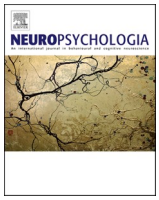
Neuropsychologia 204 (2024) 109000

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia



Using imagination and the contents of memory to create new scene and object representations: A functional MRI study

Qun Ye [a,b], Celia Fidalgo [c], Patrick Byrne [c], Luis Eduardo Muñoz [c], Jonathan S. Cant [c,1,*], Andy C.H. Lee [c,d,1,*]

[a] Intelligent Laboratory of Child and Adolescent Mental Health and Crisis Intervention of Zhejiang Province, School of Psychology, Zhejiang Normal University, Jinhua, 321004, Zhejiang, China
[b] Key Laboratory of Intelligent Education Technology and Application of Zhejiang Province, Zhejiang Normal University, Jinhua, 321004, Zhejiang, China
[c] Department of Psychology (Scarborough), University of Toronto, Toronto, Ontario, M1C 1A4, Canada
[d] Rotman Research Institute, Baycrest Centre, Toronto, Ontario, M6A 2E1, Canada

ARTICLE INFO

Keywords:
Imagination
Memory retrieval
Scene processing
Object processing
ROI analysis
Whole-volume analysis
Multivariate classification

ABSTRACT

Humans can use the contents of memory to construct scenarios and events that they have not encountered before, a process colloquially known as imagination. Much of our current understanding of the neural mechanisms mediating imagination is limited by paradigms that rely on participants' subjective reports of imagined content. Here, we used a novel behavioral paradigm that was designed to systematically evaluate the contents of an individual's imagination. Participants first learned the layout of four distinct rooms containing five wall segments with differing geometrical characteristics, each associated with a unique object. During functional MRI, participants were then shown two different wall segments or objects on each trial and asked to first, retrieve the associated objects or walls, respectively (retrieval phase) and then second, imagine the two objects side-by-side or combine the two wall segments (imagination phase). Importantly, the contents of each participant's imagination were interrogated by having them make a same/different judgment about the properties of the imagined objects or scenes. Using univariate and multivariate analyses, we observed widespread activity across occipito-temporal cortex for the retrieval of objects and for the imaginative creation of scenes. Interestingly, a classifier, whether trained on the imagination or retrieval data, was able to successfully differentiate the neural patterns associated with the imagination of scenes from that of objects. Our results reveal neural differences in the cued retrieval of object and scene memoranda, demonstrate that different representations underlie the creation and/or imagination of scene and object content, and highlight a novel behavioral paradigm that can be used to systematically evaluate the contents of an individual's imagination.

1. Introduction

Imagination is a highly complex cognitive function that enables us to mentally simulate experiences with objects, people, and scenes without relying on incoming sensory information. There are numerous benefits to imagination, such as the development of various cognitive faculties (e.g., children engaging in imaginative play), and optimizing decision-making processes in future real-world scenarios (e.g., visualizing upcoming landmarks and turns while navigating a spatial route). Despite the importance of imagination to the functioning of an individual, we do not currently have a deep understanding of how different cognitive processes, including memory retrieval and the manipulation of existing mental representations, interact to construct imagined events.

Neuroimaging research has demonstrated that a network of brain regions implicated in imagination overlaps with areas mediating other cognitive functions such as episodic memory retrieval, spatial navigation, and future thinking (for reviews, see Mullally and Maguire, 2014; Schacter et al., 2012). Indeed, substantial functional magnetic resonance imaging (fMRI) work has confirmed the engagement of the medial temporal lobes (MTL), and in particular the hippocampus, in not only the perception of scenes (for reviews, see Epstein and Baker, 2019; Lee et al., 2012a,b), but also in the use of imagination to construct fictitious

* Corresponding authors. Department of Psychology (Scarborough), University of Toronto, Toronto, Ontario M1C 1A4, Canada.
E-mail addresses: jonathan.cant@utoronto.ca (J.S. Cant), andych.lee@utoronto.ca (A.C.H. Lee).
[1] Co-senior Authors.

https://doi.org/10.1016/j.neuropsychologia.2024.109000
Received 20 March 2024; Received in revised form 9 September 2024; Accepted 10 September 2024
Available online 11 September 2024

experiences (Hassabis et al., 2007a), future scenes, and events (Addis et al., 2007; Staresina et al., 2011). Consistent with these findings, amnesic patients with bilateral hippocampal damage have been reported to be significantly impaired at recalling past and imagining new experiences (Hassabis et al., 2007b).

In addition to regions of the MTL, extrastriate visual regions important for object perception (i.e., lateral occipital area, LO) (Malach et al., 1995), scene perception (i.e., occipital place area, OPA, and parahippocampal place area, PPA) (Dilks et al., 2013; Epstein and Kanwisher, 1998) and face perception (i.e., fusiform face area, FFA) (Kanwisher et al., 1997), have been implicated in imaginative processing (Bainbridge et al., 2021). For instance, multi-voxel patterns of fMRI activity in the FFA and PPA were used to successfully distinguish between the imagination of faces and places, respectively (Ragni et al., 2021). Moreover, activity in LO and OPA showed high sensitivity to objects and scenes, respectively, during visual imagery (Bainbridge et al., 2021). In addition to this, one recent neuropsychological study showed that patients with focal parieto-occipital cortical atrophy displayed impaired imaginative scene construction performance, at a level comparable to that observed in an amnestic Alzheimer's Disease (AD) group (Ramanan et al., 2018), thus confirming the role of these extrastriate visual regions in imagination.

Although it is evident that several extrastriate and MTL areas contribute to imaginative processing, their precise contributions are unclear. The constructive episodic simulation hypothesis (Schacter and Addis, 2007) proposes that a core network encompassing a number of these and other (e.g., frontoparietal) regions supports both episodic memory retrieval and the imagination of future and hypothetical events. Specifically, it is hypothesized that different regions subserve distinct underlying processes such as the retrieval and binding of past episodic

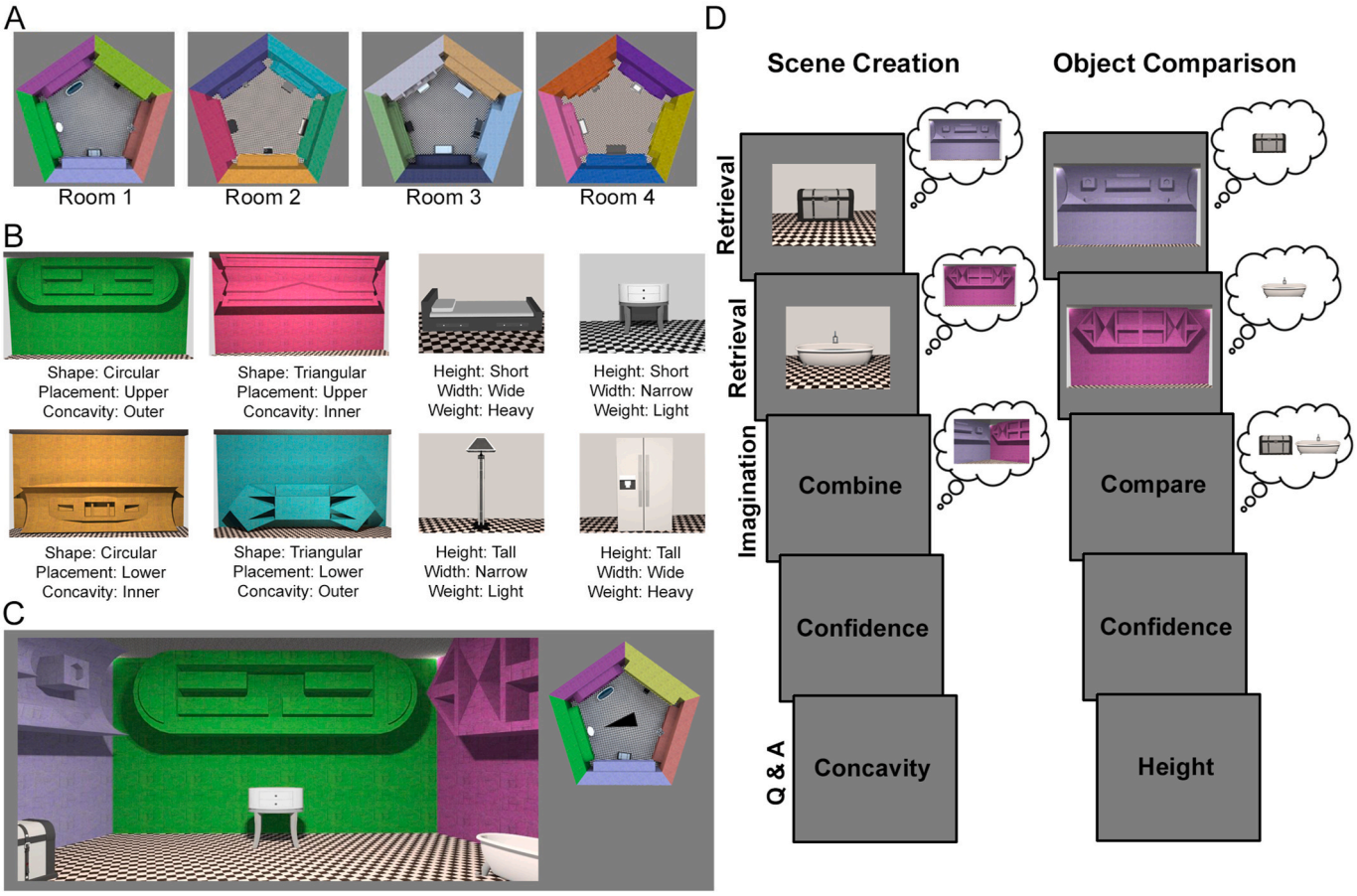


Fig. 1. Stimuli and study design. **A**. Four virtual rooms with unique wall protrusions and objects. **B**. Examples of stimulus characteristics for wall protrusions and objects. Wall protrusions all contained a shape (circular or triangular; designated based on the outer most edges of the protrusion), a placement (upper or lower; designated based on the segment of the wall the protrusion was located on), and a concavity (inner or outer; designated based on whether the protrusion caved inward at the edges or protruded outward at the edges). Objects all contained a height (tall or short; designated based on whether the object was taller than the midpoint of the wall it stood in front of), a width (wide or narrow; designated based on whether the object would span more than or less than 1 m in the real world), and a weight (heavy or light; designated based on whether an average person could reasonably lift the object off the ground and hold it comfortably for 20 s). **C**. The main viewing window along with the birds-eye map. The main viewing window through which participants viewed the detailed virtual environment, and the birds-eye map provided an overview of the room as well as an indicator of where participants were currently viewing (i.e., the black arrow) relative to the rest of the room. **D**. Schematic for task inside the scanner. Each trial began with a single image, onscreen for 8000 ms, to which participants recalled its associate. Specifically, participants either recalled the wall associated with an object in the Scene Creation condition, or conversely, the object associated with a wall in the Object Comparison condition. After a 1500 ms jittered ISI, participants saw a second image for 8000 ms, to which they recalled a second associate. After another 1500 ms jittered ISI, participants saw either a "Combine" or "Compare" cue for 7000 ms. For Scene Creation trials, participants imagined the two walls they had just retrieved in the retrieval phase and joined them together to create the corner of a room. For Object Comparison trials, participants imagined the two objects they had just retrieved side-by-side without touching, in the absence of any spatial features. After the 7000 ms imagination phase, a 1500 ms jittered ISI followed and participants were then given 3000 ms to rate their confidence in their dual-stimulus mental image on a scale from 1 to 4. They then had 5000 ms to indicate whether the two walls/objects they had imagined were the same or different in relation to a specified characteristic (the correct answers are "same" for the example Scene Creation trial with respect to the concavity of the wall protrusions and "different" for the example Object Comparison trial with respect to the height of the objects). Both examples shown here represent the Same Room Detached condition (see 'Test Session (during scanning)' for more details).

details for novel scenarios in the case of the hippocampus, and the processing of episodic spatial details by the parahippocampal cortex (Schacter et al., 2017). In contrast, since many of these regions are also implicated in spatial cognition and navigation, an alternative view, the scene-construction hypothesis (Hassabis and Maguire, 2009), posits that these regions support the construction of spatial models that are central to experienced and imagined events. Notably, while several studies have provided evidence in support of each of these viewpoints (e.g., Benoit and Schacter, 2015; Szpunar et al., 2014; Zeidman et al., 2015), we contend that the majority of the experimental paradigms that have been used provide limited insight into the complexities of the cognitive processes that underlie the imaginative process. Specifically, imagination tasks have typically required participants to form a mental image in response to one or two cue words on each trial and to subsequently provide subjective judgments of the vividness of their imaginative content, and/or rate the confidence or difficulty experienced in creating their mental image. While the aforementioned results have been illuminating, this approach has not allowed the content of a participant's imagination to be systematically evaluated. Moreover, it has provided little insight into the extent to which participants may be drawing on elements from one or more existing episodic (or conceptual) representations, how these elements are flexibly manipulated to create a novel event, and to what extent individual brain regions contribute to each of these processes.

With the above in mind, in the current neuroimaging study we designed a novel imagination task to provide deeper insight into the processes and associated neural correlates that underlie imagined content, with a particular focus on the MTL and extrastriate visual regions. Since scenes and objects dominate the sensory world around us and are fundamental components of our memories, we focused on the imagination of scene and object elements, consistent with previous studies that have examined the imagination of one or both of these stimulus categories (e.g., Zeidman et al., 2015), and with neuroimaging studies implicating object- and scene-sensitive regions in imaginative processing (e.g., Bainbridge et al., 2021) and episodic memory (e.g., Schacter et al., 2017). Importantly, we developed a multi-stage paradigm (see Fig. 1) that allowed us to manipulate the retrieval of learned scene and object elements and to dissociate the retrieval of these elements from their manipulation to create novel representations. In addition, we created an empirical metric to verify the contents of each participant's imagination, rectifying the methodological limitation discussed above. We investigated three separate questions. First, do extrastriate and MTL regions contribute to the retrieval and manipulation of learned elements during the imaginative process? Based on the literature reviewed previously (e.g., Bainbridge et al., 2021; Zeidman et al., 2015), we predicted that both regions would be implicated. Second, how, and to what extent, do the contributions of these regions differ for retrieval compared to imagination, as well as the processing of objects compared to scenes? While we were unclear on the exact nature of any potential differences in retrieval versus imagination, we predicted that lateral regions of occipitotemporal cortex (e.g., LO) would be more heavily implicated in object processing, whereas more medial regions (e.g., PPA) would be more heavily implicated in scene processing (both retrieval and imagination). With respect to MTL regions, we expected to find sensitivity to both object and scene processing in the entorhinal cortex and hippocampus, since recent studies have demonstrated functional connectivity between these MTL structures and cortical scene and object processing pathways (Grande et al., 2022). Finally, are there behavioral and neural differences observed when a new mental representation is created based on learned elements from a single representation (e.g., creating novel recombinations of the scene and object elements that were previously experienced in one spatial environment) as opposed to multiple pre-existing representations (e.g., creating novel recombinations of the scene and object elements that were previously experienced in different spatial environments)? Here we expected that combining elements across different object and scene representations would be more difficult than combining elements from within the same representation, and thus expected the behavioral and neuronal data to reflect this difference. We used both univariate analyses and multivariate analyses to address these questions because, individually, these approaches could reveal different aspects of neuronal processing (Epstein and Morgan, 2012). Hence, combining these approaches in the same study increases our ability to further our understanding of the imaginative processing of scene and object representations.

