Perception of Impossible Scenes Reveals Differential Hippocampal and Parahippocampal Place Area Contributions to Spatial Coherency

Danielle Douglas,[1]* Sathesan Thavabalasingam,[1] Zahraa Chorghay,[1] Edward B. O'Neil,[1] Morgan D. Barense,[1,2] and Andy C. H. Lee[1,2]*

ABSTRACT: Surprisingly little is known about how the brain combines spatial elements to form a coherent percept. Regions that may underlie this process include the hippocampus (HC) and parahippocampal place area (PPA), regions central to spatial perception but whose role in spatial coherency has not been explored. Participants were scanned with functional MRI while they judged whether Escher-like scenes were possible or impossible. Univariate analyses revealed differential HC and PPA involvement, with greater HC activity during spatial incoherency detection and more PPA activity during spatial coherency detection. Recognition and eye-tracking data ruled out long- or short-term memory confounds. Multivariate statistics demonstrated spatial coherency-dependent functional connectivity for the HC, but not PPA, with greater HC connectivity to various brain regions including lateral occipital complex during spatial incoherency detection. We suggest the PPA is preferentially involved during the perception of spatially coherent scenes, whereas the HC binds distinct features to create coherent representations. © 2016 Wiley Periodicals, Inc.

KEY WORDS: hippocampus; medial temporal lobe; memory; perception; fMRI

INTRODUCTION

Perceiving a coherent three-dimensional environment is central to our ability to interact with the world around us. For example, navigating a university hall likely relies on the harmonious combination of the disparate components that compose the building, including the different floor levels, staircases, walls, and corridors. Surprisingly, despite the importance of this cognitive process to everyday behavior, relatively little is known about how the brain combines features to give rise to a single, spatially coherent scene representation.

Based on their importance to spatial cognition, candidate brain regions for supporting spatial coherency include the parahippocampal place area (PPA) and hippocampus (HC). Investigations of the representational content of these two regions suggest that they each contribute to spatial processing in qualitatively different ways. The PPA is thought to represent the basic geometry and content of a scene, with these representations being intolerant to low-level shifts in information (Epstein, 2008), although this may not always be the case (see Marchette et al., 2015 for evidence for sensitivity to scene identity in the PPA). For instance, the PPA responds preferentially to whole rather than fragmented scene images (Epstein and Kanwisher, 1998), and PPA activity is known to be sensitive to manipulations of scene identity (Epstein and Higgins, 2007), boundary (Harel et al., 2013), and viewpoint (Epstein et al., 2003; Park and Chun, 2009). Crucially, to our knowledge, there has been little work exploring the role of this region in spatial coherency.

In comparison, the HC may abstract across low-level information and represent elements of a scene along with their particular configuration (Bird and Burgess, 2008). Much work on the role of the HC in spatial cognition has focused on the notion that this structure subserves a cognitive map of the external environment (O'Keefe and Nadel, 1978). Studies have reported HC cells that signal specific spatial locations (O'Keefe and Dostrovsky, 1971; Hori et al., 2003), spatial view information (Rolls et al., 1997; Ekstrom et al., 2003), and boundaries of the environment (Lever et al., 2009). Moreover, neuropsychological and neuroimaging research has implicated the HC in spatial memory tasks (Spiers et al., 2001; Burgess, 2002; King et al., 2002; Maguire et al., 2006) and spatial navigation (Maguire et al., 1997; Spiers and Maguire, 2007; Iglói et al., 2010).

More recently, theoretical accounts of HC function have attempted to account for data from paradigms that do not place an explicit demand on memory (Lee et al., 2012; Maguire and Mullally, 2013; Yonelinas, 2013). Patients with HC lesions describe

[1] **Psychology (Scarborough, St George), University of Toronto, Toronto, Ontario, Canada;** [2] **Rotman Research Institute, Baycrest Centre for Geriatric Care, Toronto, Ontario, Canada**

Additional Supporting Information may be found in the online version of this article.

Grant sponsors: Natural Sciences and Engineering Research Council of Canada (D.D., Z.C., E.O., M.D.B., A.C.H.L.), University of Toronto Excellence Award in Natural Sciences and Engineering (S.T.), and a James S McDonnell Scholar Award (M.D.B.).

*Correspondence to: Danielle Douglas; Department of Psychology, University of Toronto Scarborough, 1265 Military Trail, Toronto, ON M1C 1A4, Canada. E-mail: d.douglas@mail.utoronto.ca or Andy C. H. Lee; Department of Psychology, University of Toronto Scarborough, 1265 Military Trail, Toronto, ON M1C 1A4, Canada. E-mail: andych.lee@utoronto.ca

Accepted for publication 20 October 2016.
DOI 10.1002/hipo.22673
Published online 00 Month 2016 in Wiley Online Library (wileyonlinelibrary.com).

