations during the task.

2.4 *Participants and Procedure*

Seventeen healthy participants (10 female, 7 male) aged between 23 and 35 years (M = 27.70, SD = 3.60) were recruited for this study. Informed consent was obtained from all participants prior to their involvement in the experiment. Participants were required to have no visual impairments, as even corrected vision (e.g., glasses) could interfere with the experiment. Specifically, the lenses of glasses could reflect the infrared (IR) illumination used by the eye-tracker, creating glare in the camera (Roux, Passerieux, Ramus, et al., 2014). One participant was excluded from the study due to excessive head movement, which affected the calibration and prevented completion of the experiment. Additionally, one participant had to reschedule due to calibration issues caused by wearing eyeliner, which interfered with accurate tracking.

After consenting, participants filled out a standard demographic questionnaire and underwent an eye-tracker calibration procedure. The calibration followed a 9-dot routine as outlined in the EyeLink® 1000 Plus User Manual (SR Research, 2018). The apparatus, detailed in the Materials and Setup section, was adjusted for each participant as needed. Only after achieving a successful calibration, as determined by the eye-tracker software, were participants allowed to proceed. Participants then engaged in a practice session, which consisted of a block (3 encoding and 3 retrieval trials) identical in structure to the main experiment, ensuring they were familiar with the task requirements. Upon confirming verbally that they understood the assignment, participants began the experimental trials, which included 15 encoding and retrieval trials as previously outlined in Section 2.3. The entire procedure, including practice and experimental trials, took an average of 45 minutes to complete. At the conclusion of the experiment, participants were debriefed and asked to provide feedback on any memory techniques they had used during the trials. This feedback was collected and further addressed in the thesis' discussion section.

2.5 *Data Processing*

In this experiment, the data processing was conducted using a custom script to filter and visualise the eye-tracking data, alongside object position data during the encoding and retrieval phases of the task. The primary goal of this processing was to prepare the data for the analysis to compare participants' scanpaths during both phases and link these to memory performance, specifically focusing on the differences between encoding and retrieval phases in terms of visual exploration.

The data processing began by converting the resulting EDF files from the EyeLink 1000 Plus eye-tracking system for each participant into Matlab. There the data was organised to contain key variables such as frame number, time in seconds, x and y gaze positions, trial number, block number, and type of task (either encoding or retrieval). The table was then cleaned by removing any missing values and ensuring that it contained the correct number of columns. A smaller table for each block and trial combination was generated for each participant, allowing the data to be divided into subsets for detailed analysis. This process enabled the comparison of eye movement data within specific trials and blocks to evaluate individual scanpaths during encoding and retrieval. The raw data was then visualised, and random trials were checked as demonstrated in Figure 4. This visualisation allowed for the comparison of the scanpaths during retrieval to the original object positions, highlighting the areas where participants focused their attention. This served as a check, ensuring that participants paid attention and actually performed what was asked of them in the task.

Next I focused on filtering the data, by applying different algorithms to filter out the noise and detect fixations from eye-tracking data, using a spatial and temporal threshold to identify periods when the participant's gaze remained relatively stable. All of the algorithms used according to the previously documented procedures (Das et al. 1996; Raju et al., 2023). Fixations were identified by comparing the current gaze position with the centre of the ongoing fixation. If the spatial distance between consecutive gaze points remained below a

defined threshold (50 pixels) and the temporal gap between them was under 0.1 seconds, the gaze points were considered part of the same fixation. If the distance exceeded these thresholds, the current fixation was saved, and a new fixation began. This process was repeated for all gaze data points, resulting in a sequence of fixations with associated start and end times used to construct fixation durations, and can be viewed in Figure 5.

After identifying and segmenting fixations, the data was further transformed into data sets for each trial and each participant. These sets contained only the x, y coordinates of each fixation and the corresponding fixation duration. This format was essential for running the data through the MultiMatch algorithm (Dewhurst, Nyström, Jarodzka, et al., 2012), described in the next section, comparing encoding-retrieval scanpaths by calculating similarity scores across five dimensions: fixation position, fixation duration, saccade shape (vectors), saccade direction, and saccade length. For each trial, a table was generated containing the MultiMatch similarity scores as in Figure 6.

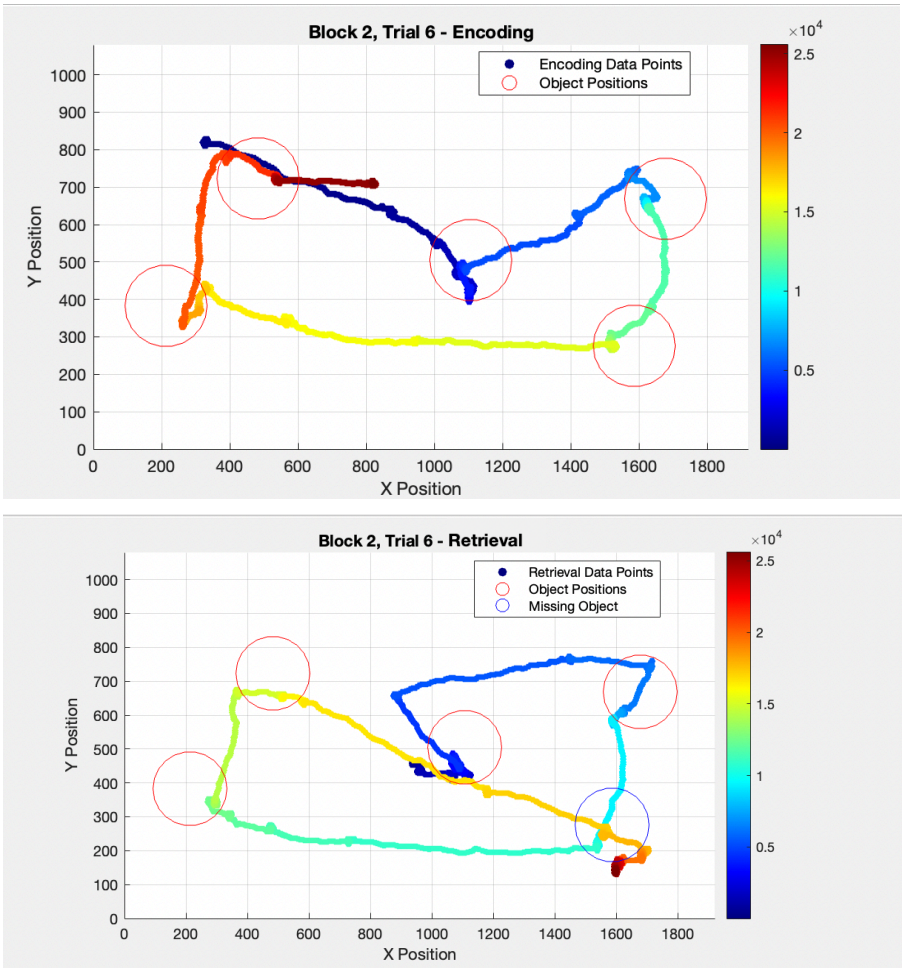

Figure 4: Raw Data Encoding and Retrieval Block 2, Trial 6. *The encoding phase is visualised by plotting the participants' eye movements as a scatter plot overlaid with the object positions. A colour gradient was applied to visualise the sequence of eye movements over time, and the positions of the objects were marked with bold red circles. For the retrieval phase, additionally, the missing object was indicated by a blue circle.*

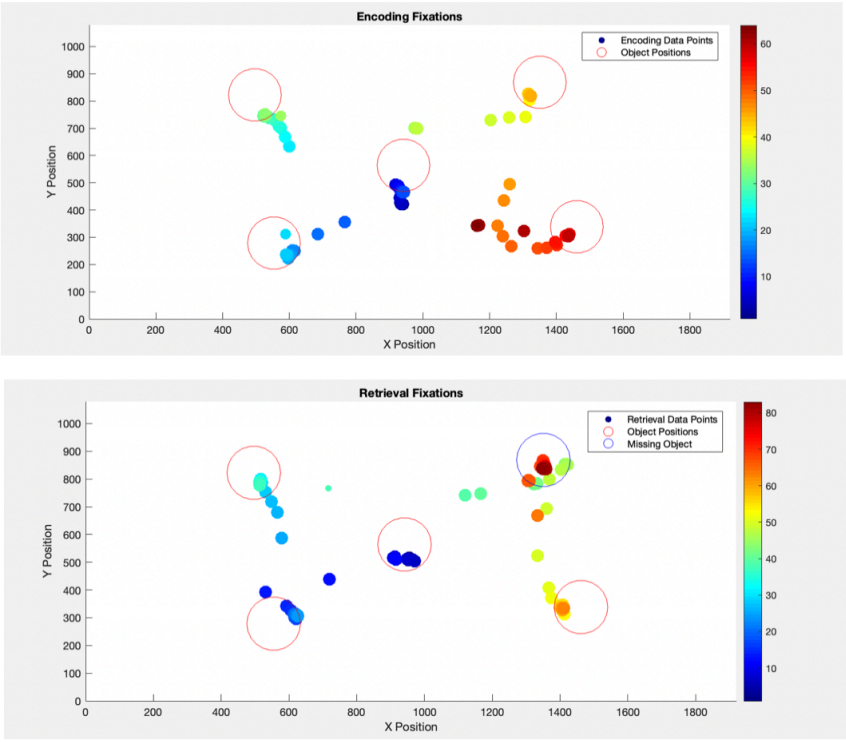

Figure 5: Fixations Data Encoding and Retrieval Participant 5, trial 13 . *The encoding phase is visualised by plotting the participants' eye movements as a scatter plot overlaid with the object positions. A colour gradient was applied to visualise the sequence of eye movements*