2. Methods

2.1. Participants

A total of 20 young adults were recruited from the University of Toronto and York University communities, with each participant receiving monetary compensation (CAD 85) for their participation. Data from one participant were excluded due to poor behavioral performance during scanning. Specifically, this participant performed below chance (50%) on the 'Question and Answer' phase of the behavioral paradigm, which was designed to assess the accuracy of participants' imagined content (see below). The remaining 19 participants had a mean age of 24.47 years (range 21–33, 12 females). All participants provided informed written consent prior to taking part and were screened for a history of psychological illness, traumatic brain injury, current use of neuroleptic medications, and MRI contraindications. Participants all had normal or corrected-to-normal vision. This study was approved by the York University (#2016-291) and University of Toronto Ethics Review Boards (#27455).

2.2. Stimuli

Four virtual rooms were constructed, which participants freely viewed and encoded (see Procedure: Study Sessions 1 and 2). Rooms were constructed using SketchUp (https://www.sketchup.com/) and rendered using iRender X. Each room had five walls that were each connected at 108° angles (i.e., pentagons from a birds-eye-view). Across all four rooms, each wall had a unique color, a geometric protrusion, and a unique object in front of it (Fig. 1A–C). Each geometric protrusion had three elements that were later used to test recall: (1) shape (circular or triangular), (2) placement (upper or lower half of the wall), and (3) concavity (concave or convex). All 20 objects were rendered in grayscale and were moveable household objects (e.g., lamp, mirror, fridge, desk, etc.). Each object had three elements that were later used to test recall: (1) height (tall or short), (2) width (wide or narrow), and (3) weight (heavy or light). These three elements were manipulated independently, such that not all wide objects were also short, and not all narrow objects were also tall, and so forth.

2.3. Procedure

The experiment was run using MATLAB (MathWorks, Natick, MA) with the Psychophysics Toolbox (Brainard, 1997). There were four separate experimental sessions, with a 24- to 64-h window between each (e.g., typically a Monday-Wednesday-Thursday-Friday schedule). The four sessions included two study sessions, one pre-test session outside the scanner, and one test session during scanning. The study and pre-test sessions were conducted on a laptop with a 12.5" LCD screen (1366 × 768 pixel resolution).

2.3.1. Study sessions 1 and 2

Study sessions 1 and 2 were identical, with the latter taking place 24–48 h after session 1. In each session, participants were informed that they would freely view and encode a series of four virtual rooms for 10 min each (Fig. 1A–C), with interspersed memory tests. They were asked to remember the visual details of the wall protrusions and the objects in each room and their memory for these was tested both onscreen with

forced-choice tests and via verbal recall. The main viewing window for each room was 1200 × 562 pixels in size. Participants were also shown a birds-eye map presented in the top right corner of the screen (200 × 193 pixels). The birds-eye map contained a birds-eye-view of the current room, as well as an arrow that indicated which direction the viewer was facing in the main viewing window. Participants used the right and left arrow keys to move the viewpoint arrow clockwise or counter-clockwise, which concurrently adjusted their viewpoint in the main viewing window by altering their viewpoint in a circular manner (i.e., adjusting viewpoint right-wards or left-wards). Each 10-min study period was divided into four smaller segments by onscreen multiple choice tests, leading to the following schedule that was identical across all participants: (1) 4 min of free viewing, followed by the first multiple choice test (wall-object relationship test); then (2) 2 min of free viewing, followed by the second multiple choice test (wall-to-wall relationship test); then (3) 2 min of free viewing, followed by the third multiple choice test (wall recall); then (4) 2 min of free viewing, and the last multiple choice test (object recall); and finally (5) a verbal recall test.

Wall-object relationship test: During this test, participants were shown a single target image in the centre of the screen (i.e., either a wall without its associated object [777-836 × 472 pixels] or object without its associated wall [433 × 344 pixels] from the current room) as well as five images along the bottom of the screen, which represented five choices mapped onto keys 1 through 5. If the target image was an object, then the five images presented along the bottom of the screen were walls, and if the target image was a wall, the five images were objects. An onscreen prompt was presented underneath the target image. The prompt directed participants to indicate which of the five images had appeared alongside the target in the room they had been viewing. The target image and prompt remained onscreen until a response was made. Once a response was made, onscreen feedback was immediately presented ("Correct" or "Incorrect") for 2000 ms, after which the next target image appeared. All five walls and five objects were presented as a target once, in randomized order, for a total of ten trials. After the last trial, participants resumed free viewing of the room for 2 min. This test prepared participants for later recall tests involving relationships between walls and objects.

Wall-to-wall relationship test: Participants were shown a single target image of a wall (777-836 × 472 pixels) in the center of the screen as well as four walls along the bottom of the screen (all without their associated objects), which represented choices mapped onto keys 1 through 4. An onscreen prompt directed participants to choose which of the four walls had appeared to the left of the target wall, or to the right of the target wall. Upon response, feedback appeared and remained onscreen for 2000 ms, after which the next target appeared. All five walls appeared as a target twice, once with a "left" prompt and once with a "right" prompt, for a total of ten trials. The walls appeared in a randomized order and the prompt remained onscreen until a response was made. After the last trial, participants resumed free viewing of the room for 2 min. This test ensured participants were binding the walls within the room together and not simply remembering each wall and object separately.

Wall recall test: Participants were shown a single target image of an object (without its associated wall) (433 × 344 pixels) in the center of the screen and were instructed to vividly recall the wall that had been behind that object in as much detail as possible. A single prompt appeared underneath the image: either "Shape," "Placement," or "Concavity." Participants saw all three prompts, one-at-a-time, in random order. Participants responded to each prompt with one of two keys on the keyboard. For "Shape," either the "T" or "C" key for triangular or circular. For "Placement," either "U" or "L" for upper or lower. Lastly for "Concavity," either "I" or "O" for inward (concave) and outward (convex). The target image and prompt remained onscreen until a response was made. Upon response, feedback was displayed on screen for 2000 ms, after which the next trial appeared. After participants responded to all three prompts for a given object, the next object appeared until all five objects in the room had been displayed, for a total of 15 responses. After the last trial, participants resumed free viewing of the room for 2 min. This test ensured participants accurately recalled the associated wall when they viewed an object cue.

Object recall test: Participants were shown a single target image of a wall (without its associated object) (777-836 × 472 pixels) in the centre of the screen and were instructed to recall and vividly imagine the object that had been in front of that wall in as much detail as possible. A single prompt appeared underneath the image: either "Height," "Width," or "Weight." Participants saw all three prompts, one at a time, in random order. Participants responded to each prompt with one of two keys on the keyboard. For "Height," either the "T" or "S" key for tall or short. For "Width," either "W" or "N" for wide or narrow. Lastly for "Weight," either "H" or "L" for heavy or light. The target image and prompt remained onscreen until a response was made. Upon response, feedback was displayed on screen for 2000 ms, after which the next trial appeared. After participants responded to all three prompts for a given wall, the next wall appeared until all five walls in the room had been displayed, for a total of 15 responses. After the last trial, participants resumed free viewing of the room for 2 min, after which they were finished studying that room. This test ensured participants accurately recalled the associated object when they viewed a cued wall.

Verbal recall test: After completing 10 min of free viewing and all multiple-choice tests, participants were instructed to verbally recall the room they had just viewed to the experimenter, who took notes of their response. They were asked to provide as many details as possible about the objects and the walls and were encouraged to describe the walls and objects in order (i.e., starting at any given wall-object combination and describing walls and objects consecutively, clockwise or counterclockwise). After the verbal recall was complete, they were given a 5-min break and proceeded to the next room. Room order was counterbalanced across participants. Performance across these four tests was used as a screening measure such that those who performed below 75% across all four tests did not proceed to Study Session 2 and were excused from the study. All scanned participants met this criterion and attained at least 75% accuracy in Study Session 2.

2.3.2. Pre-test session (outside scanner)

The third session took place between 24 and 48 h after Study Session 2 and 8–16 h prior to the scan. Participants completed a short recall test for all the walls and objects they had viewed in the virtual environments. Participants viewed a single target image onscreen and were asked to indicate the three critical characteristics of its associated image. That is, if an object was presented (433 × 344 pixels), participants recalled the wall that had been behind it and responded to each of three prompts pertaining to Shape, Placement, and Concavity. If a wall was presented (777–836 × 472 pixels), participants recalled the object that had been in front of it and responded to Height, Width, and Weight prompts. Each prompt appeared one-at-a-time underneath the target image with response options underneath (see Fig. 1B for all possible response options). The target image and prompt remained onscreen until a response was made. Participants responded by clicking on the perceived correct option using their mouse. Once they responded, feedback ("Correct" or "Incorrect") appeared onscreen for 2000 ms, and the next prompt appeared. After all three prompts were displayed, the next target image was shown. Participants viewed all 20 wall images and all 20 object images in random order and responded to all three prompts for each image, for a total of 120 responses.

2.3.3. Test session (during scanning)

Prior to entering the scanner, participants were given a brief reminder of the rooms. That is, they were given 2 min to freely view each of the four rooms, for a total of 8 min. Then they were given instructions for the task to be conducted in the scanner and completed four practice trials outside the scanner and four practice trials inside the scanner. For the test inside the scanner (Fig. 1D), each trial began with a retrieval phase, in which a single object (433 × 344 pixels) (no background wall

present) or wall (777-836 × 472 pixels) (no foreground object present) was first presented for 8000 ms. Participants were asked to recall the associated wall (Scene Creation trials) or object (Object Comparison trials), respectively, and once they had a mental image of the target in mind, they were instructed to press 1 on the button box. A jittered ISI of 1500 ms then ensued, followed by the presentation of a second image of the same stimulus category (i.e., a second object or a second wall) for 8000 ms. Participants were asked to imagine the associated wall or object in the absence of the first, and to press 1 on the button box once they had it in mind. This was followed by a jittered 1500 ms ISI, which led to the start of an imagination phase. For Scene Creation trials, the cue word "Combine" was presented for 7000 ms, to which participants were required to join together in their mind's eye the two walls they had previously imagined independently (without their associated objects) to create the corner of a room. They were instructed to always imagine the first wall they had viewed on the left and the second wall on the right. On one third of Scene Creation trials (Same Attached trials), the two walls had been previously attached in one of the virtual rooms, and their combination was, therefore, akin to retrieving a previously viewed room corner. On another third of Scene Creation trials (Same Detached trials), the two walls had not been previously attached in one of the virtual environments, but were part of the same room (e.g., Fig. 1D left). The mental image created by participants was thus a constructed scene that had never been viewed before, but was composed of two walls that had been viewed in relatively close physical proximity in one of the virtual rooms. The last third of Scene Creation trials (Different trials) involved two walls that came from different virtual rooms. Thus, the mental image created by participants was a constructed scene that had never been viewed before, and was made up of two walls from two separate mental scene representations.

In contrast, for Object Comparison trials, the onscreen verbal cue was "Compare" (7000 ms) and participants were asked to imagine the two objects sitting side-by-side without touching each other. They were instructed to imagine the objects in the absence of any spatial features from the virtual environment, for example, in front of a plain white background with no wall or floor. The trial types for Object Comparison were similar to those for Scene Creation. Specifically, one third were Same Attached trials (the two objects to be compared had been previously located in front of neighbouring walls in one of the virtual rooms), one third were Same Detached trials (the two objects to be compared had been previously located in front of non-neighbouring walls in the same virtual room), and one third were Different trials (the two objects to be compared had been previously located in front of walls from different virtual rooms).

Participants responded to the cue "Combine" or "Compare" by pressing '1' on the button box once they had imagined either two conjoined walls, or two separate objects side-by-side. Participants were instructed to hold this image in mind until the end of the 7000 ms imagination phase. Once this imagination phase was complete, a jittered 1500 ms ISI ensued, followed by the presentation of the word "Confidence" for 3000 ms (confidence rating phase), during which participants rated their confidence in the accuracy and the vividness of the dual-stimulus mental image, using buttons 1 through 4 on the button box. They were instructed to give a rating of 4 if they felt they accurately remembered and vividly envisioned both stimuli during the previous "Combine" or "Compare" phase of the trial. If, however, they felt that one or both of the stimuli were not accurately remembered or vividly imagined, they were instructed to use a lower confidence rating that reflected the overall accuracy and vividness of the two stimuli. Confidence judgments were immediately followed by a question and answer (Q&A) phase that was designed to ascertain whether the participants had retrieved the correct walls/objects earlier in the trial. For Scene Creation trials, the word "Shape," "Placement," or "Concavity" appeared whereas for Object Comparison trials, one of "Height," "Width," or "Weight" was displayed (5000 ms stimulus duration). In response, participants were required to indicate whether the two stimuli they had retrieved were from the same category or different categories with respect to the word prompt. For instance, if the prompt was "Shape", participants were to indicate whether the recalled wall protrusions contained the same overall shape or two different shapes (i.e., same for both triangular/circular or different for one triangular and one circular). Similarly, if the prompt was "Weight," participants indicated whether both recalled objects were classified similarly or differently with regard to their weight (i.e., same for both heavy/light or different for one heavy and one light). Participants responded with 1 for same, and 2 for different. The Q&A stage was the final phase of each trial. Each trial lasted 35,500 ms in total and was followed by a jittered 3000 ms ITI.

The Scene Creation and Object Comparison conditions were designed to emphasize scene and object processing, respectively, by requiring participants to create a scene from two wall segments in the former (with 'scene' defined here as a geometrically defined space) or compare two objects in the latter. Although we cannot rule out the possibility that participants may have implicitly/automatically processed additional information (e.g., associated objects during Scene Creation or associated walls during Object Comparison), we are confident that the extensive pre-scan training and explicit instructions to focus on walls (without their associated objects) during Scene Creation trials and objects (in the absence of any spatial features) during Object Comparison trials led to predominant scene and object processing in these two conditions, respectively.

The scanned experimental task contained 120 trials in total - 60 Scene Creation and 60 Object Comparison trials, each comprised of 20 "Same Attached", 20 "Same Detached", and 20 "Different" trials. Thus, the task was a fully crossed 2 by 3 design. Trials were divided equally across five scanning runs (24 trials per run) and in each run, four trials from each of the six conditions were presented in a pseudo-randomized order.

2.4. Functional localizer task

Participants were administered two independent functional localizer runs in which they passively viewed photos of faces, objects, scenes, and scrambled versions of those same stimuli. Data from these runs allowed us to identify content-specific extrastriate functional regions of interest (ROI) in each participant for analyzing the main experimental data (i.e., lateral occipital cortex [LO], occipital place area [OPA], parahippocampal place area [PPA], and fusiform face area [FFA]). There were 25 blocks per run, and each run consisted of four blocks of faces, four blocks of objects, and four blocks of scenes, all separated by blocks of scrambled images of each category. Each block lasted 14.4 s and consisted of 32 grayscale stimuli each presented for 400 ms followed by a 50 ms interstimulus interval.