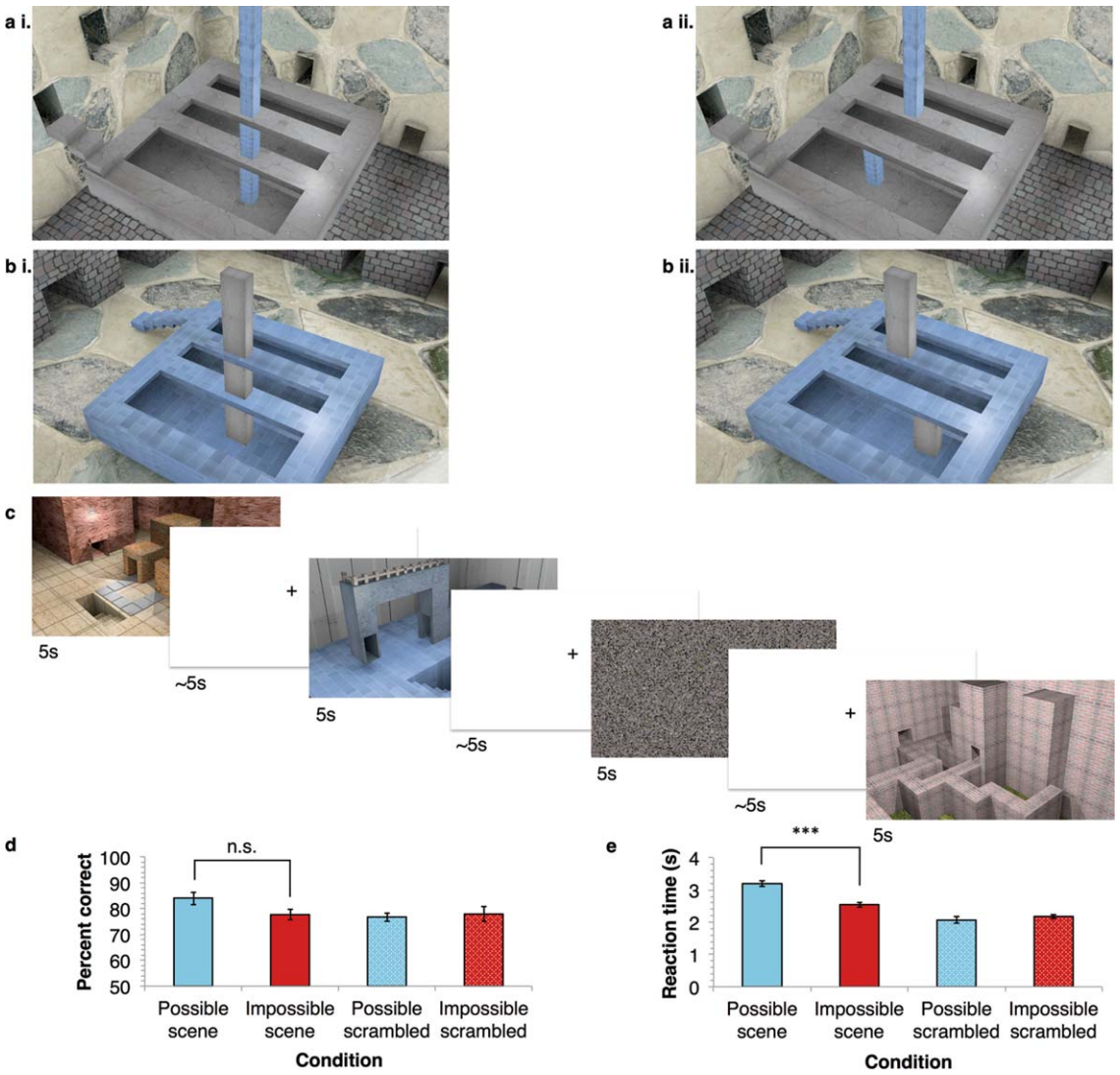

FIGURE 1. fMRI perceptual judgement task and recognition memory stimuli. Examples of (a(i)) impossible scene and (a(ii)) possible scene stimuli, and their respective memory foils (b(i)) and (b(ii)). Note that each participant saw only one version of each scene, either possible or impossible, and its respective foil. (c) Schematic diagram of the task, including scrambled scenes, for which participants were required to determine whether a semitransparent circle was located on the left or the right side of the image. See Supplementary Material for region of interest (ROI) fMRI analyses of *scenes > scrambled scenes*. (d) Accuracy (percent correct) and, (e) reaction time (s) for correct trials. Error bars represent standard error; three asterisks indicate significance at $P < 0.001$. [Color figure can be viewed at wileyonlinelibrary.com]