	Trial_1	Trial_2	Trial_3	Trial_4	Trial_5
VectorSimilarity	0.9286	0.85907	0.95004	0.9652	0.9793
DirectionSimilarity	0.89681	0.78339	0.95455	0.92391	0.96992
LengthSimilarity	0.92998	0.95295	0.92247	0.95653	0.96423
PositionSimilarity	0.86305	0.71772	0.93664	0.95174	0.97221
DurationSimilarity	1	0.9967	1	1	1

Figure 6: Similarity table for the first 5 trials for participant 12.

Additionally, another important metric was computed to quantify memory performance: the error distances from the missing object. This metric measured the spatial difference between where participants thought the missing object was and its actual location in the original encoding phase.

2.6 *MultiMatch algorithm*

The MultiMatch (MM) algorithm, introduced by Dewhurst et al. (2012), is a method for comparing eye movement scanpaths across multiple spatio-temporal dimensions. It was designed to provide a quantitative understanding of how sequential eye movements unfold, capturing both the temporal and spatial properties, unlike previous methods like, for example,

ScanMatch (Christino, et al. 2010). The MM algorithm simplifies two scanpaths, converting them into ordered sequences of connected saccadic vectors. These vectors represent eye movements between fixations, and the algorithm uses them to compare two scanpaths across several dimensions. The goal is to retain the essential structure of the scanpath while removing unnecessary noise from small saccades or closely grouped fixations. This ensures that the final comparison reflects the meaningful elements of the eye movements without being affected by minor, local deviations.

The MultiMatch algorithm was chosen as a primary analysis for my research because, unlike other simple methods that focus solely on fixation points or static eye movement patterns, MultiMatch provides a vector-based comparison that accounts for both spatial and temporal dimensions of eye movements, which is essential for understanding the nature of visual exploration. Moreover, by comparing sequential eye movement patterns across encoding and retrieval phases, MultiMatch allows us to detect these vector-based relationships and directly check the prediction of Bicanski & Burgess' (2019) computational model. In the nutshell, MultiMatch uses Dijkstra's algorithm (1959) to find the shortest path through a matrix that represents all possible vector pairings between the two scanpaths. This technique minimises the overall difference between the scanpaths by accounting for all potential pairings of saccadic vectors. The final comparison is based on five dimensions:

1. **Fixation Position**: The exact spatial locations of fixations.
2. **Fixation Duration**: Measuring how long each fixation lasted.
3. **Saccade Shape**: Geometric shape of the saccadic vector paths.
4. **Saccade Direction**: Direction of eye movements between fixations.
5. **Saccade Length**: Measuring the distance between consecutive fixations.

This approach allows us to capture the subtle variations in how participants' eyes move across stimuli. Furthermore, its capacity to simplify and align scanpaths ensures that small, irrelevant movements, or overfixations on some objects do not bias the results. The algorithm was chosen to explore how encoding-recollection similarity (SERS) manifests and how specific scanpath properties correlate to memory performance, potentially providing a deeper understanding of the connection between eye movements and memory.

2.7 *Data Analysis*

To investigate whether and how scanpaths from encoding are reinstated during recall, I first calculated the MultiMatch similarity scores across all five dimensions—position, duration, shape, direction, and length—for each participant and each trial. These scores quantified how closely participants' eye movements during the encoding phase resembled their scanpaths during retrieval, specifically focusing on how sequential reinstatement of saccades occurred.

To establish the significance of the similarity scores for the scanpaths and determine whether scanpaths are more similar within trials than between trials, I conducted a permutation analysis. This involved comparing the scanpaths produced by each participant during the

encoding of one trial with those produced during the recall phase of other trials of the same type. The method involved creating a distribution of means from the similarity scores of these non-matching encoding-retrieval pairs across all trials. To achieve this, I performed a random permutation analysis, where for each permutation, encoding data from a specific trial was paired with retrieval data from a different, randomly selected trial of the same participant, and calculated the MultiMatch similarity score for each pair. After obtaining the mean score for this permuted set, the retrieval trials were randomly shuffled again, repeating the process until a smooth distribution of means was generated (with the total number of means equaling 2000). By averaging the similarity scores from these non-matching pairs, I calculated distributions of sequential similarity scores for each participant, representing the expected level of similarity by chance across different trials. This provided a reference point to compare with the actual within-trial similarity scores for each participant.

The mean similarity score from the true encoding-recall pairs was then plotted on the distribution to observe where it fell. If the scores were to exceed the 98.75th percentile (this is due to the Bonferroni correction of a 95th percentile threshold, as explained below), it indicated that the within-trial scanpath similarity was significantly higher than could be expected by chance, supporting the hypothesis about the sequential reinstatement of eye movements during recognition memory.

Given the hypothesis that better reinstatement of scanpaths leads to better memory performance, after obtaining similarity scores for each trial, to account for individual participant variability and trial-specific effects, I employed a linear mixed-effects model. Using MATLAB's fitlme function, I specified a model with Shape, Direction, Length, and Position Similarities as fixed effects. The duration dimension was not examined in the correlation analysis because the similarity scores for this dimension were consistently high, often approaching 1. This lack of variability in the duration scores was due to the time constraints in the experiment, which limited the differences in fixation durations between encoding and retrieval phases. As a result, the duration scores provided little meaningful variation to explore in relation to memory performance. Random intercepts were included for both Participant and Trial, allowing the model to account for inherent differences across participants and the variability associated with each trial. The model was structured as follows:
DistanceError ~ Similarity1 + Similarity2 + Similarity3 + Similarity4 + (1 | Participant) + (1 | Trial).

The model's fixed-effect coefficients are used to interpret the direction and strength of each similarity measure's relationship with the distance error, and, therefore, the memory performance — a negative coefficient would indicate that higher similarity predicts better memory performance, while a positive one would suggest the opposite. Additionally, the random intercepts for Participant and Trial could reveal variability in error distances attributable to individual differences and trial-specific factors, helping to control for baseline performance.

Additionally, I conducted a correlational analysis between the error distances (i.e., how far participants' gaze deviated from the correct location of the missing object during retrieval) and four of the MultiMatch dimensions—vector, direction, length, and position—from the similarity table. To further test the hypothesis that better reinstatement of scanpaths leads to better memory performance, I performed a one-tailed test, focusing on the negative correlation between error distances (the distance between a participant's gaze and the correct location of the missing object) and the similarity scores across the four MultiMatch dimensions (vector, direction, length, and position). Therefore, for each participant, I computed four r-values and p-values corresponding to the correlation between their errors and the similarity scores in each dimension. Since each participant had four separate correlations, one for each dimension, the risk of committing a Type I error increases. To control for this risk, I applied a Bonferroni correction for adjusting the significance threshold when performing multiple comparisons. The correction ensures that the probability of obtaining a false positive remains controlled across all the tests. As the Bonferroni correction works by dividing the original alpha level (in this case, 0.05) by the number of comparisons being made, and since I conducted four comparisons per participant, the corrected significance threshold was calculated as follows:

$$\alpha_{\text{corrected}} = \frac{0.05}{4} = 0.0125$$

Thus, after applying the Bonferroni correction, only p-values below 0.0125 were considered statistically significant.

To evaluate participants' overall performance and determine whether there was a significant relationship between scanpath similarity and memory performance, I plotted each participant's correlation scores (r) between the error distances and similarity scores across four dimensions. This allowed me to visually assess the distribution of correlation values across participants. Next, I calculated the mean correlation score for all participants in each dimension and compared this mean to 0, which represents the expected correlation due to chance. Therefore, I performed a one-sample t-test. This test was used to determine whether the mean correlation score across all participants was significantly different from 0, which represents the null hypothesis of no relationship between scanpath similarity and memory performance. By conducting the t-test, I was able to assess whether the observed correlations were significantly below 0, indicating a significant relationship between scanpath reinstatement and memory performance, or whether the observed correlations could be attributed to random chance. A significant result from the t-test would provide evidence that higher scanpath similarity is indeed predictive of better memory performance, supporting the idea that eye movements during encoding and retrieval play a crucial role in recognition memory.