2.5. Image acquisition

Functional imaging was conducted using a 3T Siemens Tim Trio MRI scanner with a 32-channel head coil located at the MRI Facility of York University (Keele campus, Toronto, ON, Canada). Functional data were collected using a T2*-weighted echo-planar imagining (EPI) sequence with slices oriented parallel to the long axis of the hippocampus [25 oblique slices acquired in an interleaved order, slice thickness = 1.75 mm, interslice distance = 0 mm, voxel size = 1.5 × 1.5 × 1.75 mm, TR = 2000 ms, TE = 34 ms, matrix size = 128 × 128, field-of-view (FOV) = 192 mm, FA = 78°]. This high-resolution slice acquisition plan yielded a partial volume and provided full coverage of the temporal and occipital lobes. Seven sets of functional time-series data were collected from each participant, including five functional experiment runs and two functional localizer runs. Each functional experiment run lasted for 956 s (478 vol), and each functional localizer run lasted for 408 s (204 vol). The first four scans of each run were discarded to take into consideration the time for the MR signal to reach equilibrium. A high-resolution T1-weighted anatomical scan was also obtained for each participant

(MPRAGE sequence, slices = 192; voxel size = 1 mm^3, TR = 2300 ms; TE = 2.62 ms; FA = 9°; matrix size = 256 × 256), which was used for registration purposes and delineation of medial temporal lobe structures, including automated segmentation of the hippocampus and manual delineation of the entorhinal and perirhinal cortices (see 'ROI definitions').

2.6. Behavioral data analysis

Data were processed with custom MATLAB scripts and statistical analysis was carried out using RStudio (www.rstudio.com). The different phases of the scanned behavioral task trials (retrieval, imagination, confidence rating, and Q&A) were analyzed independently. For the retrieval phase, response times (RT) were analyzed using a two-way repeated-measures analysis of variance (ANOVA) with cue presentation (first vs. second cue) and retrieval target (wall vs. object) as within-subject factors. RTs for the imagination phase were submitted to a two-way repeated-measures ANOVA with distance (Same Attached, Same Detached, Different) and imagination type (Scene, Object) as within-subject factors.

For the confidence rating phase, we were primarily interested in confidence rating values, which were analyzed using a two-way repeated-measures ANOVA with spatial distance (Same Attached, Same Detached, Different) and imagination type (Scene, Object) as within-subject factors. Additional analyses pertaining to RTs as well as the impact of accuracy (i.e., correct vs. incorrect trials) on confidence ratings are described in the Supplemental Material.

Similarly, for the Q&A phase, we submitted accuracy to a two-way repeated-measures ANOVA with spatial distance (Same Attached, Same Detached, Different) and imagination type (Scene, Object) as within-subject factors. We also restricted our accuracy data analysis to high confidence rating trials only (rating value > 2) and analyzed the Q&A phase RT data (all trials as well as high confidence rating trials only), all of which are described fully in the Supplemental Material.

To avoid potential violations of parametric assumptions (e.g., normal distribution of data), statistical significance for all ANOVAs and post hoc tests was determined via a permutation-based approach. Specifically, original parametric statistic values (F or t values) were first calculated using standard ANOVAs or t-tests. The data were then permuted within participants to recalculate the statistic values and effect sizes (η_p^2 or Cohen's d) 10,000 times to yield a null distribution. Statistical significance (P value) was then determined by calculating the proportion of values from the null distribution that were greater than or equal to the original parametric statistic value. Similarly, the null distribution of effect sizes was used to derive 95% confidence intervals (CIs) of effect sizes. All P values for post hoc tests were corrected using the Holm–Bonferroni procedure.

2.7. fMRI data analysis

Preprocessing: Prior to statistical analysis, functional images were preprocessed using FEAT v6.00 and additional tools from FSL (Smith et al., 2004). The following preprocessing steps were conducted: (1) removal of the first four volumes to allow for stabilization of the initial signal; (2) brain extraction of anatomical images using the Brain Extraction Tool (BET) with threshold 0.5 (Smith, 2002); (3) motion correction using MCFLIRT (Jenkinson et al., 2002); (4) application of a high-pass temporal filter with default cut-off frequency of 100 s; (5) spatial smoothing with a 4-mm full width at half maximum (FWHM) Gaussian kernel for data submitted to univariate analyses (unsmoothed data were used for multivariate analyses); (6) co-registration of each participant's functional data to anatomical space using boundary based registration; and (7) normalization to the Montreal Neurological Institute 152 (MNI-152) 2-mm standard template using the FSL nonlinear registration tool (FNIRT) for univariate analyses.

Regions of Interest (ROI) definitions: Content-specific extrastriate ROIs (LO, OPA, PPA, and FFA) and medial temporal lobe ROIs (perirhinal cortex [PRC], entorhinal cortex [ERC], and hippocampus [HPC]) were defined *a priori* using functional and structural MRI data, respectively. Participant-specific masks in native space were used for multivariate classification analyses whereas group-derived probabilistic masks in standard template space were used for univariate analyses.

For the functional ROIs (Fig. 2A), data from the two functional localizer scans were fit to independent general linear models (GLMs), with one explanatory variable (EV) for each stimulus category (i.e., faces, objects, scenes, scrambled images) convolved with a gamma hemodynamic response function (HRF). Contrasts were specified to identify the LO (object > scrambled), OPA and PPA (scene > [object + face]), and FFA (face > [object + scene]) ROIs. The two functional localizer runs were then combined for each participant in a fixed-effects analysis and contiguous clusters of voxels were identified at a threshold $P < 0.001$ (uncorrected) to create participant-specific ROIs in native space. If widespread voxels were identified at this threshold, we looked at more conservative thresholds of $P < 0.0005$ and $P < 0.0001$, until a highly focal cluster survived that threshold. In contrast, if a contiguous set of voxels could not be identified at this threshold, we applied more liberal thresholds of $P < 0.005$ and $P < 0.01$, until a contiguous set of voxels was identified. Data from three participants required a more conservative threshold to delineate the PPA and OPA (two with $P < 0.0005$ and one with $P < 0.0001$). Two participants did not possess any significant clusters in the PPA or OPA, and the FFA could not be identified in one participant, even at a liberal threshold ($P < 0.01$). For the remaining participants the average number of voxels in each ROI was as follows (mean ± standard deviation): LO: 2448 ± 792.82 voxels; OPA: 601.53 ± 616.02 voxels; PPA: 1155.35 ± 635.81 voxels; FFA, 295.78 ± 101.21 voxels. Given our experimental aims and therefore the stimulus design in our main experimental task, we focused our data analysis and interpretation on object- and scene-sensitive ROIs, with the face-sensitive FFA serving as a control region.

For the anatomical ROIs (Fig. 2B-C), the HPC was localized using FreeSurfer7's recon-all function (Iglesias et al., 2015), and then visually inspected and edited by hand for accuracy for each participant. This parcellation function further segmented the HPC into the head, body and tail (head: 3508.05 ± 274.38 voxels; body: 2439.89 ± 166.98 voxels; tail: 1232.11 ± 134.78 voxels) with the parasubiculum being delineated in the head region, and the presubiculum and subiculum regions being delineated in the head and body regions. The perirhinal cortex (PRC, 4827.26 ± 1477.04 voxels) and entorhinal cortex (ERC, 1687.05 ± 372.55 voxels) were segmented manually according to the Insausti protocol (Insausti et al., 1998). To create group probabilistic ROI masks, participant-specific masks were normalized to MNI-152 space (2 mm), combined, thresholded at 50%, and binarized (LO: 1588 voxels; FFA, 173 voxels; OPA: 262 voxels; PPA: 925 voxels; HPC head: 593 voxels; HPC body: 387 voxels; HPC tail: 183 voxels; PRC: 503 voxels; ERC: 206 voxels).

Univariate analysis: The preprocessed data were denoised by deriving regressors from voxels unrelated to the experimental task and then regressing out noise components in the data using the GLMdenoise toolbox (Kay et al., 2013). Briefly, the GLMdenoise procedure implemented the following steps: first, a pool of noise voxels was identified, which consisted of voxels that demonstrated a poor initial model fit but possessed a mean signal intensity above a minimum threshold; second, a principal components analysis was performed on the time-series of these noise pool voxels to derive a set of noise components; third, these noise components were then systematically evaluated via model fit and a cross-validation approach; and fourth, the optimal number of noise components were then regressed out of the data. The denoised data for each run and participant were then submitted to individual GLMs that modeled both condition and trial phases for correct trials only (as defined by performance on the Q&A phase). Thus, there were 10 EVs in the model, each convolved with the gamma HRF: (1) Wall Retrieval; (2) Object Retrieval; (3) Scene Creation Same Attached; (4) Scene Creation

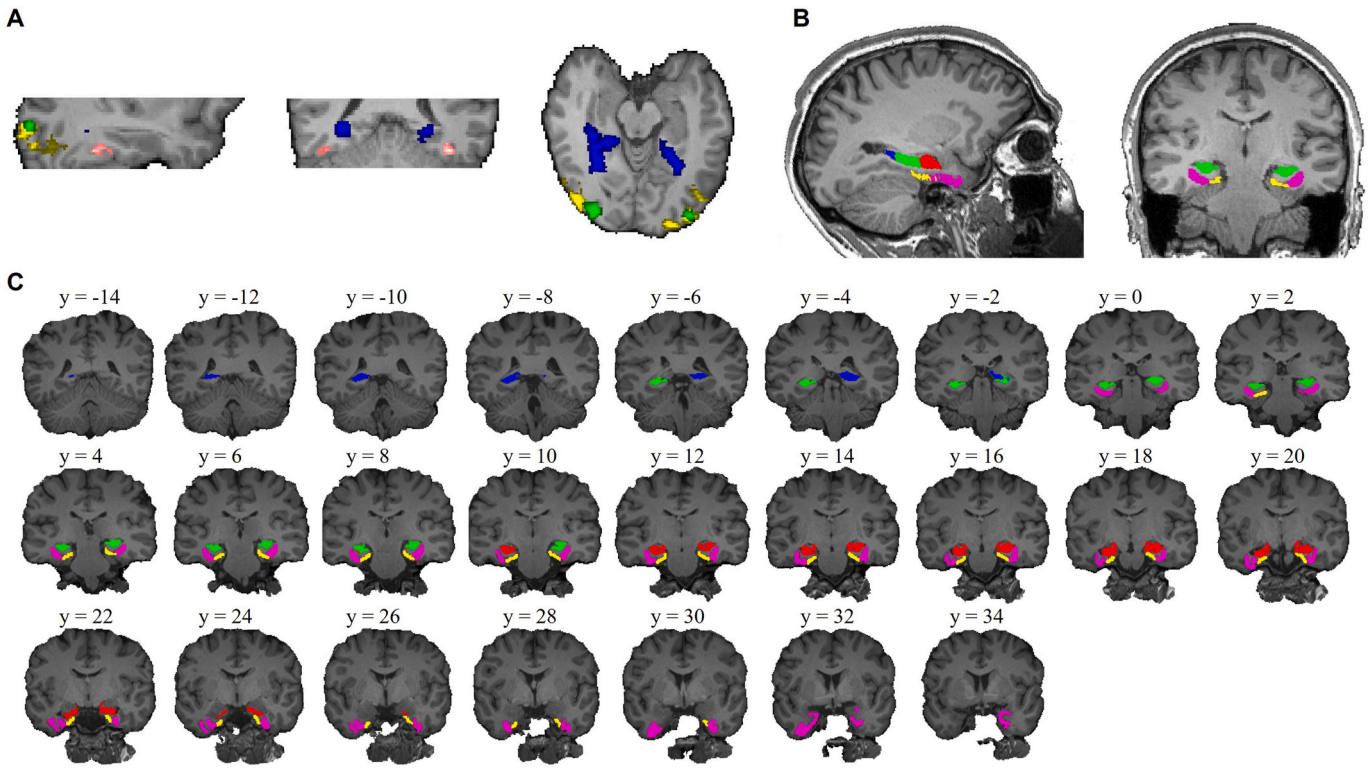


Fig. 2. Regions of interest (ROIs). **A**. Example functional ROIs (yellow: LO; green: OPA; blue: PPA; pink: FFA) from one participant, rendered in native functional space. **B-C**. Example anatomical ROIs from the same participant, rendered in native T1 space. ROIs include HPC subregions (red: head; green: body; blue: tail), PRC (purple), and ERC (yellow).

Same Detached; (5) Scene Creation Different; (6) Object Comparison Same Attached; (7) Object Comparison Same Detached; (8) Object Comparison Different; (9) Confidence Rating and Q&A phases (combined across Scene Creation and Object Comparison trials); and (10) All phases of all error trials (number of error trials for wall cue: 11.68 ± 6.53; object cue: 9.95 ± 8.16; mean ± SD).

To investigate whether our ROIs contribute to the retrieval and manipulation of learned elements during imagination, and whether this involvement differs across process and/or stimulus content, we set up contrasts to identify differences in brain activity associated with:

1. The retrieval of walls compared to objects (Wall Retrieval vs. Object Retrieval).
2. The mental creation of scenes compared to the mental comparison of objects ([Scene Creation Same Attached + Scene Creation Same Detached + Scene Creation Different] vs. [Object Comparison Same Attached + Object Comparison Same Detached + Object Comparison Different]).
3. The retrieval of walls compared to the mental creation of scenes (Wall Retrieval vs. all Scene Creation conditions).
4. The retrieval of objects compared to the mental comparison of objects (Object Retrieval vs. all Object Comparison conditions).
5. Retrieval compared to imagination ([Object Retrieval + Wall Retrieval] vs. [all Scene Creation conditions + all Object Comparison conditions]).

To investigate whether there are neural differences when a new mental scene or object representation is created based on learned elements from a single representation as opposed to multiple pre-existing representations, we set up contrasts to investigate:

6. The impact of spatial distance on the creation of scenes (pairwise comparisons between Scene Creation Same Attached vs. Scene Creation Same Detached vs. Scene Creation Different).
7. The impact of spatial distance on the mental comparison of objects (pairwise comparisons between Object Comparison Same Attached vs. Object Comparison Same Detached vs. Object Comparison Different).

For each participant, all five runs were combined in a second-level fixed effects analysis. Group-level inference was then conducted by applying a non-parametric approach implemented in FSL's randomise function in conjunction with threshold-free cluster enhancement (TFCE), with 5000 permutations and a corrected threshold of $p < 0.05$ (Smith and Nichols, 2009) for both whole-volume and ROI approaches.

Multivariate classification analysis: Statistical maps were created from denoised unsmoothed preprocessed data using a least squares – separate (LS-S) approach, in which a separate GLM was fit for each phase of each trial, with a single trial phase modeled as the regressor of interest and all other trial phases combined into a single nuisance regressor (Mumford et al., 2012). Due to the complexity of applying GLMdenoise to the LS-S approach, denoising was carried out using an independent component analysis approach as implemented by MELODIC in FSL (Beckmann and Smith, 2004). This resulted in each participant having 120 parameter estimate maps for wall retrieval, 120 parameter estimate maps for object retrieval, 60 parameter estimate maps for Scene Creation, and 60 parameter estimate maps for Object Comparison. Parameter estimate maps associated with correct trials (incorrect trials were not analyzed further) were then converted to *t*-statistic maps for classification analyses in order to decrease the contribution of noisy voxels (Misaki et al., 2010).