imagined scenes with less spatial detail than control participants, and demonstrate attenuated scene boundary extension (Hassabis et al., 2007; Mullally et al., 2012), leading to the idea that scene construction could constitute the core function of the HC (Maguire and Mullally, 2013). Moreover, discrimination deficits for scene stimuli have been associated with HC damage (Lee et al., 2005a,b; Aly et al., 2013), with convergent functional neuroimaging data from healthy adults (Lee et al., 2008; Barense et al., 2010; Aly et al., 2013). These findings have led some to suggest, rather controversially, that the HC is critical for spatial perception by way of processing complex spatial representations (Lee et al., 2012). Although the validity of these proposals has been questioned, for instance by suggestions that supporting data are confounded by mnemonic processing (Kim et al., 2011, 2015), they raise the possibility that the HC binds individual spatial elements of a scene together to create a spatially coherent representation.

Although their implication in spatial perception suggests a central role for the HC and PPA in processing spatial coherency, other regions, particularly those involved in spatial navigation, may also be involved. Two such regions are the entorhinal cortex (ERC), which is located on the medial bank of the parahippocampal gyrus (PHG), and retrosplenial cortex (RSC) (Vass and Epstein, 2013). The ERC is the primary

afferent relay between the rest of the brain and the HC, with projections from the ERC providing the majority of input to the dentate gyrus (Amaral and Witter, 1989). In particular, medial ERC represents allocentric spatial locations (Witter and Moser, 2006) via neural representations of environmental coordinates ("grid cells;" Fyhn et al., 2004; Jacobs et al., 2013) and boundaries ("border cells;" Solstad et al., 2008; Lever et al., 2009). The role of the RSC is less well understood but is thought to aid in identifying the global location of immediate environments (Epstein and Higgins, 2007; Park and Chun, 2009), and thus serves as an intermediary between parietal egocentric and HC allocentric representations of space (Aguirre and D'Esposito, 1999; Vann et al., 2009).

Here, we sought to elucidate how brain regions important for perceptual and navigational scene processing contribute to spatial coherency. We examined the neural correlates of spatial coherency using functional magnetic resonance imaging (fMRI) in conjunction with a novel experimental paradigm that was inspired by the impossible reality work of M. C. Escher (1898–1971). Neurologically healthy participants were presented with a series of single, unique virtual reality spatial scenes and were instructed to indicate whether each was "possible" (i.e., spatially coherent) or "impossible" (i.e., spatially incoherent; see Figs. 1a, c). Although paradigms using impossible objects have aided enormously in understanding how the brain represents objects (Schacter et al., 1995; Williams and Tarr, 1997; Freud et al., 2015), to our knowledge, impossible scenes have never been used to investigate scene processing. Therefore, measuring neural response to impossible scenes may provide fundamental information about how scene coherency is achieved. Univariate statistical analyses were employed to examine differences in the magnitude of activity within the PPA and HC, as well as the RSC and PHG, to spatial coherency. We expected that all regions of interest would be sensitive to the differences in spatial coherency between possible and impossible scenes. We expected greater activity to possible versus impossible scenes in the PPA, given previous findings that this region responds less to fragmented scenes. Conversely, we predicted that the HC would show greater activity to impossible compared to possible scenes, given previous findings implicating this region in binding lower-level scene elements into a conjunctive scene representation—a process taxed to a greater degree when assessing the global structure of impossible compared to possible scenes. Multivariate analyses (Partial Least Squares, PLS) were also used to examine changes in functional connectivity across conditions between our regions of interest and the rest of the brain. Particularly, we wished to assess whether upstream visual regions representing lower-level scene features might be differentially connected with the hippocampus while viewing impossible scenes, given previous findings implicating this region in binding lower-level scene elements. Importantly, we administered a surprise follow-up recognition memory task after fMRI scanning to probe whether differences in neural response to spatial coherency could be accounted for by long-term memory confounds (i.e., incidental encoding). Furthermore, a separate eye-tracking study provided insight into the manner in which participants identified impossible scenes, enabling us to address possible differences in scene information load as an alternate explanation for our neural findings.