To further investigate the sequential encoding-recollection similarity (SERS), I compared each participant's encoding scanpaths with those produced by all other participants during the

retrieval phase of the same image. This across-participant comparison helped assess whether the scanpath reinstatements were truly individual-specific or if participants tended to follow similar scanpaths when recalling the same images. The rationale behind this comparison was to understand whether within-participant scanpath similarities (between their encoding and recall phases) were significantly different from the between-participant similarities. If all participants were to generate highly similar scanpaths for the same image, it could suggest that the task itself leads to universal visual exploration patterns, which would require reinterpreting the results. On the other hand, if the within-participant SERS scores are higher than those from between-participants, this would provide stronger evidence for individual-specific scanpath reinstatement supporting memory recall.

The way this was done is similar to the permutation test described above, however, now it involved between participants comparisons. Each trial was assessed separately, comparing the scanpaths produced by each participant during the encoding of one trial with the retrieval scanpaths of other participants for the same trial. The method involved generating a distribution of means from the similarity scores of these non-matching encoding-retrieval pairs across all participants. Specifically, for a given trial, I randomly paired encoding data from a specific participant with retrieval data from a different participant, continuing this process through all participants. After obtaining the mean similarity score for this shuffled set of 16 pairings, I repeated this process 2000 times until a smooth distribution of means was generated. By averaging the similarity scores from these non-matching encoding-retrieval pairs, I calculated a distribution of sequential similarity scores for each trial, representing the expected level of similarity by chance across participants. And again the true means of the correct encoding-retrieval pairs were plotted, and if the true score exceeded the 98.75th percentile of the distribution, it indicated that the within-participant scanpath similarity was significantly higher than expected by chance.

3. RESULTS

3.1 Are Scanpaths More Similar Within Than Between Trials?

Participants' within-trial scanpath similarity scores across all five dimensions (position, duration, shape, direction, and length) were calculated using the MultiMatch method. The similarity scores for each participant and each dimension are displayed in the Supplementary Table 1.

To assess the significance of scanpath similarity within participants, I conducted a permutation analysis across four dimensions: Shape, Direction, Length, and Position. This analysis revealed 31 significant within-trial similarity scores across participants (p-values below 0.0125), indicating that observed scanpath patterns were unlikely to result from chance alone:

- **Shape Similarity** was significant for Participants 1, 2, 5, 7, 9, 10, 11, and 13.
- **Direction Similarity** was significant for Participants 1, 5, 6, 7, 9, 10, and 11.
- **Length Similarity** was significant for Participants 1, 5, 9, 10, 11, and 13.
- **Position Similarity** was significant for Participants 1, 3, 4, 5, 6, 7, 8, 9, 10, and 11.

To evaluate whether this number exceeded what could be expected by chance, a binomial distribution analysis was conducted. With the Bonferroni-corrected chance threshold of 1.25% per metric, the expected number of significant results by chance was calculated to be 0.80. As illustrated in Figure 7, which depicts the binomial distribution, the actual observed number of significant results substantially exceeded this expectation (p-value of 1.2 × 10^-41). This outcome indicates that the observed within-participant scanpath similarities are highly significant.

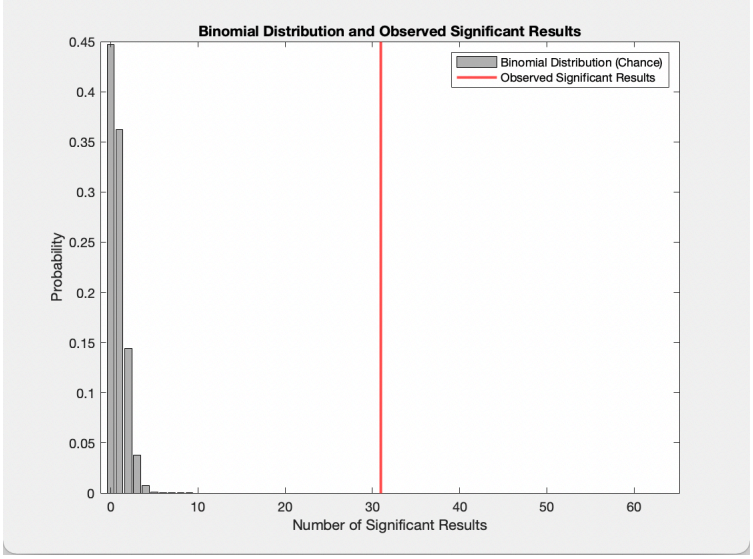

Figure 7: Binomial Distribution and Observed Significant Results. *This figure illustrates the binomial distribution of expected significant results by chance (grey bars) compared to*

the actual observed significant results (red line) across four similarity metrics and multiple participants.

3.2 Relationship Between Scanpath Reinstatement and Memory Performance

3.2.1 The Mixed-Effects Model

The mixed-effects model revealed that only Position similarity had a statistically significant effect on Distance Error ($F(1, 235) = 9.87$, $p = 0.0019$), suggesting that better positional similarity between encoding and retrieval scanpaths was associated with lower error distances, as shown in the Figure 8. The random effects covariance parameters showed the variability in error distances attributable to participants, but not trial. The random intercept standard deviation for participants is estimated at 15.92, suggesting that there are individual differences in memory performance or visual encoding strategies between participants. The standard deviation for trial variable is very close to zero (8.171e-07), indicating minimal trial-to-trial variability beyond what is explained by the fixed effects and participant-level random intercepts.

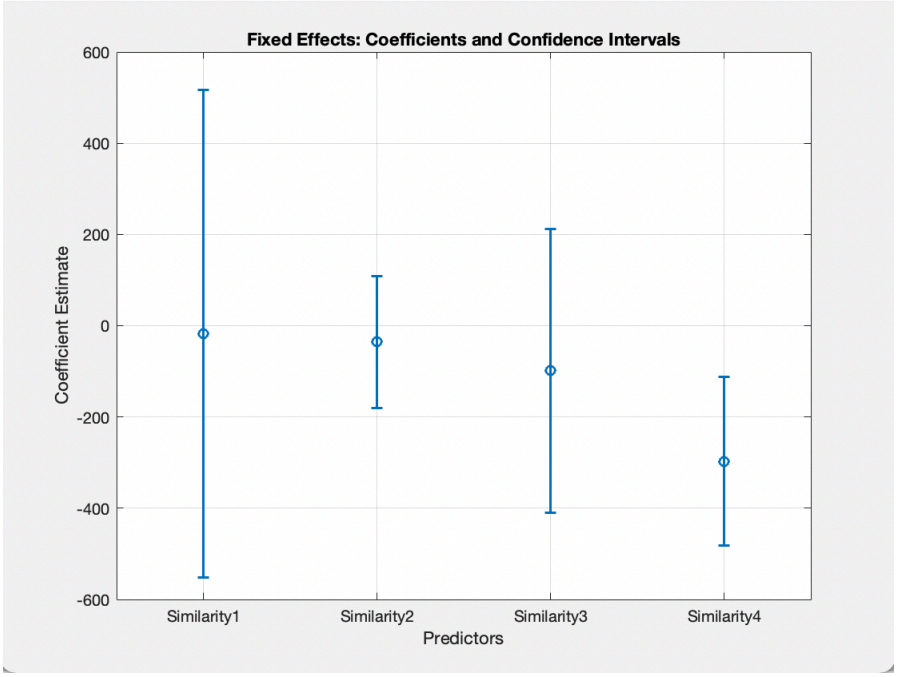

Figure 8: Fixed Effects Coefficients and Confidence Intervals for Similarity Measures in Predicting DistanceError.*The figure displays coefficient estimates and 95% confidence intervals for four similarity measures — Shape, Direction, Length, and Position Similarity — from a linear mixed-effects model predicting Error Distances.*

3.2.2. Across-Participant Correlation Analysis

For each participant, a correlation analysis was conducted between their error distances (the deviation between the participant's gaze and the correct location of the missing object) and their scanpath similarity scores across four dimensions: shape, direction, length, and position. This resulted in four correlation coefficients (r-values) and associated p-values for each participant, shown in Figure 9.