Classification analyses were conducted using the CoSMo MVPA toolbox (Oosterhof et al., 2016) and a linear support vector machine classifier (*L*2-norm regularized SVM with $C = 1$) with leave-one-run-out

cross-validation. To examine whether the retrieval and/or imagination of object and scene information were associated with differing patterns of multivariate activity, two 2-way classification analyses were implemented. Specifically, the SVM classifier was trained and tested on data from the retrieval phase to distinguish wall vs. object retrieval (i.e., Ret → Ret) and furthermore, trained and tested on data from the imagination phase to distinguish Scene Creation vs. Object Comparison (i.e., Ima → Ima). In addition to this, we explored whether patterns of activity during retrieval overlapped with those during imagination by training the classifier to distinguish wall vs. object retrieval using retrieval data and testing the trained classifier on the imagination data (i.e., Ret → Ima, cross-phase classification). Finally, to identify potential differences in activity associated with spatial distance, a 3-way classification (i.e. Same Attached vs. Same Detached vs. Different) was conducted on the Scene Creation and Object Comparison imagination data separately.

Both ROI-based and searchlight-based approaches were used in our classification analyses. For the former, classification accuracies for each participant and each ROI were entered into one-sample one-tailed *t* tests against chance (1/2 for 2-way classification analysis and 1/3 for 3-way classification analysis) to obtain original *t* values. The procedure was then repeated 10,000 times with data randomly sampled with replacement to obtain the (null) distribution of these *t* values. Bootstrapped p-values were then obtained by calculating the proportion of these resampled *t* values from the distribution that was greater than or equal to the original *t* values. Due to the selection of multiple ROIs, these bootstrapped p-values were further corrected using the false discovery rate approach. For the searchlight-based approach (Kriegeskorte et al., 2006), a sphere of 100 voxels was applied in conjunction with the linear SVM throughout the entire volume of each participant and the classification accuracy was assigned to the center voxel of the sphere. All participants' accuracy maps were normalized to MNI space and submitted to a permutation-based one-sample *t*-test as implemented in FSL's randomise function in conjunction with TFCE (5000 permutations, $p < 0.05$ corrected) for group-level inference.

3. Results

3.1. Behavioral performance

Retrieval phase: Participants were presented with two successive cues and required to recall their associated targets. A two-way repeated-measures ANOVA revealed a significant main effect of retrieval target on RTs ($F_{(1, 18)} = 5.60$, $P = 0.03$, $\eta_p^2 = 0.24$ [0, 0.25]; Fig. 3A), with participants taking longer to retrieve walls compared to objects. There was no main effect of cue presentation (i.e. RTs for 1st vs. 2nd cue; $F_{(1, 18)} = 0.06$, $P = 0.80$, $\eta_p^2 = 0.004$ [0, 0.25]) nor a two-way interaction effect ($F_{(1, 18)} = 0.09$, $P = 0.77$, $\eta_p^2 = 0.005$ [0, 0.25]).

Imagination phase: This phase required participants to imagine the items recalled during the retrieval phase either joined together to create a room corner (in the case of walls) or placed side-by-side (in the case of objects). A two-way repeated-measures ANOVA indicated an effect of imagination type on RTs ($F_{(1, 18)} = 10.90$, $P = 0.003$, $\eta_p^2 = 0.38$ [0, 0.25]; Fig. 3B), with participants taking longer for Scene Creation compared to Object Comparison. There was also a marginal main effect of spatial distance ($F_{(2, 36)} = 3.26$, $P = 0.055$, $\eta_p^2 = 0.15$ [0, 0.19]), although post-hoc comparisons did not reveal a significant difference between

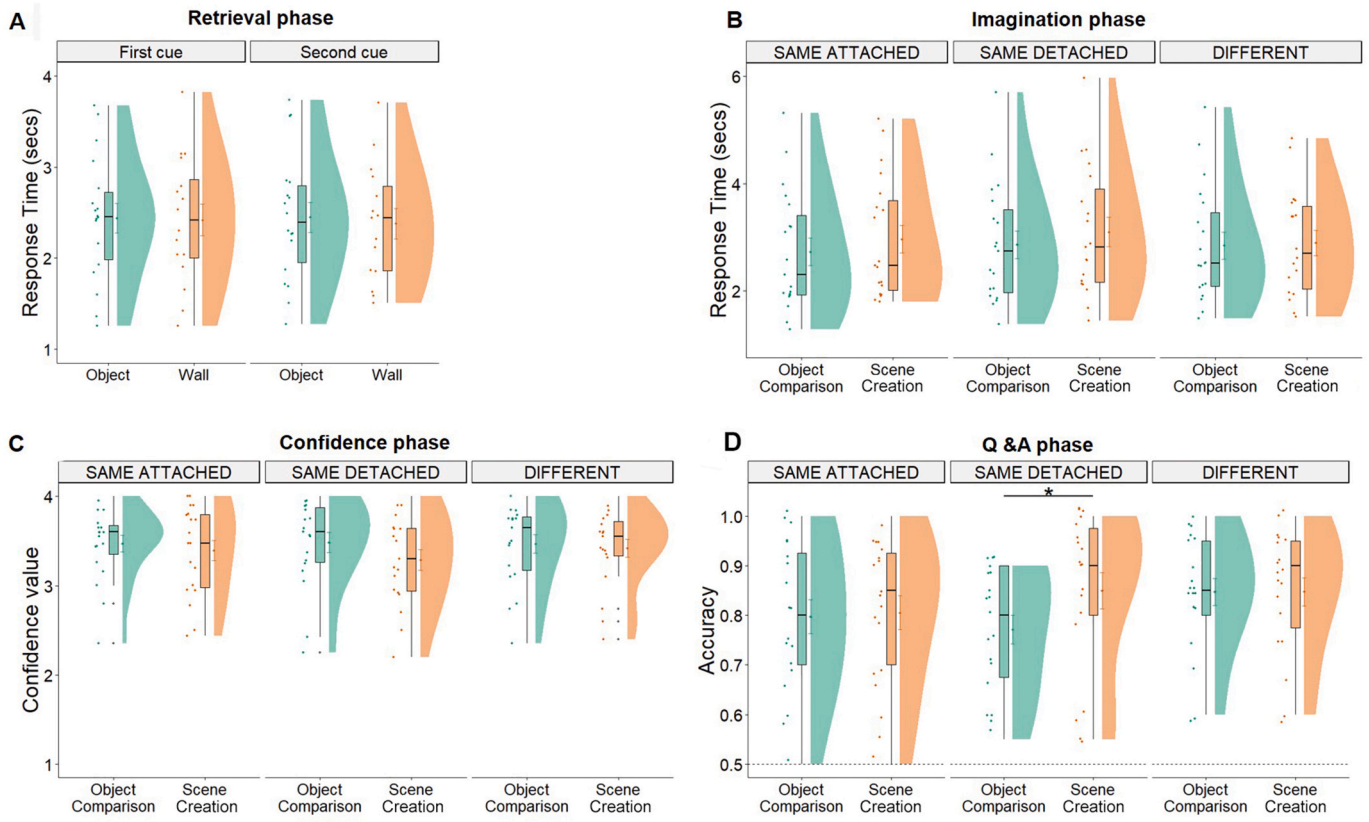


Fig. 3. Behavioral task performance for the different trial phases. Each raincloud plot provides information about individual observations, boxplots, and overall tendencies in the distribution (Allen et al., 2021). **A.** RTs for participants to retrieve objects and walls from two successive cues during the Retrieval phase **B.** RTs for participants to compare two objects (Object Comparison) or combine two walls (Scene Creation) across different spatial distances in the Imagination phase **C.** Confidence rating values reflecting participants' self-perceived accuracy and vividness of their mental images during the Imagination phase - correct and incorrect trials combined. **D.** Participant accuracy to questions in the Q & A phase pertaining to their mental images during the Imagination phase - all confidence levels combined.*$P < 0.05$.

conditions (Same Attached vs. Same Detached: $P = 0.08$; Same Attached vs. Different: $P = 0.16$; Same Detached vs. Different: $P = 0.70$). No interaction effect was found ($F_{(2, 36)} = 0$, $P = 0.996$, $\eta_p^2 = 0.0002$ [0, 0.18]).

Confidence rating phase: Following the imagination phase, participants were asked to rate their confidence in both the accuracy and vividness of the dual-stimulus mental image. A two-way repeated-measures ANOVA on the confidence rating values was conducted. There was a significant main effect of imagination type ($F_{(1, 18)} = 4.69$, $P = 0.04$, $\eta_p^2 = 0.21$ [0, 0.24]; Fig. 3C), with participants giving higher confidence rating values for Object Comparison compared to Scene Creation. There was no main effect of spatial distance ($F_{(2, 36)} = 1.05$, $P = 0.37$, $\eta_p^2 = 0.06$ [0, 0.02]), nor was there a two-way interaction effect ($F_{(2, 36)} = 2.30$, $P = 0.11$, $\eta_p^2 = 0.11$ [0, 0.02]) on confidence rating values. Analyses of RT data and impact of accuracy on confidence ratings can be found in the Supplemental Material.

Q&A phase: During this phase, participants were asked to indicate whether the two walls/objects they had imagined were from the 'same' category or 'different' categories with respect to a cued physical property. A two-way repeated-measures ANOVA on accuracy revealed a non-significant main effect of imagination type ($F_{(1, 18)} = 2.27$, $P = 0.15$, $\eta_p^2 = 0.11$ [0, 0.24]; Fig. 3D) and a marginal main effect of spatial distance ($F_{(2, 36)} = 2.94$, $P = 0.06$, $\eta_p^2 = 0.14$ [0, 0.18]), as well as a significant interaction effect ($F_{(2, 36)} = 3.64$, $P = 0.03$, $\eta_p^2 = 0.17$ [0, 0.18]). The interaction effect was driven by significantly higher accuracy for Scene Creation compared to Object Comparison in the Same Detached trials only ($t_{(18)} = 3.03$, $P = 0.007$, Cohen's d = 0.72 [−0.49, 0.50]), but not in the Same Attached trials ($t_{(18)} = 0.31$, $P = 0.76$, Cohen's d = 0.07 [−0.50, 0.49]) or Different trials ($t_{(18)} = 0$, $P = 1$, Cohen's d = 0 [−0.48, 0.50]). Analyses of RT data and high confidence trials only (accuracy and RT data) are reported in the Supplemental Material.

3.2. Univariate results

ROI analysis: We began by identifying the voxels responding to our contrasts of interest in selected ROIs. As expected, significant findings were concentrated primarily in object- and scene-sensitive ROIs, with fewer findings of note in the face-sensitive FFA. Thus, in light of our experimental questions, we limit our discussion to the former. For the retrieval of walls compared to objects, results revealed significant activations in the LO only. In contrast, significant activations were found in the LO, OPA, PPA, PRC, HPC head and body for the retrieval of objects compared to walls (Fig. 4A, Table 1). During the imagination phase, results revealed significant activations for the creation of scenes compared to the mental comparison of objects, including voxels in the LO, OPA, PPA, and HPC tail, and significant voxels in the PPA for the

Table 1
Univariate brain activation within ROIs when comparing scene and object conditions within retrieval and within imagination.

Contrast	ROI (Left/ Right)	Peak Voxel			*P(corr)* value	Cluster Size
		x	y	z		
Retrieval: Wall > Object	LO (R)	54	−64	6	0.023	25
Retrieval: Object > Wall	LO (R)	36	−84	14	<0.001	84
	LO (L)	−26	−86	20	<0.001	51
	LO (R)	42	−82	6	0.02	22
	OPA (L)	−24	−84	18	<0.001	94
	OPA (R)	34	−84	22	0.001	71
	PPA (L)	−24	−68	−12	<0.001	187
	PPA (R)	24	−42	−16	<0.001	231
	PRC (R)	30	−8	−34	0.018	11
	PRC (L)	−30	−26	−22	0.004	22
	HPC head (L)	−28	−16	−20	0.018	24
	HPC body (R)	26	−28	−8	0.012	13
Imagination: Scene > Object	LO (R)	36	−84	14	<0.001	87
	LO (L)	−26	−86	22	0.003	18
	OPA (L)	−26	−84	20	<0.001	86
	OPA (R)	36	−86	20	<0.001	70
	PPA (L)	−22	−74	−6	<0.001	53
	HPC tail (R)	22	−40	4	0.018	18
Imagination: Object > Scene	PPA (L)	−30	−46	−8	0.008	30

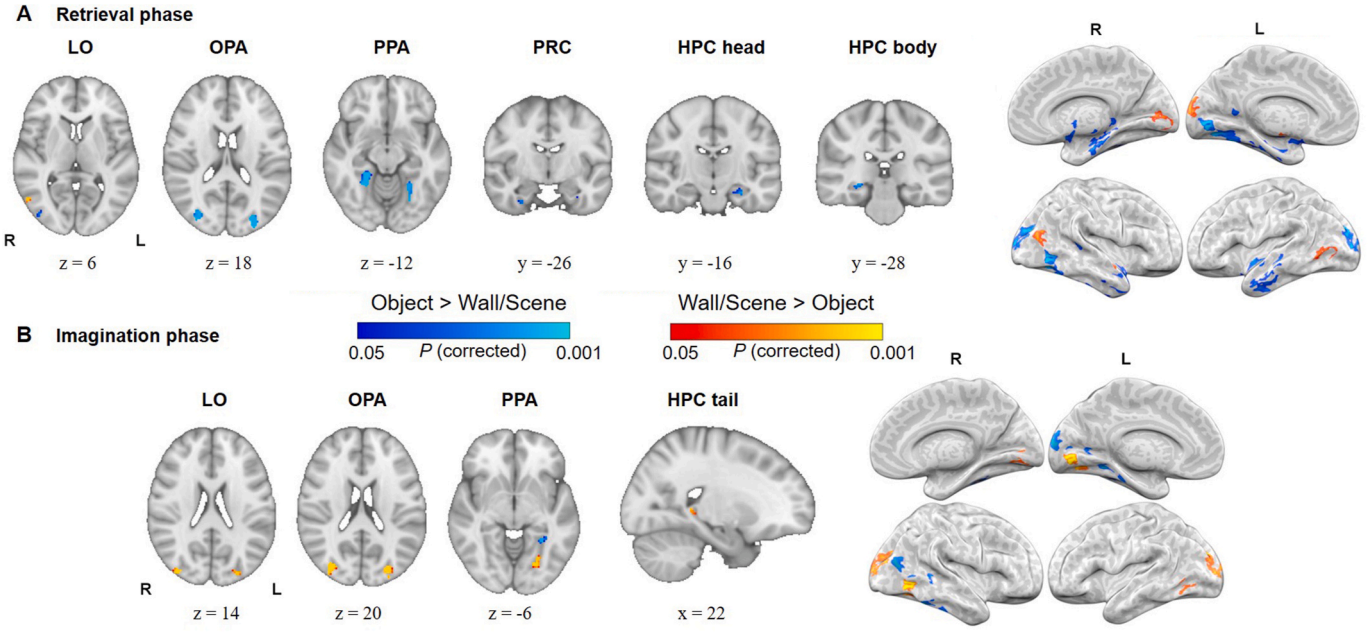


Fig. 4. Results of univariate ROI and whole-volume analyses comparing object and wall/scene conditions within the retrieval and imagination phases. **A.** For the retrieval phase, warm colors correspond to regions showing greater activation in retrieval of walls and cool colors greater activation in retrieval of objects. **B.** For the imagination phase, warm colors correspond to regions showing greater activation in creation of scenes and cool colors greater activation in mental comparison of objects. Activity surviving a corrected threshold of $P < 0.05$ (TFCE with 5000 permutations) is rendered on the MNI-152 template. R: right; L: left.

opposite direction (Fig. 4B, Table 1).