MATERIALS AND METHODS

Functional Neuroimaging

Participants

Twenty neurologically healthy, right-handed adults were recruited from the University of Toronto community. Each volunteer gave informed written consent to participate in the study and was compensated $40 for their time and travel. Four participants were removed from statistical analyses due to poor behavioral performance (two subjects performed at 50% and 54% on one of the conditions), misunderstanding of task instructions (one subject), and the loss of behavioral response recording during scan acquisition (one subject). The remaining 16 participants (nine female) were between the ages of 20 and 30 ($M = 24.06$; SD = 2.84). This study was approved by the Office of Research Ethics at the University of Toronto (Ref: 26927).

General procedure

FMRI data were acquired while each subject performed two runs of an experimental spatial coherency judgment task intermixed with a baseline target location task, and one run of a functional localizer task. The former was designed to address our theoretical question of interest, while the latter was used to identify functional regions of interest (ROI) for the statistical analysis of the fMRI data (see fMRI analyses). After scanning, a surprise recognition memory test was given outside of the scanner to allow for behavioral measurements of the memorability (e.g., incidental encoding) of the stimuli presented during the experimental possible/impossible judgment task. All tasks were programmed and administered using E-prime software (version 2.0.8.90, Psychology Software Tools).

Stimuli

Fifty-six unique, three dimensional color scenes were created with Blender (version 2.65, www.blender.org/). Each of these scenes was designed to appear spatially incoherent (and therefore "impossible") when viewed from a particular angle, and this viewpoint was captured as a screen-shot image (960 × 540 pixels; see Fig. 1a). A spatially coherent "possible" counterpart was also created for each impossible scene and captured from the same viewpoint in order to minimize any differences between scene conditions, apart from the spatial incoherency. Thus, the impossible and possible versions of each pair were similar in terms of their overall layout, scene elements, and textures, and only differed with respect to the disruption of spatial

coherency in the impossible scene. Both versions of each scene were then spatially rearranged, and textures switched between elements within the scene, to create foil images for the subsequent recognition task (Fig. 1b). The stimulus set was divided into two and counterbalanced across participants such that each participant was never exposed to both the impossible and possible version of any given scene. Additionally, baseline non-scene images (960 × 540 pixels), matched for low-level visual properties, were creating by scrambling each of the scenes using Photoshop (Adobe CS 5.1). A partially transparent circle (20 × 20 pixels) was overlaid at a pseudo-random location on each of the scrambled scenes, such that half of these dots were found on the left and half on the right of the baseline scrambled images. These images were designed to create a baseline task that was similar in difficulty to the main experimental task (see below).

Spatial coherency judgment and baseline task

Two runs were administered per subject. In each run, participants were presented with images in the centre of the screen, one at a time, for 5000 ms, followed by a fixation cross during a jittered inter-stimulus interval (ISI; mean 5000 ms). Each run consisted of 28 experimental possible/impossible judgment trials intermixed with 28 baseline target location trials, with each task indicated by the type of stimulus being presented (scene versus scrambled scene). In the experimental task trials, participants were asked to indicate whether a given scene was possible or impossible in the real world (14 impossible, 14 possible scenes per run). In contrast, the baseline task required participants to determine whether a semi-transparent circle was located on the left or the right side of the image. This task was chosen to encourage participants to scan different portions of the scrambled image as they might while searching for an impossible portion of a scene. Participants indicated their responses via pre-specified buttons on a button box held in their right hand and all stimuli were projected onto a screen situated at the back of the scanner, which could be seen via a mirror mounted on the MRI head coil.

Prior to scanning, all participants were administered a short practice task on a laptop computer (12″ × 7.5″ screen display, 960 × 540 pixel resolution). This practice version consisted of four trials each of possible and impossible judgments, and four trials of the scrambled dot search task. All practice trials used a different set of stimuli to those presented in the scanner.

Functional localizer task

A functional localizer task was administered after the experimental paradigm in order to isolate brain voxels that activate preferentially to scenes (Epstein and Kanwisher, 1998), independently of whether they respond to the experimental task. Participants were shown randomized blocks of sequentially presented gray-scale stimuli: scenes, objects, and faces. Twenty stimuli were presented within each block, each for 300 ms (ISI 450 ms). Four blocks of each stimulus type were presented, and a fixation cross (15,000 ms) was presented between every three blocks. Participants were instructed to press a preassigned button on a hand-held button box whenever an object was presented twice in a row (1-back task), to ensure attendance to stimuli.