	ShapeSimilarity	DirectionSimilarity	LengthSimilarity	PositionSimilarity
P1	0,03975	0,08407	0,00797	0,09541
P2	0,02080	0,15170	0,39727	0,04334
P3	0,63684	0,25108	0,87547	0,64429
P4	0,00693	0,30432	0,01467	0,00073
P5	0,01035	0,03530	0,29189	0,31058
P6	0,63663	0,01979	0,75430	0,07275
P7	0,79234	0,56978	0,23008	0,85313
P8	0,02606	0,02213	0,12103	0,20155
P9	0,01420	0,02147	0,60324	0,00029
P10	0,03777	0,00508	0,09402	0,25590
P11	0,91294	0,92004	0,28513	0,61294
P12	0,13558	0,86620	0,29739	0,16423
P13	0,81643	0,15364	0,08521	0,76209
P14	0,77408	0,36907	0,24788	0,00872
P15	0,89734	0,44567	0,18040	0,02242
P16	0,87729	0,49994	0,31344	0,97425

Figure 9: Participants' p-values. *The table presents the p-values for each participant's correlation scores between error distances and similarity scores across four dimensions: ShapeSimilarity, DirectionSimilarity, LengthSimilarity, and PositionSimilarity. Each row represents a participant, and each column corresponds to one of the four dimensions. P-values that are significant after Bonferroni correction ($p < 0.0125$) are highlighted in red, while p-values that are significant at the standard 0.05 level are highlighted in green.*

Out of the 16 participants, several showed significant negative correlations at the $p < 0.05$ level across different dimensions, with a smaller number remaining significant after applying the stricter Bonferroni correction ($p < 0.0125$). Position Similarity demonstrated the strongest effect, with several participants retaining significance even after correction. Fewer participants showed significant correlations in Shape, Direction, and Length dimensions. Notably, some participants showed significant correlations in multiple dimensions simultaneously, like, participant number 2. The results across participants showed considerable variability, making it difficult to decisively reject the null hypothesis based solely on individual participant correlations. While some participants demonstrated significant correlations in multiple dimensions, the overall inconsistency—both before and after the Bonferroni correction—suggests that the relationship between scanpath similarity and memory performance may not be that uniform across individuals. However, the Across-Participant Mean Correlation Analysis might shed more light on the general trend.

To assess the overall relationship between scanpath similarity and memory performance, the mean correlation scores (mean r-values) were calculated for each of the four dimensions by averaging the individual r-values from all participants. The mean r-values per participant were then compared to 0 (the null hypothesis of no relationship between scanpath similarity and memory performance), using one-sample t-tests.

As shown in Figure 11, the Across-Participant Mean Correlation Analysis provided strong evidence supporting the hypothesis that greater scanpath similarity between encoding and

retrieval phases is associated with better memory performance. All three dimensions—ShapeSimilarity ($p = 0.0034$), DirectionSimilarity ($p = 0.0011$), and PositionSimilarity ($p = 0.0001$)—demonstrated significant negative correlations with memory performance, with PositionSimilarity showing the strongest effect. However, LengthSimilarity ($p = 0.0228$) did not reach statistical significance after the correction, indicating that saccade length may be less relevant to memory performance in this task. Notably, ShapeSimilarity appears to exhibit two distinct groups of participants, as visualised in the data, suggesting that different participants may have employed varying visual encoding strategies, highlighting the potential importance of individual differences.

3.3 Are Scanpaths More Similar Within Than Between Participants?

To assess whether the scanpath reinstatements were participant-specific or whether all participants followed similar visual exploration patterns during recall, each participant's encoding scanpaths was compared with the retrieval scanpaths produced by other participants for the same object locations. This analysis revealed a total of 29 significant results across the trials, far exceeding the expected 0.75 results by chance ($p < 5.07 \times 10^{(-39)}$), indicating that the observed scanpath patterns across participants were unlikely to be random, as in the Figure 10.

- **Shape Similarity** was significant for Trials 1, 2, 4, 6, 8, 9, 10, 11, 12, 13, 15.
- **Direction Similarity** was significant for Trials 3, 5, 6, 7, 11, 12, 13, 15.
- **Length Similarity** was significant for Trials 1, 5, 9, 10, 11, 13.
- **Position Similarity** was significant for Trials 2, 6, 7, 8, 10, 11, 12, 13, 15.

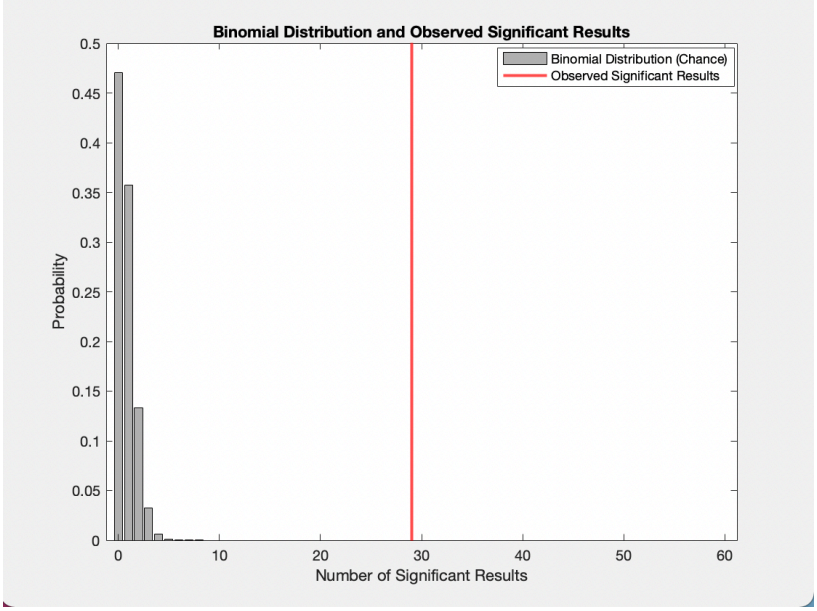

Figure 10: Binomial Distribution for Across Participants Analysis and Observed Results. *This figure illustrates the binomial distribution of expected significant results by chance (grey bars) compared to the actual observed significant results (red line) across four similarity metrics, multiple participants and trials.*