When comparing Wall Retrieval and Scene Creation, the former was associated with significant activity in a number of ROIs including the LO, OPA, PPA, HPC body, PRC and ERC, whereas the latter was associated with significant activity in the HPC tail ROI only (Table 2). In contrast, for object stimuli, significantly greater activity was observed during retrieval compared to mental comparison in all of the ROIs examined (Table 2), with the reverse contrast yielding no significant differences. Comparing retrieval and imagination for objects and walls combined produced significant differences in association with retrieval only in the LO, OPA PPA, FFA, PRC, ERC, HPC body and HPC tail (Fig. 5A, Table 2).

No significant voxels sensitive to the contrasts of spatial distance for the mental creation of scenes or mental comparison of objects were found within these ROIs.

Whole-volume analysis: To examine whether the patterns of activity observed within our selected ROIs during wall/object retrieval and scene/object imagination were accompanied by changes in activity in other cortical/subcortical regions, we examined whole-volume maps of activity using our contrasts of interest. During retrieval, more widespread activity was observed for objects compared to walls, whereas the reverse pattern was found during the imagination phase, with Scene Creation being associated with a greater extent of activity in contrast to Object Comparison. Specifically, for the retrieval of walls compared to objects (Fig. 4A, warm colors), we identified regions of significant activity outside of our selected ROIs in the bilateral occipital pole (left peak x, y, z = −2, −100, 12; right peak x, y, z = 4, −94, 14). In comparison, for the reverse contrast (Fig. 4A, cool colors), we observed significantly greater activity in the left perirhinal cortex (−30, −10, −28) and left hippocampus head (−24, −6, −20), converging with our ROI results. Moreover, we also identified a wider distribution of significant activity beyond our ROIs, including bilateral occipital pole (−30, −94, 20; 26, −102, 4), bilateral lingual gyrus (−14, −74, −8; 10, −84, −4), bilateral temporal occipital fusiform cortex (−30, −56, −18; 34, −44, −22), middle temporal gyrus (−50, −4, −32; 68, −24, −12), and left temporal fusiform cortex (−38, −18, −24) (see Table 3 for full report of significant clusters). For the mental creation of scenes compared to the mental comparison of objects, significant clusters of activity associated with the former (Fig. 4B, warm colors) were observed in the region of the right PPA (30, −56, −8) and extending across the LO and OPA bilaterally (−26, −82, −20; 34, −80, 12), consistent with the ROI results. In addition, we also identified a widespread network of activity across the occipital and temporal lobes, including bilateral lingual gyrus (−6, −78, −4; 2, −84, −2), right occipital fusiform gyrus (28, −70, −8), bilateral occipital pole (−18, −102, 12; 36, −90, 12), and right inferior and middle temporal gyrus (54, −60, −10). In contrast, for the mental comparison of objects compared to the creation of scenes (Fig. 4B, cool colors), highly focal clusters were identified beyond our selected ROIs in bilateral occipital pole (−2, −98, 16; 8, −98, 18).

Broadly speaking, comparable patterns of whole-volume activity were observed when retrieval was contrasted with imagination either within stimulus type or with scene and object conditions combined, with retrieval associated with clusters of activity overlapping with our chosen ROIs, and imagination associated with significant activity primarily in lateral temporal and inferior parietal (i.e., angular gyrus, precuneus) regions. For example, compared to retrieval, imagination (scene and object conditions combined; Fig. 5B, warm colors) was found to be associated with significant bilateral activity in the middle temporal gyrus (−48, −28, −6; 48, −36, 0), superior temporal gyrus (−60, −40, 40; 58, −28, −2) and left supramarginal gyrus (−48, −48, 10). On the other hand, a large swathe of activity was observed during retrieval (Fig. 5B, cool colors) encompassing the LO (52, −76, −6), FFA/PPA (−22, −62, −16; 32, −58, −18), hippocampus (20, −30, −6; −14, −32, −6), and extending into the lingual gyrus (−18, −66, −4; 16, −72, −10) and occipital pole (16, −98, 4; 0, −100, 6) (see Table 4 for a full report of significant clusters).

Table 2
Univariate brain activation within ROIs when comparing imagination and retrieval within each stimulus category as well as with both categories combined.

Contrast	Region (Left/ Right)	Peak Voxel			*P* value	Cluster Size
		x	y	z		
Wall Retrieval > Scene Creation	LO (R)	48	−78	−6	<0.001	471
	LO (L)	−42	−78	−4	0.001	236
	LO (L)	−34	−88	16	<0.001	174
	LO (R)	40	−84	14	<0.001	141
	LO (L)	−28	−80	24	0.025	4
	OPA (L)	−26	−88	14	<0.001	124
	OPA (R)	34	−80	14	<0.001	82
	PPA (L)	−24	−38	−18	<0.001	469
	PPA (R)	32	−30	−20	<0.001	412
	FFA (R)	46	−46	−24	<0.001	83
	HPC body (R)	18	−32	−6	<0.001	6
	HPC body (L)	−20	−30	−8	0.004	2
	HPC body (L)	−16	−32	−8	0.018	1
	HPC body (L)	−16	−34	−4	0.015	1
	PRC (R)	30	−28	−26	<0.001	128
	PRC (L)	−30	−30	−24	<0.001	82
	PRC (L)	−30	−6	−32	0.006	43
	ERC (L)	−26	−28	−24	0.005	28
	ERC (R)	26	−28	−24	0.002	12
Scene Creation > Wall Retrieval	HPC tail (R)	22	−40	2	0.047	1
Object Retrieval > Object Comparison	LO (R)	44	−70	−8	<0.001	390
	LO (L)	−42	−78	−4	0.001	224
	LO (L)	−36	−86	14	<0.001	175
	LO (R)	40	−84	14	<0.001	141
	LO (L)	−28	−78	22	0.015	5
	OPA (L)	−26	−88	14	<0.001	124
	OPA (R)	34	−80	14	<0.001	82
	PPA (L)	−22	−38	−18	<0.001	459
	PPA (R)	32	−30	−20	<0.001	409
	FFA (R)	42	−62	−10	<0.001	59
	FFA (L)	−38	−52	−20	0.029	1
	HPC head (L)	−24	−10	−26	0.022	15
	HPC head (R)	34	−16	−22	0.029	3
	HPC head (R)	36	−20	−18	0.039	1
	HPC body (R)	24	−28	−8	<0.001	25
	HPC body (L)	−20	−30	−8	<0.001	14
	HPC body (L)	−22	−22	−16	0.024	6
	HPC body (R)	36	−24	−16	0.046	1
	HPC body (L)	−34	−32	−12	0.04	1
	HPC tail (R)	22	−36	2	0.007	23
	HPC tail (L)	−16	−36	−2	0.018	2
	HPC tail (L)	−22	−36	0	0.021	1
	HPC tail (L)	−18	−36	4	0.031	1
	PRC (R)	30	−28	−26	<0.001	177
	PRC (L)	−32	−26	−24	<0.001	93
	PRC (L)	−30	−6	−32	0.002	55
	ERC (L)	−22	−26	−26	<0.001	44
	ERC (R)	28	−26	−28	<0.001	20
Object Comparison > Object Retrieval	NULL					
Retrieval > Imagination	LO (R)	44	−70	−8	<0.001	436
	LO (L)	−42	−72	−6	0.001	231
	LO (L)	−36	−86	14	<0.001	178

(*continued on next page*)

Table 2 (*continued*)

Contrast	Region (Left/ Right)	Peak Voxel x	y	z	*P* value	Cluster Size
(Scene and Object combined)	LO (R)	40	−84	14	<0.001	141
	LO (L)	−28	−80	24	0.016	5
	OPA (L)	−26	−88	14	<0.001	124
	OPA (R)	34	−80	14	<0.001	82
	PPA (L)	−24	−34	−20	<0.001	469
	PPA (R)	32	−30	−20	<0.001	418
	FFA (R)	42	−46	−24	<0.001	79
	FFA (L)	−38	−52	−20	0.035	1
	HPC body (L)	−20	−30	−8	<0.001	18
	HPC body (R)	18	−32	−6	<0.001	12
	HPC body (L)	−34	−32	−12	0.037	1
	HPC tail (R)	24	−36	2	0.024	7
	HPC tail (R)	14	−36	−2	0.02	2
	HPC tail (L)	−14	−36	−2	0.029	2
	HPC tail (L)	−18	−36	4	0.037	1
	PRC (R)	32	−28	−26	<0.001	182
	PRC (L)	−32	−30	−26	<0.001	180
	ERC (L)	−22	−28	−24	<0.001	59
	ERC (R)	26	−28	−26	<0.001	20
Imagination > Retrieval (Scene and Object combined)	NULL					

Finally, similar to our ROI analysis, no significant clusters of activity were found at the whole-volume level in association with any of the contrasts pertaining to spatial distance in relation to the mental creation of scenes or the mental comparison of objects.

3.3. Multivariate results

Scene vs. object information 2-way classification: When the classifier was trained and tested on the retrieval phase data (Ret → Ret) in our ROI-based analysis, significant above chance decoding accuracies were observed in the FFA ($t_{(17)} = 3.07$, $P_{FDR} = 0.006$), OPA ($t_{(16)} = 4.98$, $P_{FDR} < 0.001$), PPA ($t_{(16)} = 5.72$, $P_{FDR} < 0.001$), LO ($t_{(18)} = 11.62$, $P_{FDR} < 0.001$), ERC ($t_{(18)} = 3.07$, $P_{FDR} = 0.044$), HPC head ($t_{(18)} = 3.24$, $P_{FDR} = 0.005$) and HPC body ($t_{(18)} = 2.55$, $P_{FDR} = 0.015$), but not in PRC and HPC tail (PRC: $t_{(18)} = 1.20$, $P_{FDR} = 0.14$; HPC tail: $t_{(18)} = 0.73$, $P_{FDR} = 0.24$) (Fig. 6A left). In contrast, when the imagination phase data were submitted to the classifier (Ima → Ima; Fig. 6A middle), significant above chance decoding accuracies were found in the PPA ($t_{(16)} = 4.46$, $P_{FDR} < 0.001$) and LO ($t_{(18)} = 8.44$, $P_{FDR} < 0.001$), but not in other ROIs (FFA: $t_{(17)} = 1.70$, $P_{FDR} = 0.097$; OPA: $t_{(16)} = 2.15$, $P_{FDR} = 0.053$; ERC: $t_{(18)} = 1.26$, $P_{FDR} = 0.17$; PRC: $t_{(18)} = 2.19$, $P_{FDR} = 0.053$; HPC head: $t_{(18)} = 0.82$, $P_{FDR} = 0.27$; HPC body: $t_{(18)} = -0.12$, $P_{FDR} = 0.62$; HPC tail: $t_{(18)} = -1.19$, $P_{FDR} = 0.88$). Lastly, cross-phase decoding (i.e., training on retrieval data and testing on imagination data; Ret → Ima) revealed significant above chance decoding accuracy in the HPC tail only (Fig. 6A right, $t_{(18)} = 4.22$, $P_{FDR} = 0.002$; FFA: $t_{(17)} = 1.17$, $P_{FDR} = 0.38$; OPA: $t_{(16)} = -2.69$, $P_{FDR} = 0.99$; PPA: $t_{(16)} = -0.23$, $P_{FDR} = 0.76$; LO: $t_{(18)} = -2.67$, $P_{FDR} = 0.99$; ERC: $t_{(18)} = 0.99$, $P_{FDR} = 0.38$; PRC: $t_{(18)} = 0.41$, $P_{FDR} = 0.52$; HPC head: $t_{(18)} = 0.44$, $P_{FDR} = 0.52$; HPC body: $t_{(18)} = 1.47$, $P_{FDR} = 0.36$).

A searchlight-based approach revealed a widespread network of occipito-temporal regions during the retrieval phase, including temporal occipital fusiform cortex, lateral occipital cortex, inferior and middle temporal gyrus, parahippocampal cortex, and throughout the HPC (Fig. 6B and C left). During the imagination phase, significant clusters of above-chance decoding were observed in lateral occipital cortex, occipital fusiform gyrus, lingual gyrus, and inferior and middle temporal gyri (Fig. 6B and C middle). For the cross-phase searchlight-based decoding analysis, significant above-chance decoding was found predominantly in the left hemisphere, including inferior temporal gyrus, temporal fusiform cortex, parahippocampal cortex including in the region of the PPA, and the HPC body and tail (Fig. 6B and C right). Full details of significant clusters are provided in Table 5.

Spatial distance (i.e. Same Attached vs. Same Detached vs. Different) 3-way classification: Within Scene Creation and Object Comparison, it was not possible to decode spatial distance on the basis of the imagination phase data. Mean decoding accuracies were not significantly above chance in any of the ROIs and similarly, a whole-volume searchlight-based approach did not reveal any significant clusters for either Scene Creation or Object Comparison.

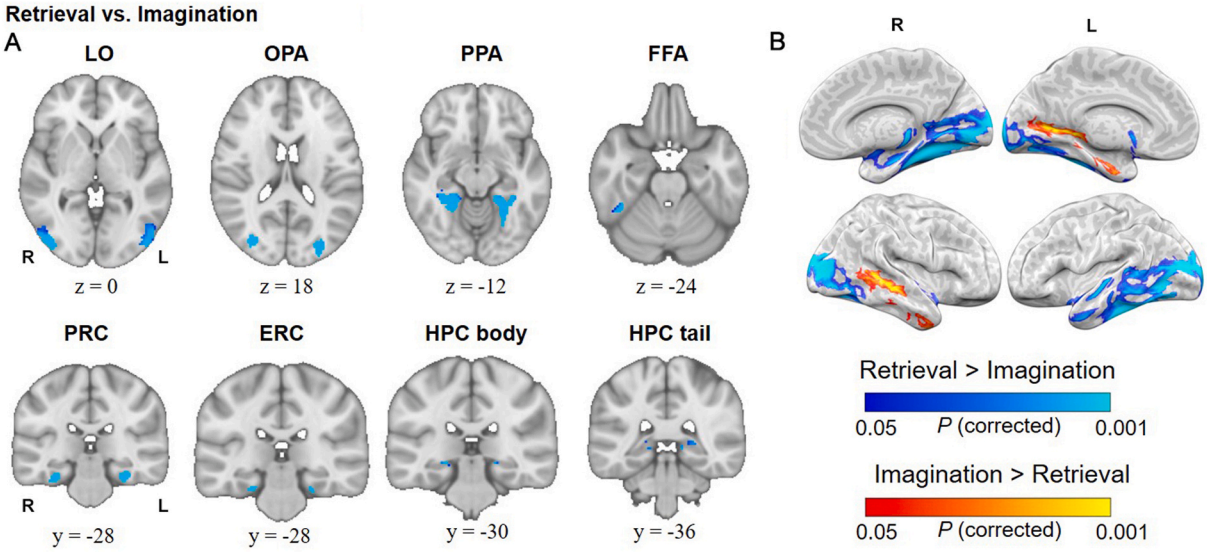


Fig. 5. Results of univariate **A.** ROI and **B.** whole-volume analyses comparing retrieval and imagination phases for object and wall/scene conditions combined. Cool colors correspond to regions showing greater activation during retrieval compared to imagination, whereas warm colors depict areas showing the reverse pattern of activity. Activity surviving a corrected threshold of $P < 0.05$ (TFCE with 5000 permutations) are rendered on the MNI-152 template.

Table 3
Univariate whole brain activation when comparing scene and object conditions within retrieval and within imagination.