Recognition memory task

In order to explore activity related to incidental memory encoding, participants were administered a surprise recognition memory test for the spatial coherency judgment and baseline task stimuli immediately after scanning. Participants were shown color images of possible and impossible scenes, one at a time on a laptop computer (12" × 7.5" screen display, 960 × 540 pixel resolution). Half of these scenes had been presented during scanning, with the other half being new (i.e., 50% targets and 50% foils), for a total of 112 previously presented stimuli (56 scenes, 56 scrambled images) and 112 newly presented stimuli (56 scenes, 56 scrambled images). Participants were instructed to categorize each scene as old or new using pre-specified keys on the laptop keyboard, and were allowed as much time as needed for each trial.

Scanning procedure

Neuroimaging data were acquired with a 3T Signa MRI (GE Medical Systems) at the MRI Unit of the Centre for Addiction and Mental Health Research Imaging Centre, Toronto, Canada. The experimental portion consisted of 378 functional image volumes were collected over two runs using a T2*-weighted Blood Oxygen Level Dependent (BOLD) Spiral In/Out pulse sequence (Glover, 2012) (number of slices = 47, voxel resolution= 3.5 × 3.5 × 3.5 mm, matrix size = 64 × 64, repetition time [TR] = 3000 ms, echo time [TE] = 30 ms, FA = 60°, inter-slice distance = 0 mm). The same Spiral In/Out pulse sequence was also used to collect 86 functional image volumes while participant performed a localizer task (see below). High-resolution whole-brain T1-weighted 3D anatomical images were acquired using a BRAVO sequence comprising 200 slices with acquisition parameters: voxel resolution = 0.9 × 0.9 × 0.9 mm^3, matrix size = 256 × 256 TR = 6.7 ms, TE = 3 ms, FA =8°.

Behavioral analyses

Performance accuracy during fMRI acquisition (percent of trials correct) was analyzed using a two-way repeated measures (RM) ANOVA with within-subjects factors of condition (possible and impossible) and task (spatial coherency judgment, baseline task). Reaction times during fMRI acquisition were entered into a three-way RM ANOVA with within-subjects factors of accuracy (correct, incorrect), condition (possible, impossible) and task (spatial coherency judgment, baseline task). Pairwise comparisons meeting Bonferroni corrected alpha levels were used to investigate effects between conditions within each ANOVA. For the recognition memory task, the proportion of hits (H, correct identification of repetition/old stimuli) and false alarms (FA, incorrect identification of a new stimulus as a

repetition/being old) were calculated for each participant. Signal detection theory was then used to derive a behavioral measure of sensitivity for the recognition memory task d' = Z(H) – Z(FA). All false alarm rates or hits equal to 0 or 1 were adjusted by adding or subtracting half of a trial, respectively (Macmillan and Kaplan, 1985). We investigated any potential differences in sensitivity across conditions with a 2 × 2 × 2 RM ANOVA with within-subject factors of accuracy at encoding (correct, incorrect), task-type (experimental task, baseline task), and spatial coherency (possible, impossible).

Imaging Analyses

Data pre-processing

Brain image data pre-processing and analysis were performed using the FMRI Expert Analysis Tool (FEAT version 5.0.1) part of the FMRIB Software Library (www.fmrib.ox.ac.uk/fsl; Smith et al., 2004). The data for all three EPI runs for each participant (two experimental and one functional localizer run) were processed with the following steps: (1) BET mesh deformation approach for segmentation of brain from non-brain matter (Smith, 2002), (2) MCFLIRT motion correction, in which image volumes were realigned to a central volume with rigid body registration (Jenkinson et al., 2002), (3) spatial smoothing using a Gaussian kernel (full width at half-maximum, 6 mm), (4) normalization of the EPI dataset to grand-mean intensity with a single multiplicative factor, (5) high-pass temporal filtering (50 s) to remove noise due to low frequency scanner drift, and (6) registration of each participant's EPI data to their high-resolution anatomical T1 3D spatial scan and to a standard Montreal Neurological Institute (MNI-152) template image using a combination of linear and non-linear approaches including boundary-based registration (Greve and Fischl, 2009). Thus, all stereotactic coordinates are reported in MNI space (*x*, *y*, *z*).

Univariate fMRI data analysis

After pre-processing, the experimental functional data for each participant (two runs per subject) were submitted to two separate general linear models (GLM) using FEAT. The first was designed to investigate the spatial coherency task collapsed across subsequent memory performance, and a second to investigate subsequent memory performance collapsed across performance on the spatial coherency judgment task. The functional localizer data were analyzed using a third GLM.