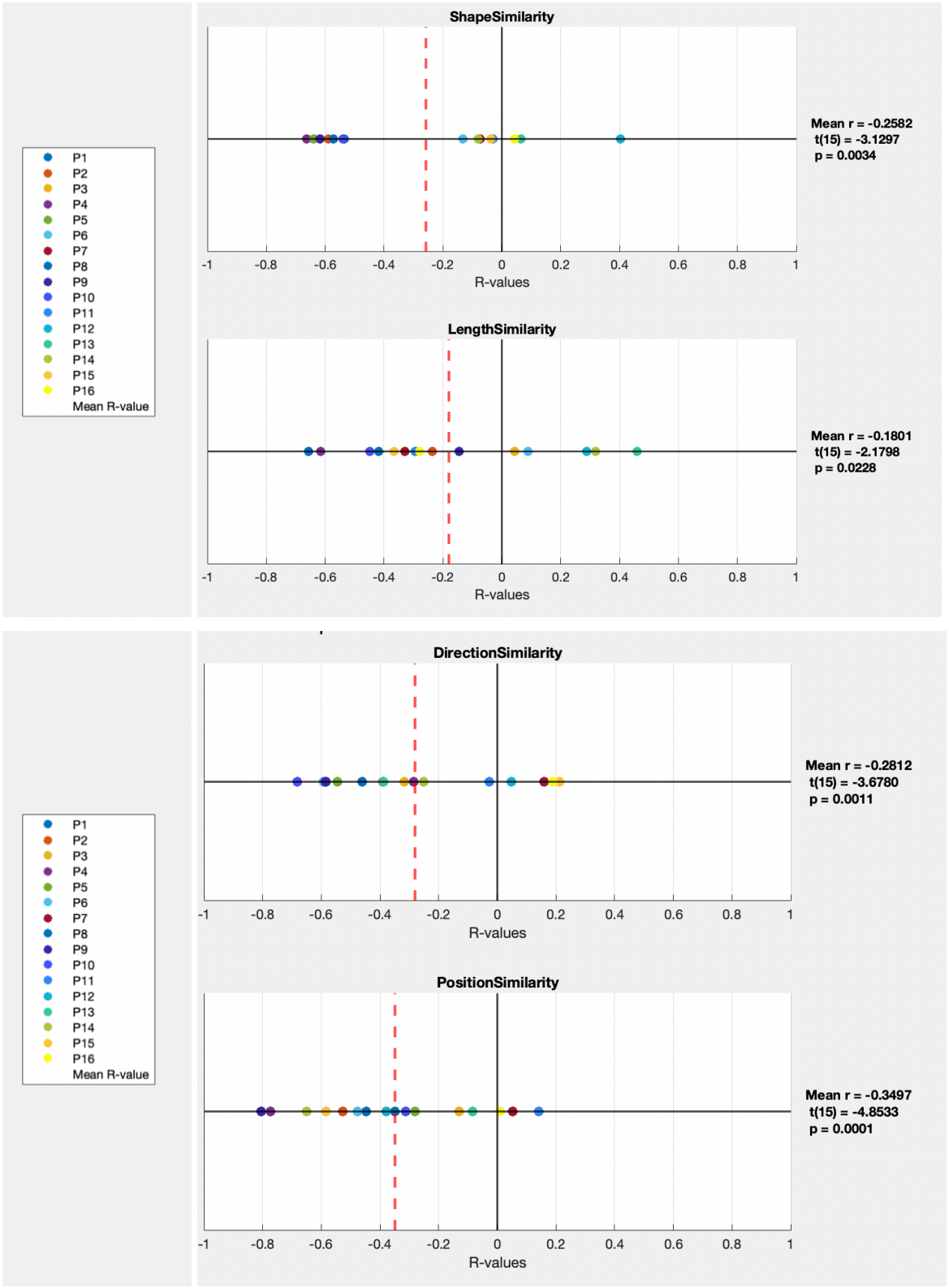

Figure 11: Participants' Correlation Scores. *The figure shows four scatter plots with participants' correlations between error distances and similarity scores across four dimensions: Shape, Direction, Length, and Position. Each dot represents a participant, coloured uniquely. A solid black vertical line marks r = 0, representing chance-level correlation, while a dashed red vertical line shows the mean correlation for each dimension.*

4. DISCUSSION

In this project I designed a novel memory task that required participants to identify the missing object from an array, aiming to compare eye movements during encoding and retrieval. The main hypothesis was that scanpaths during these two phases would exhibit similarity, and that this similarity would correlate with memory performance. This hypothesis comes from a model of visual recognition memory in which the relative locations of visual features are encoded as a series of movements between them, and retrieval consists of testing for the presence of predicted features at the end of eye movements from other features.

4.1 *Are Scanpaths More Similar Within Than Between Trials?*

The results of this study provide strong evidence that support the hypothesis that eye movements are reinstated during recognition memory, revealing a sequential encoding-recollection similarity across multiple dimensions (Shape, Direction, Length, and Position). The permutation analysis of within-participants scanpath similarities found significant within-trial similarity scores, suggesting that these patterns reflect intentional gaze behaviour rather than random eye movements. It also implies that the similarity scores obtained reflected memory-driven behaviour rather than individual visual tendencies. These findings were expected, as they fully support previous research on SERS (Wynn, Shen, & Ryan, 2019; Johansson & Nyström, 2022).

4.2 *Are Scanpaths More Similar Within Than Between Participants?*

The results from the between-participant analysis further clarify the extent to which individual-specific scanpath patterns support memory recall. The significant scanpath similarities across multiple trials and dimensions suggest that, while certain visual exploration patterns are consistently caused by specific stimuli, the participant's memory recall process remains unique. These results highlight the role of individual-specific SERS and support the hypothesis that, despite some universal exploration tendencies, participants rely heavily on personalised visual strategies, which are often reinstated between encoding and recall.

4.3 *Relationship Between Scanpath Reinstatement and Memory Performance*

As for the relationship between SERS and memory performance, the linear mixed-effects model and the across-participant mean correlation analysis offer complementary but somewhat differing perspectives on the relationship between scanpath similarity and memory performance.

The linear mixed-effects model analysis provided the first insight into the relationship between scanpath similarity and memory performance. This model indicated that only Position Similarity was significantly associated with reduced error distances, suggesting that alignment in positional eye movements between encoding and retrieval was the primary factor linked to improved memory performance. It is important to note also that the coefficient estimate interval for Shape similarity was substantially larger than those of the

other dimensions. This high variability, as well as the significance of the random intercept for participants, suggested that it is possible that some subjects were indeed aligning their scanpath shapes with memory performance, but others were not, prompting further investigation. Given the potential importance of Shape similarity, it seemed necessary to examine the relationship between memory performance and SERS through a more detailed analysis to capture the nuance of individual differences, leading to a correlation analysis of these dimensions.

The across-participant mean correlation analysis provides strong evidence that scanpath similarity across several dimensions, specifically shape, direction, and position, is significantly associated with memory performance in a visual recognition task. Participants who demonstrated greater similarity in their eye movements between the encoding and retrieval phases tended to perform better, as indicated by lower error distances in locating the missing objects. The fact that shape, direction, and position dimensions all showed significant relationships with memory performance suggests that participants are not only revisiting the same locations during retrieval but are also replicating the spatial path and movement patterns from encoding. This pattern of reinstatement aligns with the previous theories on the relationship between SERS and memory performance (Johansson & Nyström, 2022; Bone, Ritchie, Millin, & O'Connor, 2022; Foulsham, Barton, & Kingstone, 2018), including the possible involvement of grid cells in guiding eye movements (Bicanski & Burgess, 2019). By replicating the saccadic paths from encoding, participants may be utilising a cognitive mechanism similar to how the brain processes spatial relationships during navigation, further indicating that eye movements serve as a spatial mapping tool for memory retrieval.

It is important to note also that in individual correlation scores for the Shape dimension, there appears to be a clustering pattern, with participants forming two distinct groups: one with negative correlations between distance error and similarity score and the other closer to zero, but also negative, as shown in the Figure 11. This clustering suggests that participants may be using different visual encoding and memory strategies, which aligns with the findings from the linear mixed-effects model, further indicating variability and less consistency in how participants engage with this aspect of scanpath similarity. This variability may imply that, while some participants display strong alignment in Shape similarity with improved memory performance, others do not. This variability could explain why, despite previous finding, the Shape similarity did not reach significance in the mixed-effects model in our experiment. To explore this further, the correlation scores between Shape and Direction similarity were analysed, yielding a significant result ($R = 0.6874$, $p = 0.0033$). Notably, this was the only significant correlation observed among the different dimension pairs, indicating that Shape and Direction may share a unique relationship in how they support memory performance, and that there might indeed be two distinct groups of participants. These findings for the Shape dimension can be explained also by the small sample size in this study, and, therefore, increasing the number of participants could help clarify these trends and provide a more robust assessment of the Shape dimension's role in memory performance.

The strongest correlation with memory performance in the current study was found in the

position similarity dimension. This result is surprising, given that during the retrieval phase, participants could still view the objects (except for the missing one) on the screen, unlike in the Johansson and Nyström's (2022) study, where participants were required to recall the scene while staring at a blank screen. The presence of the objects likely made it easier for participants to match their eye movements with previously encoded positions, which should have resulted in consistently high (if not uniform) similarity scores for the position dimension, but it did not. Therefore, it is interesting to see the stronger correlations for the position similarities, and it requires closer examination in future research to fully understand its implications. One possible explanation is that the large size of the objects on the screen allowed participants who focused on and revisited precise details of each object, rather than only approximate positions, to achieve better memory performance. This suggests that focused attention on specific object features, rather than general positional revisits, may enhance memory recall by supporting more detailed visual encoding.

It is also important to note that, similar to the Johansson and Nyström (2022) study's findings, SERS in length and duration were insignificant in our study. They argued that these properties are typically dependent on visual features from the image, which are absent or obscured, when participants recall from a blank screen. This likely applies to our study as well, and that would explain why the length of saccades was not found to play a significant role in aiding memory performance.

Overall, these results provide robust support for the hypothesis that certain dimensions of eye movement patterns, particularly shape, direction, and position are reinstated during memory recall. Position similarity, in particular, was also found to play a crucial role in memory performance. However, the observed insignificant result and variability within the shape dimension asks for closer examination in future studies to understand the nuances in individual encoding and retrieval strategies that might influence this relationship.

4.4 Directions for Future Research

One of the most promising directions for future research is to extend the current paradigm to face recognition. Similar to navigating between objects, it would be valuable to explore how participants navigate between salient facial features during encoding and retrieval phases, also following the theoretical framework suggested by Bicanski & Burgess' (2019) model. Future studies could replicate the current experiment using faces as stimuli, allowing us to see if participants' eye movements during face recognition also exhibit sequential reinstatement similar to that observed with object recognition. It is important to note that the large body of literature already exists identifying specific brain regions—such as the fusiform face area (FFA)—that are activated during face recognition (Freiwald, 2020). These regions are thought to be specialised for holistic face processing, which differs from the way relational memory is typically studied with object locations. Therefore, the new experimental paradigm would be crucial in ensuring that the face is not perceived as a whole, but rather as a set of individual salient features, such as eyes, nose, and mouth.