Contrast	Region (Left/Right)	Peak Voxel			*P(corr)* value	Cluster Size
		x	y	z		
Retrieval: Wall > Object	Occipital Pole (L)	−2	−100	12	0.024	103
	Occipital Pole (R)	4	−94	14	0.028	
Retrieval: Object > Wall	Occipital Pole (L)	−30	−94	20	<0.001	4713
	Occipital Pole (R)	26	−102	4	<0.001	
	Lingual Gyrus (R)	10	−84	−4	<0.001	
	Lingual Gyrus (L)	−14	−74	−8	<0.001	
	Occipital Fusiform Gyrus (R)	22	−74	−8	<0.001	
	Temporal Occipital Fusiform Cortex (R)	34	−44	−22	<0.001	
	Temporal Occipital Fusiform Cortex (L)	−30	−56	−18	<0.001	
	Middle Temporal Gyrus (L)	−50	−4	−32	0.008	475
	Temporal Pole (L)	−56	4	−28	0.015	
	Superior Temporal Gyrus (L)	−60	−4	−10	0.028	
	Amygdala, Hippocampus (L)	−24	−6	−20	0.022	441
	Frontal Orbital Cortex (L)	−30	12	−18	0.027	
	Insular Cortex (L)	−34	8	−16	0.032	
	Perirhinal Cortex (L)	−30	−10	−28	0.033	
	Hippocampus (L)	−28	−16	−20	0.035	
	Temporal Fusiform Cortex (L)	−30	−10	−36	0.036	
	Temporal Fusiform Cortex (L)	−38	−18	−24	0.023	166
	Perirhinal Cortex (L)	−30	−26	−22	0.023	
	Middle Temporal Gyrus (R)	68	−24	−12	0.035	29
	Insula Cortex (L)	−40	−6	−8	0.026	18
	Frontal Orbital Cortex (R)	30	8	−18	0.04	7
	Insular Cortex (R)	40	−18	−2	0.042	5
	Hippocampus (L)	−34	−20	−16	0.049	2
	Precuneus Cortex (L)	−4	−56	6	0.046	2
	Insular Cortex (L)	−40	−12	−8	0.046	1
Imagination: Scene > Object	Lingual Gyrus (L)	−6	−78	−4	<0.001	998
	Lingual Gyrus (R)	2	−84	−2	<0.001	
	Occipital Fusiform Gyrus (R)	28	−70	−8	0.004	
	Occipital Fusiform Gyrus (L)	−22	−74	−6	0.007	
	Temporal Occipital Fusiform Cortex (R)	30	−56	−8	0.009	
	Lateral Occipital Cortex (R)	34	−80	12	<0.001	572
	Occipital Pole (R)	36	−90	12	0.002	
	Occipital Pole (L)	−18	−102	12	0.008	283
	Lateral Occipital Cortex (L)	−30	−90	4	0.015	
	Lateral Occipital Cortex (L)	−26	−82	20	0.007	92
	Inferior Temporal Gyrus (R)	54	−60	−10	0.018	77
	Middle Temporal Gyrus (R)	64	−56	−6	0.036	
Imagination: Object > Scene	Occipital Pole (R)	8	−98	18	0.019	240
	Occipital Pole (L)	−2	−98	16	0.02	

4. Discussion

Using fMRI with a novel behavioral paradigm that was designed to control and provide a systematic evaluation of imagined content, the current study provides insight into the neural correlates associated with the retrieval and manipulation of object and spatial scene information.

Both ROI and whole volume analyses revealed that the retrieval of objects was associated with a greater extent of activity compared to the retrieval of walls. Although we cannot determine to what extent the observed activity was driven by the perception of the presented cues versus the retrieval of content from memory, it is plausible that our use of everyday objects as target stimuli resulted in the spontaneous retrieval of semantic information, which was particularly emphasized during the object retrieval condition, leading to significant activity in regions involved in the processing of object-related perceptual, conceptual and contextual information such as LO, PRC and parahippocampal cortex (e.g., Aminoff et al., 2013; Clarke and Tyler, 2014; Grill-Spector et al., 2001; Martin et al., 2018). In contrast, the ROI analysis revealed that only LO showed significantly greater activity for the retrieval of walls compared with objects. In this condition, participants were shown an object cue and were then required to retrieve the associated wall. Given that LO is known to be involved in the perceptual processing of objects (Kourtzi and Kanwisher, 2001; Malach et al., 1995), it is conceivable that the observed wall retrieval-related neural activity can be explained, at least in part, by the cued recall format of the retrieval phase and the perceptual processing of the onscreen retrieval cue, which would have naturally elicited activation in ventral-stream regions mediating visual perception (of course, this argument also applies to the interpretation of the results in the object retrieval condition, which elicited activity in regions known to be involved in the perceptual processing of scenes). Neural activity related to the retrieval of the associated stimulus from memory would have followed this perceptual-related activity, and consistent with this two-stage cued recall process, we observed significant activity in 'traditional' perceptual- and memory-related brain regions (across both ROI and

Table 4
Univariate whole brain activation when comparing imagination and retrieval within each stimulus category as well as with both categories combined.

Contrast	Region (Left/Right)	Peak Voxel			*P(corr)* value	Cluster Size
		x	y	z		
Wall Retrieval > Scene Creation	Lingual Gyrus (R)	18	−68	−12	<0.001	15486
	Lingual Gyrus (L)	−16	−68	−12	<0.001	
	Lateral Occipital Cortex	−36	−88	4	<0.001	
	Occipital Pole (R)	14	−98	4	<0.001	
	Occipital Fusiform Gyrus (R)	26	−70	−10	<0.001	
	Occipital Pole (L)	−2	−104	8	<0.001	
	Occipital Fusiform Gyrus	−22	−80	−4	<0.001	
	Hippocampus (R)	20	−28	−6	<0.001	
	Temporal Pole, Parahippocampal Gyrus (L)	−26	4	−22	0.037	42
Scene Creation > Wall Retrieval	Middle Temporal Gyrus (L)	−48	−30	−6	0.004	724
	Superior Temporal Gyrus (L)	−62	−30	0	0.004	
	Angular Gyrus (L)	−62	−56	12	0.004	
	Middle Temporal Gyrus (R)	48	−36	0	0.004	461
	Superior Temporal Gyrus (R)	58	−28	−2	0.004	
	Supramarginal Gyrus (R)	54	−40	6	0.006	
	Precuneus (L)	−26	−52	10	0.01	30
	Middle Temporal Gyrus (R)	52	−50	10	0.048	2
Object Retrieval > Object Comparison	Temporal Occipital Fusiform Cortex (L)	−24	−56	−18	<0.001	15981
	Occipital Pole (R)	38	−92	4	<0.001	
	Occipital Fusiform Gyrus (R)	32	−62	−16	<0.001	
	Temporal Occipital Fusiform Cortex (R)	36	−60	−16	<0.001	
	Parahippocampal Gyrus (L)	−18	−30	−6	<0.001	
	Lateral Occipital Cortex (R)	34	−84	2	<0.001	
	Lingual Gyrus (L)	−24	−90	2	<0.001	
	Occipital Pole (L)	−28	−92	2	<0.001	
	Lateral Occipital Cortex (L)	−30	−86	0	<0.001	
	Amygdala (R)	22	−4	−12	0.019	298
	Planum Polare (R)	40	−18	−2	0.02	
	Temporal Pole (R)	34	12	−22	0.021	
	Frontal Orbital Cortex (R)	28	10	−20	0.025	
	Insular (R)	38	4	−14	0.038	
Object Comparison > Object Retrieval	Middle Temporal Gyrus (L)	−48	−38	0	0.005	630
	Superior Temporal Gyrus (L)	−66	−40	4	0.008	
	Supramarginal Gyrus (L)	−48	−48	10	0.012	
	Angular Gyrus (L)	−62	−56	12	0.013	
	Lateral Occipital Cortex (L)	−58	−64	18	0.023	
	Superior Temporal Gyrus (R)	48	−36	2	0.008	186
	Precuneus (L)	−24	−50	12	0.011	25
	Precuneus (R)	28	−48	12	0.047	2
	Lateral Occipital Cortex (L)	−48	−62	12	0.05	1
Retrieval > Imagination (combined)	Temporal Occipital Fusiform Cortex (R)	32	−58	−18	<0.001	17054
	Occipital Pole (R)	16	−98	4	<0.001	
	Hippocampus (R)	20	−30	−6	<0.001	
	Occipital Fusiform Gyrus (R)	32	−62	−16	<0.001	
	Temporal Occipital Fusiform Cortex (L)	−22	−62	−16	<0.001	
	Hippocampus (L)	−14	−32	−6	<0.001	
	Lateral Occipital Cortex (R)	52	−76	−6	<0.001	
	Lingual Gyrus (L)	−18	−66	−14	<0.001	
	Occipital Fusiform Gyrus (L)	−28	−78	−16	<0.001	
	Occipital Pole	0	−100	6	<0.001	
	Lingual Gyrus (R)	16	−72	−10	<0.001	
Imagination > Retrieval (combined)	Middle Temporal Gyrus (L)	−48	−28	−6	0.003	801
	Superior Temporal Gyrus (L)	−60	−40	4	0.004	
	Supramarginal Gyrus (L)	−48	−48	10	0.019	
	Middle Temporal Gyrus (R)	48	−36	0	0.004	437
	Superior Temporal Gyrus (R)	58	−28	−2	0.005	
	Precuneus (L)	−26	−52	10	0.004	47

whole-volume analyses), reflecting the success of our paradigm at using visual stimuli to cue the retrieval of mnemonic representations. The next step in the evolution of our paradigm would be to modify it so that visual content is matched across different conditions in the retrieval phase (e.g., the same visual content is used to elicit object versus scene retrieval). In doing so, we could further explore if our current retrieval results are weighted more towards bottom-up perceptual input versus top-down mnemonic retrieval. Although this will be an important step moving forward, we contend that our current retrieval results are not solely explained by perceptual processing of the onscreen retrieval cue, as our behavioral results demonstrate that participants were complying with task instructions and each participant had substantial behavioral training on the task prior to conducting the fMRI experiment. In other words, our participants were clearly retrieving object and scene representations from memory, and our results are entirely consistent with a large body of literature showing the involvement of ventral stream perceptual regions in visual imagery (e.g., Albers et al., 2013; Cichy et al., 2012; Dijkstra et al., 2017; Johnson and Johnson, 2014; Lee et al., 2012a,b; Reddy et al., 2010; Stokes et al., 2011).

During the imagination phase, which did not contain markedly different visual perceptual content across conditions (i.e., the words "combine" vs. "compare"), both the ROI and whole-volume analyses revealed more widespread activity for the mental creation of scenes compared with the mental comparison of objects. Specifically, mentally

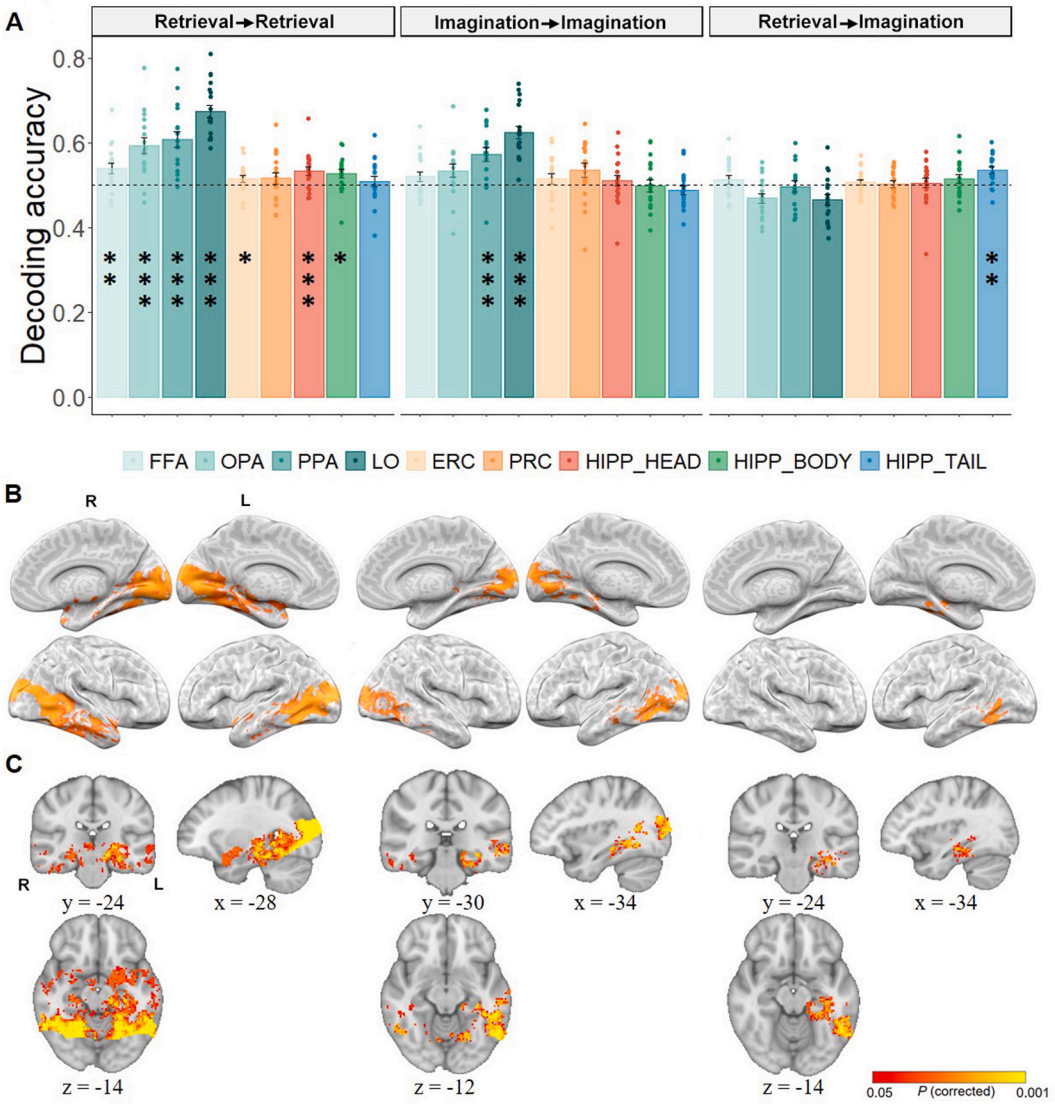


Fig. 6. Multivariate classification results for the 2-way classification of objects versus scenes. **A**. Mean decoding accuracy within each ROI for the retrieval phase (left), imagination phase (middle), and cross-phase decoding (i.e., train retrieval, test imagination) (right). The dashed line indicates chance-level performance (0.5). Error bars represent the standard error of the mean (SEM) across participants. FDR corrected: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. **B-C.** Results of the searchlight-based MVPA with warm colors corresponding to regions showing significantly above chance decoding accuracy (TFCE, *P(corrected)* < 0.05) across the whole volume for the retrieval phase (left), imagination phase (middle), and cross-phase decoding (right). For visualization purposes in B, the maps have been projected onto an inflated surface mesh as implemented in NeuroElf v1.1 (Colin Holmes' 27-scan average brain image, https://neuroelf.net/). In C, the maps have been rendered on the MNI152 template, with select coronal, sagittal, and transverse slices being shown. R: right; L: left.

combining discrete wall segments elicited greater activity in regions classically associated with the perceptual processing of scenes (from the ROI analysis: LO, PPA, and OPA; Dilks et al., 2013; Epstein and Kanwisher, 1998; MacEvoy and Epstein, 2011; from the whole-volume analysis: the lingual gyrus; Aguirre et al., 1998) as well as the tail region of the HPC, a structure that has traditionally been associated with scene memory but has also been suggested to play a critical role in complex scene perception (e.g., Lee et al., 2013; Lee et al., 2012a,b). Interestingly, recent work has demonstrated that the PPA is split into posterior and anterior segments, with the former involved in the visual perception of currently viewed scene features and the latter implicated in nonvisual tasks that move beyond the perception of currently visible information, such as scene memory and navigation (Baldassano et al., 2016). Moreover, functional connectivity analyses revealed that the anterior, but not posterior, PPA is strongly connected to the hippocampus. With this in mind, our results suggest that regions of the human scene-processing network are not only involved in the visuoperceptual processing of scene features and scene memory-guided behavior such as spatial navigation (for review, see Epstein and Baker, 2019), but also in the imagination of novel scene content and the processing of their associated underlying neural representations. Future studies could utilize our novel scene creation paradigm along with functional connectivity analyses to examine if the current results are more attributable to neural processing in anterior, compared to posterior, PPA, as well as investigate the contribution of another important node in the human scene-processing network, the retrosplenial complex (RSC; Epstein, 2008).