Examining neural activity associated with spatial coherency judgments. For each run of each participant, one predictor was convolved with a double-gamma model of the human hemodynamic response function (HRF) for each condition (or explanatory variable, EV). There were 9 EVs: one each for correct trials of each of the four conditions (possible scene, impossible scene, possible scrambled, impossible scrambled), one each for incorrect trials of each of the four conditions, and one EV for fixation. Parameter estimates relating the height of the HRF response to each EV were calculated on a voxel by voxel basis, via a multiple linear regression of the response time-course, creating one beta image for each regressor. Since we were interested in successful scene processing, we focused on correct trials and created beta images for the following contrasts of interest. First, to confirm that each of our regions of interest (ROI) responded to scene stimuli, we investigated whether each of our ROIs showed activity greater to scenes compared to scrambled baseline images: "*(correct impossible scenes + correct possible scenes) > (correct possible scrambled + correct impossible scrambled)*" (see Supplementary Material). Next, to examine our main interest, neural activity particular to impossible compared to possible scenes was probed with the contrasts "*correct impossible > correct possible scenes*," and "*correct possible > correct impossible scenes*." Since accuracy was modelled as part of our GLM, we also conducted a number of contrasts pertaining to incorrect trials. The findings for these are reported in full in the Supplementary Material, although notably, no consistent pattern of results emerged when comparing incorrect and correct coherency judgments. This may not be surprising since there were fewer incorrect trials (i.e., leading to reduced statistical power) and moreover, a number of factors could have potentially contributed to inaccurate coherency judgments (e.g., incorrect button presses, momentary lack of concentration, etc.). Thus, since incorrect trials can be difficult to interpret and as we were primarily interested in neural activity when participants processed spatial coherency successfully, we have chosen to focus our discussion on the correct trial analyses.

A second-level statistical analysis was then conducted across the two functional sessions for each subject using a standard weighted fixed effects model, and with each functional session as a higher level regressor. The resulting parameter estimate images for each participant were then combined in a higher-level (group) analysis using a method of permutation-based non-parametric inference (Randomise function of FSL: http://www.fmrib.ox.ac.uk/fsl/randomise, version 2.1; 5000 permutations used) known as threshold-free cluster enhancement (TFCE: Smith and Nichols, 2009). TFCE identifies clusters of activation from standard GLM designs and does not require the size and extent of clusters to be pre-defined beforehand. As such, the use of TFCE can avoid problems/bias related to identifying significant activations according to some pre-selected arbitrary threshold. To correct for multiple comparisons, TFCE applies a multi-threshold meta-analysis of random field theory cluster-*P* values. In the current study, clusters with a corrected family-wise (FWE) *P*-value of < 0.05 within each of our predefined ROIs — HC, PPA, PHG, and RSC — were considered statistically significant (i.e., small volume correction, SVC; see "fMRI Regions of Interest" section later).

Examining neural activity associated with subsequent memory. For each run of each participant, one predictor was convolved with a double-gamma model of the HRF for each EV. Since we were interested in successful versus unsuccessful coding, trials were modelled according to subsequent memory