Furthermore, given that current research does not yet investigate the direct role of grid cells in memory-guided eye movements. A logical next step would be to combine the Object-Location Memory Test paradigm developed here with neuroimaging techniques, such as fMRI, to directly observe whether grid-like representations in the entorhinal cortex are activated during visual recognition tasks. This could provide further evidence connecting grid cell activity with visual memory in humans. The same can be done with the face recognition experiment. In a study that came out in September 2024, called "Entorhinal grid-like codes for visual space during memory formation", researchers did a similar experiment investigated the role of saccade-based grid-like codes in the human entorhinal cortex during memory formation (Graichen et al., 2024). Similarly to what Bicanski & Burgess (2019) predicted they found that these grid-like codes were consistently present as participants studied visual scenes. However, it remains unclear whether the signals encode relational information between scene elements or simply respond to eye movements. Interestingly the researchers found that stronger grid signals were associated with worse individual recognition memory. The way researchers explained this result was that participants with better recognition memory may have used a different strategy, integrating prior knowledge or schemas to enhance memory performance, rather than relying solely on visuospatial encoding. Our new paradigm is well-suited to explore this possibility, as it controls for such external influences by restricting peripheral vision and forcing participants to rely primarily on eye movement-based encoding, allowing for a clearer future assessment of grid-cell involvement in memory.

Additionally, it would be fascinating to explore the role of grid cells in memory recognition in mental illnesses. For example, recent research has indicated that grid-like theta modulation is reduced in patients with schizophrenia (Convertino et al., 2023). And if grid-like representations are confirmed to aid memory performance, this could suggest that grid cell dysfunction may not be limited to spatial processing but could also affect cognitive functions, in psychiatric conditions. A logical extension of this research would be to apply the current paradigm to populations with mental illnesses, particularly schizophrenia, to see if there are deficits in the reinstatement of scanpaths and whether these deficits are associated with altered grid-like neural activity.

Another intriguing area for future investigation is whether grid cells play a role in other cognitive functions, besides navigation and memory. Theoretical research by Chen et al. (2022) such as "*Are Grid-Like Representations a Component of All Perception and Cognition?"* suggests that these neural mechanisms, grid cells, may be supporting varied perceptual and cognitive functions across different sensory inputs. Exploring this possibility would provide a more comprehensive understanding of whether grid cells are truly a universal representation system.

Lastly, studies like Schroeger et al (2020) and Hwang et al. (2012) on spatial and temporal eye–hand coordination emphasises the role of the parietal reach region (PRR) in coordinating eye and hand movements. This region could also be explored in future research in relation to grid cell activity, combining eye and hand tracking with neuroimaging to determine if the

entorhinal cortex and PRR work together to guide coordinated eye-hand movements during spatial memory tasks and exhibit grid-like representations.

5. REFERENCES

Anderson, N. C., Anderson, F., Kingstone, A., & Bischof, W. F. (2015). A comparison of scanpath comparison methods. *Behavior Research Methods, 47*(4), 1377–1392. https://doi.org/10.3758/s13428-014-0550-3

Andersson, R., Larsson, L., Holmqvist, K., Stridh, M., & Nyström, M. (2017). One algorithm to rule them all? An evaluation and discussion of ten eye movement event-detection algorithms. *Behavior Research Methods, 49*(2), 616–637. https://doi.org/10.3758/s13428-016-0738-9

Armson, M. J., Barker, R., & Phillips, M. (2020). Memory for gallery tours: The influence of eye movements on detailed recall in high visual imagers. *Memory & Cognition, 48*(5), 872–883. https://doi.org/10.3758/s13421-019-00988-1

Barker, R., & Armson, M. J. (2024). Eye movements precede detail production during memory recall: Implications for high imagers. *PsyArXiv Preprints.* https://psyarxiv.com/4t9vj/

Bicanski, A., & Burgess, N. (2019). A computational model of visual recognition memory via grid cells. *Current Biology, 29*, 979–990. https://doi.org/10.1016/j.cub.2019.01.077

Bobak, A. K., Mileva, V. R., & Hancock, P. J. B. (2022). Face-Information Sampling in Super-Recognizers. *Psychological Science, 33*(4), 567–580. https://doi.org/10.1177/09567976221096320

Bone, M. B., Ritchie, J. B., Millin, R., & O'Connor, A. R. (2022). Eye-movement replay supports episodic remembering. *Proceedings of the Royal Society B: Biological Sciences, 289*(1977), 20220964. https://doi.org/10.1098/rspb.2022.0964

Chen, Z. S., Zhang, X., Long, X., & Zhang, S. J. (2022). Are grid-like representations a component of all perception and cognition? *Frontiers in Neural Circuits, 16*, 924016. https://doi.org/10.3389/fncir.2022.924016

Constantinescu, A. O., et al. (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science, 352*(6292), 1464–1468. https://doi.org/10.1126/science.aaf0941

Convertino, L., Bush, D., Zheng, F., Adams, R. A., & Burgess, N. (2023). Reduced grid-like theta modulation in schizophrenia. *Brain, 146*(5), 2191–2198. https://doi.org/10.1093/brain/awac416

De Brouwer, A. J., Gallivan, J. P., Flanagan, J. R., & Bowman, N. (2022). Eye movements reveal spatiotemporal dynamics of visually-informed planning in navigation. *eLife, 11*, Article 73097. https://doi.org/10.7554/eLife.73097

Dewhurst, R., Nyström, M., Jarodzka, H., Foulsham, T., Johansson, R., & Holmqvist, K. (2012). It depends on how you look at it: Scanpath comparison in multiple dimensions with MultiMatch, a vector-based approach. *Behavior Research Methods, 44*(4), 1079–1100. https://doi.org/10.3758/s13428-012-0212-2

Doeller, C., Barry, C., & Burgess, N. (2010). Evidence for grid cells in a human memory network. *Nature, 463*, 657–661. https://doi.org/10.1038/nature08704

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience, 30*, 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Emrich-Mills, L. (2024). Developing a novel spatial memory task for early Alzheimer's diagnosis with immersive virtual reality and eye-tracking. (Doctoral thesis, University College London, UCL).

Foulsham, T., Barton, J. J. S., & Kingstone, A. (2018). Distinct roles of eye movements during memory encoding and retrieval. *Cognitive Psychology, 105*, 28–51. https://doi.org/10.1016/j.cogpsych.2018.08.001

Freiwald, W. A. (2020). The neural mechanisms of face processing: cells, areas, networks, and models. *Current Opinion in Neurobiology, 60*, 184–191. https://doi.org/10.1016/j.conb.2019.12.007

Graichen, L. P., Linder, M. S., Keuter, L., Jensen, O., Doeller, C. F., Lamm, C., Staudigl, T., & Wagner, I. C. (2024). Entorhinal grid-like codes for visual space during memory formation. *bioRxiv*. https://doi.org/10.1101/2024.09.27.615339

Hafting, T., Fyhn, M., Molden, S., Moser, M. B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature, 436*(7052), 801–806. https://doi.org/10.1038/nature03721

Hannula, D. E., Ryan, J. D., Tranel, D., & Cohen, N. J. (2017). Visual sampling predicts hippocampal activity. *The Journal of Neuroscience, 37*(3), 599–609. https://doi.org/10.1523/JNEUROSCI.4546-15.2017

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition, 33*, 98–106.

Heiting, G. (2020). Dominant eye test. *All About Vision*. https://www.allaboutvision.com/en-gb/resources/dominant-eye-test/

Howett, D., Castegnaro, A., Krzywicka, K., Hagman, J., Marchment, D., Henson, R. N. A., Rio, M., King, J. A., Burgess, N., & Chan, D. (2019). Differentiation of mild cognitive impairment using an entorhinal cortex-based test of VR navigation. *Brain, 142*, 1751–1766. https://doi.org/10.1093/brain/awz116

Hwang, E. J., Hauschild, M., Wilke, M., & Andersen, R. A. (2012). Spatial and temporal eye–hand coordination relies on the parietal reach region. *Journal of Neuroscience, 32*(38), 12281–12292. https://doi.org/10.1523/JNEUROSCI.1554-12.2012

Julian, J. B., Keinath, A. T., Frazzetta, G., & Epstein, R. A. (2017). Human entorhinal cortex represents visual space using a boundary-anchored grid. *Nature Neuroscience, 20*(2), 239–251. https://doi.org/10.1038/s41593-017-0049-1

Killian, N. J., Jutras, M. J., & Buffalo, E. A. (2012). A map of visual space in the primate entorhinal cortex. *Nature, 491*(7426), 761–764. https://doi.org/10.1038/nature11587

Kümmerer, M., Wallis, T. S. A., & Bethge, M. (2022). DeepGaze III: Modeling free-viewing human scanpaths with deep learning. *Journal of Vision, 22*(9), 1–24. https://doi.org/10.1167/jov.22.9.3

Ladyja-Wojcik, E., Mendez, M., & Dade, M. (2022). Eye movements and scene construction: Free viewing improves effective connectivity between the hippocampus and frontal eye fields (HPC-FEF). *Journal of Cognitive Neuroscience, 34*(4), 567–578.