In contrast to the widespread activity observed during Scene Creation, only the PPA (ROI analysis) and occipital pole (whole-volume analysis) were significantly active when mentally comparing objects versus creating scenes. The interpretation of these results is less obvious than the Scene Creation findings, and potentially speaks to an imbalance in how our stimuli competed for neural resources. Specifically, during the retrieval phase more widespread activity was observed during object

Table 5
Searchlight-based decoding for 2-way classification (object vs scene).

Train → Test	Region (Left/Right)	Peak Voxel			*P* value	Cluster Size
		x	y	z		
Retrieval → Retrieval	Temporal Occipital Fusiform Cortex (L)	−44	−54	−18	<0.001	34087
	Inferior Temporal Gyrus (R)	52	−52	−18	<0.001	
	Parahippocampal Cortex (L)	−32	−40	−2	<0.001	
	Middle Temporal Gyrus (R)	68	−42	−2	<0.001	
	Occipital Fusiform Gyrus (R)	36	−62	−16	<0.001	
	Inferior Lateral Occipital Cortex (L)	−46	−84	−2	<0.001	
	Inferior Lateral Occipital Cortex (R)	28	−84	−2	<0.001	
	Occipital Pole (R)	16	−94	2	<0.001	
	Thalamus, Hippocampus (R)	16	−32	0	<0.001	
Imagination → Imagination	Occipital Fusiform Gyrus (L)	−34	−60	−4	<0.001	16601
	Lingual Gyrus (R)	8	−86	−2	<0.001	
	Occipital Pole (R)	14	−90	0	<0.001	
	Inferior Lateral Occipital Cortex (L)	−28	−88	0	<0.001	
	Middle Temporal Gyrus (L)	−58	−46	−12	<0.001	
	Inferior Temporal Gyrus (L)	−50	−56	−14	<0.001	
	Inferior Lateral Occipital Cortex (L)	−58	−64	−14	<0.001	
Retrieval → Imagination	Inferior Temporal Gyrus (L)	−46	−52	−8	<0.001	3569
	Inferior Lateral Occipital Cortex (L)	−54	−64	−12	<0.001	
	Posterior Inferior Temporal Gyrus (L)	−48	−42	−10	<0.001	
	Posterior Temporal Fusiform Cortex (L)	−36	−26	−14	<0.001	
	Parahippocampal Cortex (L)	−28	−36	−14	<0.001	
	Posterior Temporal Fusiform Cortex, Hippocampus (L)	−36	−34	−14	<0.001	

compared to scene retrieval, whereas during the imagination phase Scene Creation was associated with a greater activity compared with Object Comparison. This pattern of results suggests a potential dominance (i.e., more widespread activity) of scene-related processing, since object retrieval involved the perception of an on-screen scene element (i.e., a wall) and reflecting this, some of the associated activation occurred within known scene-sensitive regions (i.e., PPA, OPA). Moreover, even though no stimulus was present on-screen during the imagination phase, creating scenes yielded more widespread activation than comparing objects. Stimulus-based explanations for these results could relate to, at least in part, an imbalance in the number of 'space defining' (e.g., a stove) versus 'space ambiguous' (e.g., a vase) objects utilized across experimental conditions (Mullally and Maguire, 2011), or differences in the ease of imageability of object and scene features across our stimulus set. Alternatively, this difference in activation in the imagination phase may not relate to scene versus object processing per se, but rather to task-related differences in the processes of combining versus comparing mental representations. One possibility is that the Scene Creation task was more cognitively demanding than Object Comparison although notably, the data on this are equivocal and suggest that such an explanation is overly simplistic. Specifically, while some of our behavioral indices suggest greater effort was involved in Scene Creation versus Object Comparison (i.e., longer RTs and lower confidence ratings for the former versus the latter process), others suggest equivalent performance across tasks (i.e., no overall difference in accuracy for Scene Creation versus Object Comparison in the Q&A phase; but note that higher accuracy was observed for Scene Creation when examining only trials with high confidence ratings – see Supplemental Material). In short, further research is needed to investigate how different experimental tasks (e.g., comparing vs. combining) impact the neural correlates of imagination and in doing so, this will bring more clarity to our understanding of the neural mechanisms underlying the imaginative processing of scenes versus objects.

Manipulating spatial distance only had a subtle effect on behavioral performance and no observable impact on neural activity. Behaviorally, there were only trend level effects when comparing response times and accuracy for trials in which participants imagined walls or objects that were adjacent to each other in one of the previously learned rooms (Same Attached), with those for trials in which participants imagined two non-connecting walls or non-neighbouring objects taken from the same room (Same Detached) or different rooms (Different). Neurally, there were no significant differences in univariate activity between these conditions and moreover, it was not possible to differentiate the multivariate patterns of activity associated with them using a classifier. The absence of a significant effect of spatial distance is somewhat surprising given that participants retrieved previously experienced content during the Same Attached condition, particularly during Scene Creation (i.e., an existing wall configuration), whereas they generated novel content during the Same Detached and Different conditions (i.e., a novel wall recombination). Although we cannot rule out the possibility that this null finding is the product of a methodological limitation in our study, we suggest that it may reflect the nature of the neural substrates underlying the mental visualization of existing and newly generated spatial scenes. That is, similar neural mechanisms underlie both, at least within the timescale of the current behavioral paradigm, converging with the notion that overlapping core brain regions support both the retrieval of past episodic information as well as the imagination of future and hypothetical events (e.g., Schacter and Addis, 2007; Schacter et al., 2017), perhaps with differences in the direction of information flow between areas (e.g., Byrne et al., 2007). The lack of a difference between spatial distance conditions, in particular with regards to hippocampal involvement, is also not inconsistent with the idea that the hippocampal-dependent process of scene construction is central to memory retrieval and imagination (Mullally and Maguire, 2014; Zeidman and Maguire, 2016), although it is important to note the current data are unable to arbitrate between different theoretical viewpoints regarding how individual regions contribute to retrieval and imagination.

Supportive of the suggestion that overlapping neural substrates underlie both the retrieval of previously experienced content and the imagination of novel content, a 2-way classifier trained on the retrieval data was able to successfully differentiate patterns of activity associated with the imagination of scenes from that of objects. While a targeted ROI approach (in which all voxels were considered within each region) only revealed significant classification accuracy in the hippocampal tail, a more localized searchlight approach identified informative voxels throughout the left posterior hippocampus in the body and tail regions, as well as the parahippocampal cortex including the PPA, and regions in the lateral temporal lobe. The observation of predominantly posterior hippocampus involvement converges with previous work that demonstrated greater activity in the posterior hippocampus when participants constructed a newly imagined event containing details across disparate

memories relative to an event containing details previously simulated (Gaesser et al., 2013), and aligns with the notion of a gradient in representational detail along the longitudinal axis of the hippocampus (Poppenk et al., 2013). Specifically, it has been proposed that while the anterior portion of the hippocampus subserves coarse or abstract representations, the posterior portion is important for processing representations with more fine-grained detail (Poppenk et al., 2013). According to this viewpoint, the former may be critical for inferential reasoning and generalizing across previously learnt information (e.g., Preston et al., 2004; Schlichting et al., 2015) while the latter may be recruited when fine details are required, for instance when reconstructing previous memories (e.g., Sheldon and Levine, 2016) or constructing novel scenarios such as in the Same Detached and Different conditions in the present study. This is consistent with recent findings that view the hippocampus as a heterogenous structure, wherein different task demands and stimulus configurations dictate the pattern of sub-region activity observed (Dalton et al., 2018). Pertinent to the results of this study, these recent findings demonstrate that different regions of the hippocampus were engaged when mental imagery was used for scene construction or object imagination (Dalton et al., 2018). Our results, along with these findings, underscore the importance of using a variety of stimulus types and tasks to interrogate the functional organization of hippocampal subregions with respect to theories of scene construction (Zeidman and Maguire, 2016) versus theories of associative processing such as the constructive episodic simulation hypothesis (Roberts et al., 2018).

The scene stimuli in our study were designed to promote the processing of spatial/geometric features (e.g., shape of protrusion, location of protrusion, and convexity/concavity of protrusion on wall segments), an approach that is consistent with studies of scene perception that led to the development of the spatial layout hypothesis, which argues that scene representation in PPA is dominated by spatial features (for review, see Epstein, 2008). However, recent evidence shows that scene regions in both the human (particularly PPA) and monkey brain (the lateral and medial place patches, putative homologues of the human PPA) not only process spatial geometry, but also nonspatial visual features such as texture (Cant and Xu, 2012, 2017; Kornblith et al., 2013; Lowe et al., 2017; Park and Park, 2017), and the importance of each feature in the human brain varies according to perceived scene category (i.e., geometry contributes more to the representation of manufactured scenes, such as indoor rooms or cityscapes, whereas geometry and texture equally contribute to the representation of natural scenes, such as caves or deserts; Lowe et al., 2016). Moreover, previous research has demonstrated that attention to the texture and material properties of objects differentially activates PPA, whereas attention to the shape of the same objects activates LO (Cant et al., 2009; Cant and Goodale, 2007, 2011). Interestingly, the Object Comparison task in our study utilized both spatial/geometric (i.e., height, width) and material property (weight) object features, which differs from the Scene Creation task that only utilized spatial features. In light of this, it would be interesting to examine imaginative processing further by manipulating both geometric and textural/material features of objects and scenes, in both manufactured and natural environments.

Although our study focused on the MTL and extrastriate regions, it is important to acknowledge that other areas have also been implicated in imaginative processing. For instance, recent work has demonstrated that the ventromedial prefrontal cortex (vmPFC) is implicated in both semantic and spatially constructive scene processing (McCormick and Maguire, 2021), as a node that connects to lateral temporal cortex to support semantic scene processing and the hippocampus to support spatial scene construction. Our scanning protocol (i.e., slices oriented parallel to the long axis of the hippocampus) did not provide full coverage of the frontal lobes and as such, our current data cannot speak to the involvement of vmPFC. In our paradigm, we also did not use prompts that would encourage semantic processing of scene features, although our object comparison condition did include such prompts (i.e., comparisons of object weight). It is possible, therefore, that some of the brain activity observed along a lateral-to-medial gradient in extrastriate and MTL regions is explained by semantic versus spatial processing of our object and scene stimuli, but a future study designed to directly test this hypothesis would be necessary to provide more definitive data. Moreover, by using a paradigm that includes semantic probes after both scene creation and object comparison in conjunction with whole brain data acquisition, we could also examine the degree of separability in the vmPFC for mental processes involved in spatial versus semantic aspects of scene and object processing.

Finally, as a methodological point, our study highlights the strengths of adopting multiple approaches when analyzing fMRI data. We observed multiple incidences of converging findings across the different analyses we implemented, including ROI and whole volume univariate contrasts, and ROI and whole volume searchlight multivariate classification. There were, however, also results that were unique to a single analysis, therefore underlining the complementary nature of our different analyses. For instance, the multivariate whole volume searchlight analysis, which assigned a classification accuracy to each voxel in the data volume, revealed greater involvement of several brain regions during imagination compared to retrieval, which were not evident from the broader multivariate ROI approach that assigned a single classification accuracy on the basis of all voxels within a predefined functional/anatomical region. Future studies should extend this approach by utilizing additional tools, such as functional connectivity analysis, to illuminate the neural mechanisms subserving imaginative processing.

To summarize, our study utilized a novel experimental task to explore the neural mechanisms mediating the imaginative processing of object and scene stimuli. More extensive neural activity was observed when retrieving object stimuli from memory (cued by the perception of a scene segment shown on screen), and when using imagination to mentally create scenes. This may reflect a dominance of scene compared with object processing, evidenced by more widespread neural activity observed for the former, or alternatively, differences in cognitive demand between the scene and object conditions, for instance differences in task difficulty, the balance of 'space defining' versus 'space ambiguous' objects across conditions, the ease of imageability of different scene and object features, perception vs. mnemonic processes as the source of activation in cued retrieval, and the contribution of different types of tasks and visual features in imaginative processing. Importantly, our study provides a methodological advance to the field, as we developed a novel behavioral paradigm to systematically evaluate the contents of an individual's imagination, rather than relying on participants' subjective responses. Using this paradigm, we observed consistencies across our behavioral and neural results (e.g., very little effect of spatial distance when using imagination to create scenes or compare objects), and our multi-pronged analytical approach (i.e., using ROI and whole-volume analyses with both univariate and multivariate methods) generated a more comprehensive picture of the neural correlates of imaginative processing than would have been garnered using only one or two types of analyses. Taken together, this demonstrates the utility of our novel imagination paradigm and analytical approach, and shows that the contents of an individual's imagination can be more systematically evaluated in studies moving forward.

CRediT authorship contribution statement

Qun Ye: Writing – review & editing, Writing – original draft, Visualization, Formal analysis. **Celia Fidalgo:** Methodology, Investigation, Formal analysis. **Patrick Byrne:** Methodology, Conceptualization. **Luis Eduardo Muñoz:** Investigation. **Jonathan S. Cant:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Andy C.H. Lee:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Data availability

Data will be made available on request.

Acknowledgements

We thank all participants for their time and staff at the MRI Facility of York University for their help with data collection. This research was funded by Discovery Grants from the Natural Sciences and Engineering Research Council of Canada to JSC and ACHL.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuropsychologia.2024.109000.

References

Addis, D.R., Wong, A.T., Schacter, D.L., 2007. Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. Neuropsychologia 45 (7), 1363–1377. https://doi.org/10.1016/j.neuropsychologia.2006.10.016.

Aguirre, G.K., Zarahn, E., D'Esposito, M., 1998. An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. Neuron 21 (2), 373–383. https://doi.org/10.1016/S0896-6273(00)80546-2.

Albers, A.M., Kok, P., Toni, I., Dijkerman, H.C., De Lange, F.P., 2013. Shared representations for working memory and mental imagery in early visual cortex. Curr. Biol. 23 (15), 1427–1431.

Allen, M., Poggiali, D., Whitaker, K., Marshall, T.R., van Langen, J., Kievit, R.A., 2021. Raincloud plots: a multi-platform tool for robust data visualization. Wellcome Open Res. 4 (63) https://doi.org/10.12688/wellcomeopenres.15191.2 eCollection 2019.