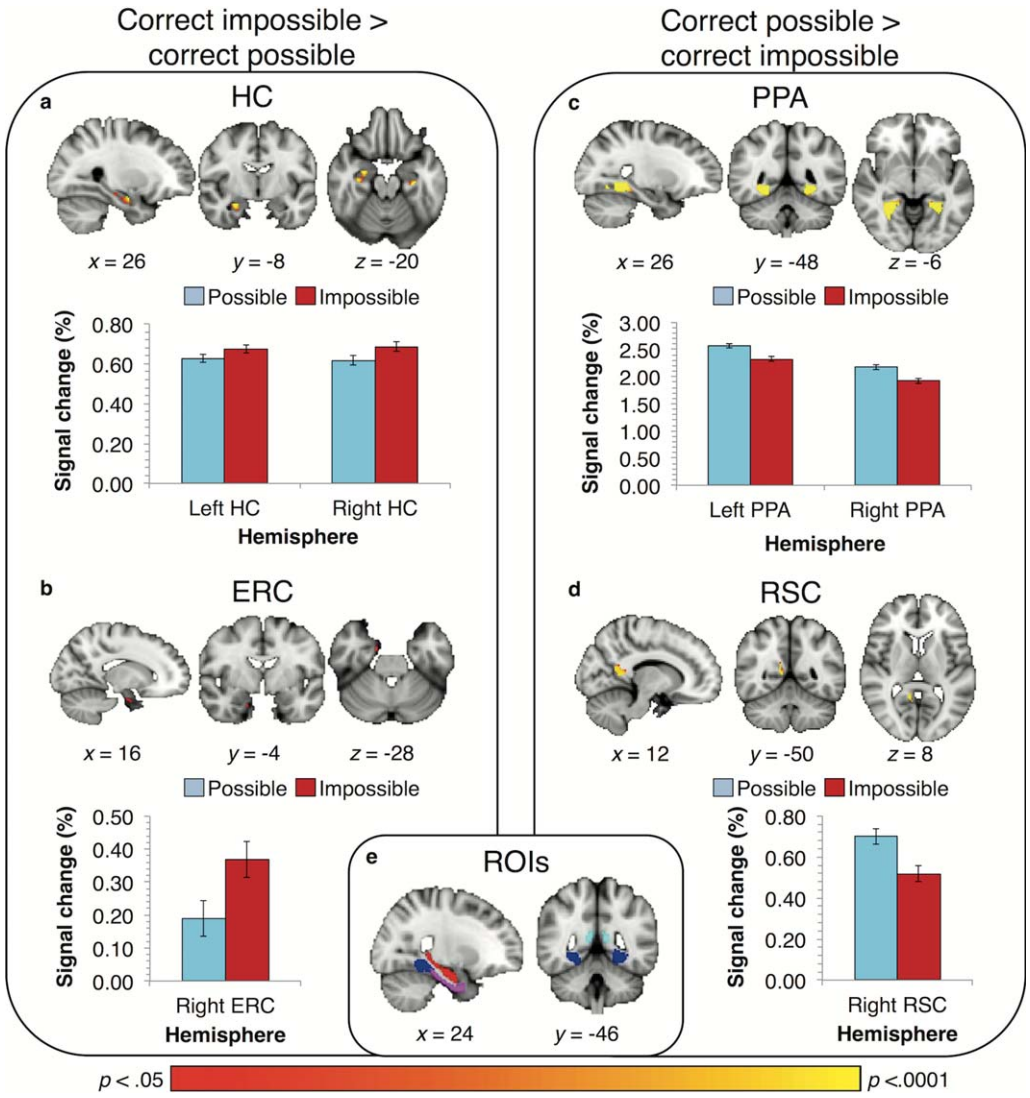

FIGURE 2. Sagittal, coronal, and horizontal views of significant clusters in each ROI above bar graphs illustrating the mean percent signal change extracted from the maximum peak voxel for each ROI. Clusters resulting from the contrast of *correct impossible > correct possible scenes* are depicted for (a) HC and (b) ERC. Clusters resulting from the contrast of *correct possible > correct impossible scenes* in the (c), PPA and (d) RSC. Error bars represent SEM of the within-subject difference between impossible and possible conditions for each correct and incorrect trials respectively. All clusters are depicted at $P < 0.05$, FWE-svc and rendered on MNI152 template. 0 percent signal change denotes the baseline activity (mean signal intensity) across the scanning session. ROIs are depicted in (e): HC (red), PHG (fuchsia), PPA (blue), and RSC (light blue). [Color figure can be viewed at wileyonlinelibrary.com]

irrespective of coherency accuracy. There were 9 EVs: one each for trials that were subsequently recognized or forgotten for each condition (recognition hit or miss for possible scene, impossible scene, possible scrambled, impossible scrambled), and one EV for fixation. Parameter estimates relating the height of the HRF response to each EV were calculated on a voxel by voxel basis, via a multiple linear regression of the response time-course, creating one beta image for each regressor. Beta images were also created to investigate whether each of our ROIs showed activity greater to remembered scenes compared to forgotten scenes: "*remembered scene (impossible + possible) > forgotten scene (impossible + possible)*." We also analyzed activity specific to recognition of each spatial coherency condition: *remembered possible scene > forgotten possible scene*, remembered *impossible scene > forgotten impossible scene*. A second-level statistical analysis was conducted with the same specifications as above (see "Examining neural activity associated with spatial coherency judgments").

Functional localizer fMRI analysis

For each participant, a predictor was convolved with a double-gamma model for each stimulus category (scenes, objects, faces) and one baseline condition (fixation; i.e., 4 EVs in total). Parameter estimates relating the height of the HRF response to each EV were calculated on a voxel by voxel basis,

via a multiple linear regression of the response time-course. This created one beta image for each EV as well as the planned contrast of interest "*scene – (object + face)*" to isolate scene-selective voxels in the parahippocampal cortex (parahippocampal place area, PPA). An additional contrast of "*object - (scene + face)*" was used to isolate object sensitive voxels in the lateral occipital cortex (LOC) to further investigate effects found in the multivariate connectivity analysis (see "Results—fMRI study"). The resulting parameter estimate images for each participant were combined in a higher-level (group) analysis with a mixed effects model, and significant activity in the parahippocampal cortex, and lateral occipital cortex were identified using a random field-based voxel-wise threshold of $P < 0.01$, family-wise error corrected for multiple comparisons. This scene-sensitive region was subsequently used as a group-level functional ROI (fROI).