Ladyka-Wojcik, N., Liu, Z. X., & Ryan, J. D. (2022). Unrestricted eye movements strengthen effective connectivity from hippocampal to oculomotor regions during scene construction. *NeuroImage, 260*, 119497. https://doi.org/10.1016/j.neuroimage.2022.119497

Laeng, B., & Teodorescu, D. S. (2018). Eye movements actively reinstate spatiotemporal mnemonic content. *Vision, 3*(2), 21. https://doi.org/10.3390/vision3020021

Mathôt, S. (2020). An adaptive algorithm for fast and reliable online saccade detection. *Behavior Research Methods, 52*(2), 376–394. https://doi.org/10.3758/s13428-019-01304-3

Nau, M., Navarro Schröder, T., Bellmund, J. L. S., & Doeller, C. F. (2018). Hexadirectional coding of visual space in human entorhinal cortex. *Nature Neuroscience, 21*, 188–190. https://doi.org/10.1038/s41593-017-0050-8

Olsen, R. K., Lee, Y., Kube, J., Rosenbaum, R. S., Grady, C. L., Moscovitch, M., & Ryan, J. D. (2015). The role of relational binding in item memory: Evidence from face recognition in a case of developmental amnesia. *Journal of Neuroscience, 35*(13), 5342–5350. https://doi.org/10.1523/JNEUROSCI.4566-14.2015

Roux, P., Passerieux, C., & Ramus, F. (2014). An eyetracking investigation of intentional motion perception in schizophrenia. *Journal of Psychiatry & Neuroscience, 40*(2), 118. https://www.jpn.ca/content/jpn/suppl/2021/09/26/40.2.118.DC1/jpn-140065-1-at.pdf

Schroeger, A., Goettker, A., Braun, D. I., & Gegenfurtner, K. R. (2020). Eye and hand movements when playing a dynamic computer game. *Journal of Vision, 20*(11), 1–13. https://doi.org/10.1167/jov.20.11.1

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences of the United States of America, 93*(24), 13515–13522. https://doi.org/10.1073/pnas.93.24.13515

SR Research. (2018). EyeLink® 1000 Plus User Manual (Version 1.5.2). https://www.hse.ru/mirror/pubs/share/560338728.pdf

Staudigl, T., Hartley, T., & Burgess, N. (2018). Hexadirectional modulation of high-frequency electrophysiological activity in the human anterior medial temporal lobe maps visual space. *Current Biology, 28*(20), 3325–3329. https://doi.org/10.1016/j.cub.2018.07.042

Wynn, J. S., Shen, K., & Ryan, J. D. (2019). Eye movements actively reinstate spatiotemporal mnemonic content. *Vision, 3*(2), 21. https://doi.org/10.3390/vision3020021

SUPPLEMENTARY MATERIALS

Table 1. ***The similarity scores for each participant and each dimension.***

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10	Trial 11	Trial 12	Trial 13	Trial 14	Trial 15	
P1	0,808	0,837	0,826	0,846	0,847	0,874	0,852	0,955	0,870	0,932	0,918	0,854	0,823	0,849	0,930	**Shape**
P1	0,578	0,546	0,772	0,714	0,742	0,870	0,690	0,891	0,829	0,857	0,913	0,698	0,685	0,752	0,872	**Direction**
P1	0,815	0,927	0,821	0,845	0,870	0,894	0,884	0,956	0,826	0,942	0,958	0,877	0,822	0,916	0,952	**Length**
P1	0,739	0,814	0,790	0,819	0,746	0,829	0,793	0,920	0,849	0,824	0,921	0,833	0,725	0,800	0,934	**Position**
P1	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	**Duration**
P2	0,898	0,920	0,863	0,863	0,904	0,833	0,856	0,791	0,890	0,939	0,871	0,837	0,925	0,926	0,843	**Shape**
P2	0,857	0,827	0,623	0,855	0,772	0,592	0,606	0,711	0,835	0,921	0,764	0,770	0,937	0,857	0,821	**Direction**
P2	0,884	0,937	0,864	0,879	0,893	0,888	0,908	0,913	0,893	0,934	0,878	0,856	0,916	0,931	0,760	**Length**
P2	0,814	0,884	0,734	0,746	0,740	0,732	0,731	0,612	0,781	0,925	0,818	0,688	0,850	0,897	0,738	**Position**
P2	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,633	1,000	1,000	1,000	**Duration**
P3	0,858	0,883	0,886	0,828	0,744	0,727	0,879	0,723	0,914	0,942	0,918	0,810	0,846	0,841	0,956	**Shape**
P3	0,624	0,786	0,918	0,657	0,714	0,798	0,682	0,594	0,859	0,928	0,833	0,893	0,607	0,857	0,870	**Direction**
P3	0,853	0,906	0,893	0,869	0,828	0,595	0,910	0,628	0,932	0,887	0,943	0,752	0,950	0,937	0,976	**Length**
P3	0,822	0,839	0,922	0,826	0,687	0,728	0,787	0,732	0,728	0,944	0,902	0,767	0,811	0,890	0,938	**Position**
P3	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	**Duration**
P4	0,829	0,905	0,619	0,915	0,869	0,896	0,883	0,846	0,843	0,911	0,862	0,900	0,931	0,855	0,952	**Shape**
P4	0,755	0,818	0,515	0,582	0,726	0,785	0,648	0,876	0,745	0,909	0,661	0,762	0,864	0,915	0,877	**Direction**
P4	0,884	0,953	0,493	0,923	0,932	0,933	0,918	0,931	0,849	0,944	0,948	0,902	0,926	0,882	0,941	**Length**
P4	0,759	0,919	0,602	0,939	0,850	0,901	0,774	0,790	0,601	0,932	0,918	0,914	0,908	0,743	0,927	**Position**
P4	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,550	1,000	1,000	1,000	1,000	1,000	1,000	**Duration**
P5	0,946	0,894	0,974	0,739	0,690	0,933	0,891	0,787	0,887	0,894	0,840	0,871	0,953	0,888	0,953	**Shape**
P5	0,937	0,815	0,897	0,872	0,756	0,941	0,964	0,546	0,833	0,696	0,362	0,704	0,953	0,700	0,874	**Direction**
P5	0,930	0,883	0,972	0,643	0,517	0,935	0,955	0,923	0,934	0,935	0,888	0,929	0,911	0,904	0,951	**Length**
P5	0,941	0,808	0,948	0,759	0,782	0,928	0,935	0,856	0,690	0,928	0,795	0,897	0,935	0,913	0,957	**Position**
P5	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	**Duration**
P6	0,909	0,949	0,637	0,961	0,961	0,885	0,860	0,938	0,925	0,888	0,960	0,829	0,853	0,961	0,938	**Shape**