Aminoff, E.M., Kveraga, K., Bar, M., 2013. The role of the parahippocampal cortex in cognition. Trends Cogn Sci 17 (8), 379–390. https://doi.org/10.1016/j.tics.2013.06.009.

Bainbridge, W.A., Hall, E.H., Baker, C.I., 2021. Distinct representational structure and localization for visual encoding and recall during visual imagery. Cereb Cortex 31 (4), 1898–1913. https://doi.org/10.1093/cercor/bhaa329.

Baldassano, C., Esteva, A., Fei-Fei, L., Beck, D.M., 2016. Two distinct scene-processing networks connecting vision and memory. eNeuro 3 (5). https://doi.org/10.1523/ENEURO.0178-16.2016.

Beckmann, C.F., Smith, S.M., 2004. Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Trans Med Imaging 23 (2), 137–152. https://doi.org/10.1109/TMI.2003.822821.

Benoit, R.G., Schacter, D.L., 2015. Specifying the core network supporting episodic simulation and episodic memory by activation likelihood estimation. Neuropsychologia 75, 450–457. https://doi.org/10.1016/j.neuropsychologia.2015.06.034.

Brainard, D.H., 1997. The Psychophysics toolbox. Spat Vis 10 (4), 433–436. https://www.ncbi.nlm.nih.gov/pubmed/9176952.

Byrne, P., Becker, S., Burgess, N., 2007. Remembering the past and imagining the future: a neural model of spatial memory and imagery. Psychol. Rev. 114 (2), 340–375. https://doi.org/10.1037/0033-295X.114.2.340.

Cant, J.S., Arnott, S.R., Goodale, M.A., 2009. fMR-adaptation reveals separate processing regions for the perception of form and texture in the human ventral stream. Exp. Brain Res. 192 (3), 391–405. https://doi.org/10.1007/s00221-008-1573-8.

Cant, J.S., Goodale, M.A., 2007. Attention to form or surface properties modulates different regions of human occipitotemporal cortex. Cereb Cortex 17 (3), 713–731. https://doi.org/10.1093/cercor/bhk022.

Cant, J.S., Goodale, M.A., 2011. Scratching beneath the surface: new insights into the functional properties of the lateral occipital area and parahippocampal place area. J. Neurosci. 31 (22), 8248–8258. https://doi.org/10.1523/JNEUROSCI.6113-10.2011.

Cant, J.S., Xu, Y., 2012. Object ensemble processing in human anterior-medial ventral visual cortex. J. Neurosci. 32 (22), 7685–7700. https://doi.org/10.1523/JNEUROSCI.3325-11.2012.

Cant, J.S., Xu, Y., 2017. The contribution of object shape and surface properties to object ensemble representation in anterior-medial ventral visual cortex. J Cogn Neurosci 29 (2), 398–412. https://doi.org/10.1162/jocn_a_01050.

Cichy, R.M., Heinzle, J., Haynes, J.D., 2012. Imagery and perception share cortical representations of content and location. Cerebr. Cortex 22 (2), 372–380.

Clarke, A., Tyler, L.K., 2014. Object-specific semantic coding in human perirhinal cortex. J. Neurosci. 34 (14), 4766–4775. https://doi.org/10.1523/JNEUROSCI.2828-13.2014.

Dalton, M.A., Zeidman, P., McCormick, C., Maguire, E.A., 2018. Differentiable processing of objects, associations, and scenes within the hippocampus. J. Neurosci. 38 (38), 8146–8159.

Dijkstra, N., Zeidman, P., Ondobaka, S., van Gerven, M.A., Friston, K., 2017. Distinct top-down and bottom-up brain connectivity during visual perception and imagery. Sci. Rep. 7 (1), 5677.

Dilks, D.D., Julian, J.B., Paunov, A.M., Kanwisher, N., 2013. The occipital place area is causally and selectively involved in scene perception. J. Neurosci. 33 (4), 1331–1336a. https://doi.org/10.1523/JNEUROSCI.4081-12.2013.

Epstein, R., Kanwisher, N., 1998. A cortical representation of the local visual environment. Nature 392 (6676), 598–601. https://doi.org/10.1038/33402.

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

Epstein, R.A., Baker, C.I., 2019. Scene perception in the human brain. Annu Rev Vis Sci 5, 373–397. https://doi.org/10.1146/annurev-vision-091718-014809.

Epstein, R.A., Morgan, L.K., 2012. Neural responses to visual scenes reveals inconsistencies between fMRI adaptation and multivoxel pattern analysis. Neuropsychologia 50 (4), 530–543.

Gaesser, B., Spreng, R.N., McLelland, V.C., Addis, D.R., Schacter, D.L., 2013. Imagining the future: evidence for a hippocampal contribution to constructive processing. Hippocampus 23 (12), 1150–1161. https://doi.org/10.1002/hipo.22152.

Grande, X., Sauvage, M.M., Becke, A., Düzel, E., Berron, D., 2022. Transversal functional connectivity and scene-specific processing in the human entorhinal-hippocampal circuitry. Elife 11, e76479.

Grill-Spector, K., Kourtzi, Z., Kanwisher, N., 2001. The lateral occipital complex and its role in object recognition. Vision Res 41 (10–11), 1409–1422. https://doi.org/10.1016/s0042-6989(01)00073-6.

Hassabis, D., Kumaran, D., Maguire, E.A., 2007a. Using imagination to understand the neural basis of episodic memory. J. Neurosci. 27 (52), 14365–14374. https://doi.org/10.1523/JNEUROSCI.4549-07.2007.

Hassabis, D., Kumaran, D., Vann, S.D., Maguire, E.A., 2007b. Patients with hippocampal amnesia cannot imagine new experiences. Proc Natl Acad Sci U S A 104 (5), 1726–1731. https://doi.org/10.1073/pnas.0610561104.

Hassabis, D., Maguire, E.A., 2009. The construction system of the brain. Philos. Trans. R. Soc. Lond. B Biol. Sci. 364 (1521), 1263–1271. https://doi.org/10.1098/rstb.2008.0296.

Iglesias, J.E., Augustinack, J.C., Nguyen, K., Player, C.M., Player, A., Wright, M., Roy, N., Frosch, M.P., McKee, A.C., Wald, L.L., Fischl, B., Van Leemput, K., Alzheimer's Disease Neuroimaging, I., 2015. A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: application to adaptive segmentation of in vivo MRI. Neuroimage 115, 117–137. https://doi.org/10.1016/j.neuroimage.2015.04.042.

Insausti, R., Juottonen, K., Soininen, H., Insausti, A.M., Partanen, K., Vainio, P., Laakso, M.P., Pitkanen, A., 1998. MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. AJNR Am J Neuroradiol 19 (4), 659–671. https://www.ncbi.nlm.nih.gov/pubmed/9576651.

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17 (2), 825–841. https://doi.org/10.1016/s1053-8119(02)91132-8.

Johnson, M.R., Johnson, M.K., 2014. Decoding individual natural scene representations during perception and imagery. Front. Neurosci. 8, 59.

Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17 (11), 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997.

Kay, K.N., Rokem, A., Winawer, J., Dougherty, R.F., Wandell, B.A., 2013. GLMdenoise: a fast, automated technique for denoising task-based fMRI data. Front. Neurosci. 7, 247. https://doi.org/10.3389/fnins.2013.00247.

Kornblith, S., Cheng, X., Ohayon, S., Tsao, D.Y., 2013. A network for scene processing in the macaque temporal lobe. Neuron 79 (4), 766–781. https://doi.org/10.1016/j.neuron.2013.06.015.

Kourtzi, Z., Kanwisher, N., 2001. Representation of perceived object shape by the human lateral occipital complex. Science 293 (5534), 1506–1509. https://doi.org/10.1126/science.1061133.

Kriegeskorte, N., Goebel, R., Bandettini, P., 2006. Information-based functional brain mapping. Proc Natl Acad Sci U S A 103 (10), 3863–3868. https://doi.org/10.1073/pnas.0600244103.

Lee, A.C.H., Brodersen, K.H., Rudebeck, S.R., 2013. Disentangling spatial perception and spatial memory in the hippocampus: a univariate and multivariate pattern analysis fMRI study. J Cogn Neurosci 25 (4), 534–546. https://doi.org/10.1162/jocn_a_00301.

Lee, A.C.H., Yeung, L.K., Barense, M.D., 2012a. The hippocampus and visual perception. Front. Hum. Neurosci. 6, 91. https://doi.org/10.3389/fnhum.2012.00091.

Lee, S.H., Kravitz, D.J., Baker, C.I., 2012b. Disentangling visual imagery and perception of real-world objects. Neuroimage 59 (4), 4064–4073.

Lowe, M.X., Gallivan, J.P., Ferber, S., Cant, J.S., 2016. Feature diagnosticity and task context shape activity in human scene-selective cortex. Neuroimage 125, 681–692. https://doi.org/10.1016/j.neuroimage.2015.10.089.

Lowe, M.X., Rajsic, J., Gallivan, J.P., Ferber, S., Cant, J.S., 2017. Neural representation of geometry and surface properties in object and scene perception. Neuroimage 157, 586–597. https://doi.org/10.1016/j.neuroimage.2017.06.043.

MacEvoy, S.P., Epstein, R.A., 2011. Constructing scenes from objects in human occipitotemporal cortex. Nat. Neurosci. 14 (10), 1323–1329. https://doi.org/10.1038/nn.2903.

Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc Natl Acad Sci U S A 92 (18), 8135–8139. https://doi.org/10.1073/pnas.92.18.8135.

Martin, C.B., Douglas, D., Newsome, R.N., Man, L.L., Barense, M.D., 2018. Integrative and distinctive coding of visual and conceptual object features in the ventral visual stream. Elife 7, e31873. https://doi.org/10.7554/eLife.31873.

McCormick, C., Maguire, E.A., 2021. The distinct and overlapping brain networks supporting semantic and spatial constructive scene processing. Neuropsychologia 158, 107912.
Misaki, M., Kim, Y., Bandettini, P.A., Kriegeskorte, N., 2010. Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. Neuroimage 53 (1), 103–118. https://doi.org/10.1016/j.neuroimage.2010.05.051.
Mullally, S.L., Maguire, E.A., 2011. A new role for the parahippocampal cortex in representing space. J. Neurosci. 31 (20), 7441–7449.
Mullally, S.L., Maguire, E.A., 2014. Memory, imagination, and predicting the future: a common brain mechanism? Neuroscientist 20 (3), 220–234. https://doi.org/10.1177/1073858413495091.
Mumford, J.A., Turner, B.O., Ashby, F.G., Poldrack, R.A., 2012. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. Neuroimage 59 (3), 2636–2643. https://doi.org/10.1016/j.neuroimage.2011.08.076.
Oosterhof, N.N., Connolly, A.C., Haxby, J.V., 2016. CoSMoMVPA: multi-modal multivariate pattern analysis of neuroimaging data in matlab/GNU octave. Front Neuroinform 10, 27. https://doi.org/10.3389/fninf.2016.00027.
Park, J., Park, S., 2017. Conjoint representation of texture ensemble and location in the parahippocampal place area. J. Neurophysiol. 117 (4), 1595–1607. https://doi.org/10.1152/jn.00338.2016.
Poppenk, J., Evensmoen, H.R., Moscovitch, M., Nadel, L., 2013. Long-axis specialization of the human hippocampus. Trends Cogn Sci 17 (5), 230–240. https://doi.org/10.1016/j.tics.2013.03.005.
Preston, A.R., Shrager, Y., Dudukovic, N.M., Gabrieli, J.D., 2004. Hippocampal contribution to the novel use of relational information in declarative memory. Hippocampus 14 (2), 148–152. https://doi.org/10.1002/hipo.20009.
Ragni, F., Lingnau, A., Turella, L., 2021. Decoding category and familiarity information during visual imagery. Neuroimage 241, 118428. https://doi.org/10.1016/j.neuroimage.2021.118428.
Ramanan, S., Alaeddin, S., Goldberg, Z.L., Strikwerda-Brown, C., Hodges, J.R., Irish, M., 2018. Exploring the contribution of visual imagery to scene construction - evidence from Posterior Cortical Atrophy. Cortex 106, 261–274. https://doi.org/10.1016/j.cortex.2018.06.016.
Reddy, L., Tsuchiya, N., Serre, T., 2010. Reading the mind's eye: decoding category information during mental imagery. Neuroimage 50 (2), 818–825.
Roberts, R.P., Schacter, D.L., Addis, D.R., 2018. Scene construction and relational processing: separable constructs? Cerebr. Cortex 28 (5), 1729–1732.
Schacter, D.L., Addis, D.R., 2007. The cognitive neuroscience of constructive memory: remembering the past and imagining the future. Philos. Trans. R. Soc. Lond. B Biol. Sci. 362 (1481), 773–786. https://doi.org/10.1098/rstb.2007.2087.
Schacter, D.L., Addis, D.R., Hassabis, D., Martin, V.C., Spreng, R.N., Szpunar, K.K., 2012. The future of memory: remembering, imagining, and the brain. Neuron 76 (4), 677–694. https://doi.org/10.1016/j.neuron.2012.11.001.
Schacter, D.L., Benoit, R.G., Szpunar, K.K., 2017. Episodic future thinking: mechanisms and functions. Curr Opin Behav Sci 17, 41–50. https://doi.org/10.1016/j.cobeha.2017.06.002.
Schlichting, M.L., Mumford, J.A., Preston, A.R., 2015. Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. Nat. Commun. 6, 8151. https://doi.org/10.1038/ncomms9151.
Sheldon, S., Levine, B., 2016. The role of the hippocampus in memory and mental construction. Ann. N. Y. Acad. Sci. 1369 (1), 76–92. https://doi.org/10.1111/nyas.13006.
Smith, S.M., 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17 (3), 143–155. https://doi.org/10.1002/hbm.10062.
Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23 (Suppl. 1), S208–S219. https://doi.org/10.1016/j.neuroimage.2004.07.051.
Smith, S.M., Nichols, T.E., 2009. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage 44 (1), 83–98. https://doi.org/10.1016/j.neuroimage.2008.03.061.
Staresina, B.P., Duncan, K.D., Davachi, L., 2011. Perirhinal and parahippocampal cortices differentially contribute to later recollection of object- and scene-related event details. J. Neurosci. 31 (24), 8739–8747. https://doi.org/10.1523/JNEUROSCI.4978-10.2011.
Stokes, M., Saraiva, A., Rohenkohl, G., Nobre, A.C., 2011. Imagery for shapes activates position-invariant representations in human visual cortex. Neuroimage 56 (3), 1540–1545.
Szpunar, K.K., St Jacques, P.L., Robbins, C.A., Wig, G.S., Schacter, D.L., 2014. Repetition-related reductions in neural activity reveal component processes of mental simulation. Soc Cogn Affect Neurosci 9 (5), 712–722. https://doi.org/10.1093/scan/nst035.
Zeidman, P., Maguire, E.A., 2016. Anterior hippocampus: the anatomy of perception, imagination and episodic memory. Nat. Rev. Neurosci. 17 (3), 173–182. https://doi.org/10.1038/nrn.2015.24.
Zeidman, P., Mullally, S.L., Maguire, E.A., 2015. Constructing, perceiving, and maintaining scenes: hippocampal activity and connectivity. Cereb Cortex 25 (10), 3836–3855. https://doi.org/10.1093/cercor/bhu266.