P6	0,942	0,939	0,774	0,974	0,936	0,936	0,618	0,915	0,813	0,674	0,970	0,884	0,744	0,908	0,911	**Direction**
P6	0,924	0,899	0,475	0,941	0,937	0,865	0,911	0,927	0,900	0,937	0,951	0,927	0,927	0,957	0,901	**Length**
P6	0,869	0,957	0,923	0,920	0,940	0,914	0,821	0,822	0,855	0,838	0,954	0,965	0,839	0,959	0,915	**Position**
P6	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	**Duration**
P7	0,928	0,935	0,950	0,925	0,884	0,871	0,941	0,751	0,922	0,983	0,819	0,845	0,976	0,915	0,814	**Shape**
P7	0,964	0,873	0,749	0,900	0,473	0,700	0,893	0,617	0,809	0,976	0,315	0,882	0,971	0,826	0,657	**Direction**
P7	0,869	0,945	0,944	0,950	0,959	0,880	0,894	0,761	0,954	0,970	0,855	0,839	0,975	0,888	0,789	**Length**
P7	0,903	0,924	0,885	0,906	0,666	0,795	0,927	0,700	0,908	0,980	0,936	0,955	0,967	0,796	0,828	**Position**
P7	1,000	1,000	1,000	1,000	0,022	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	**Duration**
P8	0,821	0,785	0,869	0,774	0,806	0,864	0,788	0,967	0,908	0,784	0,848	0,970	0,884	0,808	0,934	**Shape**
P8	0,688	0,580	0,846	0,581	0,450	0,746	0,834	0,969	0,874	0,505	0,606	0,977	0,868	0,430	0,947	**Direction**
P8	0,839	0,791	0,819	0,823	0,896	0,867	0,769	0,944	0,829	0,824	0,865	0,942	0,952	0,919	0,976	**Length**
P8	0,832	0,757	0,728	0,579	0,635	0,812	0,811	0,968	0,828	0,889	0,740	0,973	0,867	0,520	0,948	**Position**
P8	1,000	1,000	0,906	1,000	0,856	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	**Duration**
P9	0,966	0,821	0,858	0,915	0,923	0,860	0,861	0,835	0,861	0,887	0,952	0,947	0,973	0,807	0,926	**Shape**
P9	0,942	0,233	0,726	0,962	0,834	0,751	0,743	0,616	0,705	0,770	0,937	0,955	0,979	0,508	0,876	**Direction**
P9	0,974	0,934	0,893	0,909	0,926	0,903	0,936	0,945	0,929	0,929	0,973	0,936	0,967	0,920	0,931	**Length**
P9	0,955	0,754	0,761	0,907	0,916	0,708	0,958	0,731	0,739	0,854	0,929	0,911	0,938	0,787	0,940	**Position**
P9	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	**Duration**
P10	0,961	0,938	0,919	0,850	0,839	0,940	0,934	0,873	0,830	0,943	0,906	0,903	0,969	0,870	0,948	**Shape**
P10	0,948	0,903	0,706	0,666	0,898	0,851	0,934	0,731	0,837	0,955	0,845	0,898	0,941	0,783	0,906	**Direction**
P10	0,950	0,932	0,918	0,833	0,876	0,951	0,898	0,930	0,819	0,959	0,916	0,909	0,974	0,950	0,949	**Length**
P10	0,938	0,929	0,844	0,739	0,647	0,940	0,922	0,844	0,791	0,943	0,962	0,906	0,964	0,894	0,961	**Position**
P10	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	**Duration**
P11	0,935	0,855	0,947	0,971	0,975	0,899	0,969	0,944	0,835	0,799	0,913	0,855	0,793	0,932	0,908	**Shape**
P11	0,857	0,770	0,973	0,947	0,973	0,976	0,984	0,916	0,499	0,409	0,843	0,534	0,506	0,823	0,886	**Direction**
P11	0,920	0,941	0,938	0,958	0,961	0,934	0,948	0,908	0,883	0,646	0,874	0,834	0,898	0,941	0,910	**Length**
P11	0,850	0,714	0,945	0,952	0,961	0,915	0,932	0,893	0,631	0,792	0,923	0,722	0,703	0,830	0,842	**Position**
P11	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,011	1,000	1,000	1,000	**Duration**
P12	0,858	0,898	0,897	0,891	0,891	0,890	0,849	0,799	0,825	0,899	0,833	0,866	0,864	0,807	0,888	**Shape**
P12	0,612	0,736	0,703	0,735	0,663	0,801	0,781	0,636	0,686	0,830	0,674	0,801	0,753	0,643	0,925	**Direction**

P12	0,849	0,916	0,948	0,892	0,958	0,895	0,774	0,684	0,769	0,886	0,907	0,859	0,847	0,733	0,870	**Length**
P12	0,815	0,725	0,782	0,672	0,736	0,775	0,803	0,619	0,781	0,770	0,748	0,721	0,796	0,781	0,829	**Position**
P12	1,000	1,000	0,822	1,000	0,481	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	**Duration**
P13	0,858	0,875	0,895	0,834	0,866	0,840	0,843	0,835	0,833	0,902	0,889	0,874	0,832	0,816	0,856	**Shape**
P13	0,522	0,616	0,722	0,564	0,605	0,792	0,780	0,635	0,725	0,836	0,807	0,808	0,499	0,649	0,808	**Direction**
P13	0,848	0,891	0,908	0,874	0,882	0,807	0,771	0,857	0,780	0,906	0,856	0,849	0,926	0,746	0,915	**Length**
P13	0,754	0,777	0,725	0,663	0,708	0,772	0,767	0,648	0,763	0,764	0,706	0,738	0,772	0,779	0,827	**Position**
P13	0,430	0,668	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,718	**Duration**
P14	0,840	0,888	0,896	0,896	0,852	0,879	0,913	0,899	0,916	0,885	0,852	0,878	0,921	0,887	0,888	**Shape**
P14	0,753	0,618	0,803	0,815	0,783	0,721	0,860	0,786	0,919	0,646	0,602	0,688	0,793	0,686	0,759	**Direction**
P14	0,904	0,912	0,937	0,922	0,902	0,943	0,902	0,886	0,913	0,904	0,958	0,920	0,951	0,951	0,903	**Length**
P14	0,793	0,740	0,708	0,762	0,661	0,765	0,853	0,733	0,802	0,813	0,563	0,706	0,869	0,560	0,769	**Position**
P14	0,988	1,000	0,950	1,000	1,000	1,000	0,888	1,000	1,000	1,000	0,030	0,432	1,000	0,997	0,833	**Duration**
P15	0,865	0,882	0,909	0,916	0,902	0,902	0,876	0,828	0,917	0,915	0,861	0,913	0,874	0,868	0,818	**Shape**
P15	0,766	0,843	0,754	0,841	0,799	0,792	0,736	0,630	0,795	0,731	0,757	0,773	0,709	0,791	0,754	**Direction**
P15	0,840	0,883	0,903	0,936	0,907	0,919	0,910	0,864	0,920	0,956	0,924	0,937	0,935	0,900	0,828	**Length**
P15	0,574	0,794	0,783	0,796	0,762	0,834	0,763	0,653	0,726	0,860	0,598	0,792	0,725	0,806	0,753	**Position**
P15	1,000	1,000	1,000	1,000	1,000	1,000	0,982	0,897	1,000	1,000	0,917	0,987	0,840	0,503	0,935	**Duration**
P16	0,828	0,871	0,877	0,874	0,819	0,889	0,874	0,848	0,872	0,865	0,891	0,860	0,868	0,817	0,902	**Shape**
P16	0,642	0,756	0,831	0,555	0,780	0,653	0,774	0,704	0,681	0,630	0,879	0,782	0,536	0,592	0,851	**Direction**
P16	0,777	0,857	0,859	0,930	0,884	0,912	0,932	0,870	0,943	0,918	0,856	0,847	0,940	0,905	0,924	**Length**
P16	0,604	0,717	0,797	0,628	0,724	0,764	0,694	0,756	0,747	0,786	0,829	0,855	0,811	0,814	0,818	**Position**
P16	1,000	1,000	1,000	1,000	0,636	1,000	1,000	1,000	1,000	0,777	1,000	1,000	1,000	1,000	1,000	**Duration**

Object-Location Memory Experiment Procedure.
***Step-by-Step Process*:**

1. **Pre-Experiment Setup:**
    - Participants are informed about the task and time constraints in the instructions before the experiment begins.
    - A calibration session takes place.
    - An untimed practice session is conducted to allow participants to familiarize themselves with the task, ask questions, and identify any calibration issues with the equipment.
2. **Start of the Experiment:**
    - The main experiment consists of 5 testing blocks, each containing three encoding-retrieval pairs (3 encoding trials followed by 3 retrieval trials).
3. **Encoding Phase (3 trials):**
    - In the encoding phase, as demonstrated in Figure S1, participants are looking at the dark screen where a randomly generated arrangement of objects is hidden. Their vision is constrained to a small circular area of focus, so only the objects in the area of focus are visible, meaning participants can only see objects when they look at them directly.
    - Participants are instructed to locate all 5 hidden objects and remember their location.
4. **Retrieval Phase (3 trials):**
    - In each block, after 3 encoding trials, participants are shown the same object arrangements in the 3 trials of the retrieval phase.
    - In this phase, one object is missing, and participants need to identify where the missing object was located in the original arrangement and focus on its location.
    - The foveated vision constraint is still active, requiring participants to focus on specific areas of the screen to find the missing object.
    - If a participant focuses on the same area for 5 seconds continuously, the program automatically proceeds to the next trial, indicating they have likely located the missing object.
5. **Trial Completion:**
    - Each encoding or retrieval trial has a 30-second time limit.
    - If the participant has not identified the missing object by the end of the 30 seconds, the trial will automatically move to the next one.
6. **Block Completion:**
    - Each testing block consists of three encoding-retrieval pairs. Once the three encoding and three retrieval trials are completed, participants can take a short break.
    - Before starting a new block, participants are reminded of the task instructions.

7. **Completion of the Experiment:**
    - After completing all blocks, the experiment ends.
    - Participants are given a final opportunity to ask questions and receive debriefing.

***Figure S1**: **Visual Setup of the Experimental Trial.** The figure shows a screenshot of the Object-Location Memory Experiment during the encoding or retrieval phase. The image illustrates an object within a circular area of focus, simulating the foveated vision condition. Only the object within this area is visible, while the surrounding space is obscured. This demonstrates how the participant's gaze is tracked, with the circle following eye movements to reveal the objects in focus.*