fMRI Regions of Interest. Four bilateral masks were used; one fROI and three anatomical ROIs (aROI), in order to investigate neural activity during spatial coherency judgments in *a priori* brain regions of interest (Fig. 2). The fROI comprised the PPA as identified by the functional localizer (PPA: 707 voxels). The HC and PHG aROIs were each created from the Harvard Cortical and Sub-Cortical Spatial Atlases. For the former, the right and left hippocampi were combined and thresholded at 50% (1013 voxels). The PHG mask was created using a combination of the anterior and posterior parahippocampal gyri in both hemispheres, thresholded at 50% (876 voxels). The RSC mask was created from the MRIcron Brodmann Template (www.mccauslandcenter.sc.edu/mricro/mricron), by combining Brodmann areas 29 and 30; 318 voxels). Finally, the LOC mask was identified by the functional localizer and used to interrogate the Multivariate PLS data (226 voxels).

Multivariate fMRI data analyses

In order to explore possible functional interactions between regions identified in the univariate analyses and the rest of the brain, functional connectivity was investigated employing a partial least squares approach (seed-PLS; McIntosh et al., 1996; McIntosh and Lobaugh, 2004). Seed-PLS is a covariance-based multivariate analysis technique that examines how patterns of activity vary across the brain with respect to activity in a seed ROI over time, in this case 5 time windows (each corresponding to a TR) from the start of each trial. We adopted a non-rotated (i.e., contrast-based as opposed to data-driven) approach (McIntosh et al., 1996), which allowed us to examine the possibility that there were reliable differences in connectivity that related to a contrast of interest. To achieve this, correlation matrices capturing the relationship between a seed (e.g., HC peak voxel identified from the univariate analysis) and the brain data (BOLD signal), as well as a contrast of interest (i.e., correct possible versus correct impossible) and brain data (BOLD signal) were calculated (Krishnan et al., 2011), and stacked to form a single matrix. Next, singular-value decomposition was completed, resulting in a set of orthogonal latent variables (LVs) that describe the relation between the seed/design and data matrices. Each LV comprised: (1) a singular-value indicating the strength of the LV; (2) a linear contrast between the task conditions coding the effect depicted by the LV; and (3) a singular image detailing the weight, or "salience" of each voxel to the LV within each time lag proportional to the covariance of activity with the contrast. The statistical significance of each LV was formally assessed via a permutation technique that randomly reassigned each subject's data to each condition, prior to rerunning the analysis. With each reordering a new singular value was created, and in the current analysis this reordering was carried out 500 times (i.e., 500 permutations). As such, the significance for a given LV reflects the exact probability that the singular value from the permuted data exceeded the original LV singular value, applying a threshold of $P < 0.05$. The reliability of the LV voxel saliences from the singular image was established using bootstrap estimation of the standard error (SE). This procedure involved randomly resampling subjects with replacement, and rerunning the PLS analysis to determine new saliences. This bootstrap procedure was repeated 200 times, and the bootstrap SE was calculated from the resampled saliences. As the ratio of the observed salience to the bootstrap SE approximates a *z*-score, clusters comprising ≥ 15 voxels with a bootstrap ratio (BSR) > 2.81 were deemed reliable.

Eye-Tracking

Participants

Fourteen undergraduates (eight female; aged 18–22 years, $M = 19.86$; S.D. $= 1.23$) were recruited from the University of Toronto community. Each volunteer gave informed written consent to participate in the study and was compensated class credit for their time. This study was approved by the Office of Research Ethics at the University of Toronto (Ref: 26827).

General procedure

Eye movements were recorded during a self-paced spatial coherency judgment task. Participants were presented with single spatial scene images and were instructed to determine whether each stimulus was "possible" or "impossible" using pre-specified buttons on a keypad. The task was self-paced so that trials terminated once participants responded. A Tobii T120 system (Tobii Technology), comprising a 17 in. LCD monitor (1280 × 1024 pixel resolution) with an inbuilt 120 Hz infrared eye-tracker, was used in conjunction with Tobii Studio version 3.0.9 software to administer the task, and to record participant responses and eye movements.

Stimuli

Scene images consisted of the same virtual reality images used in the fMRI study.

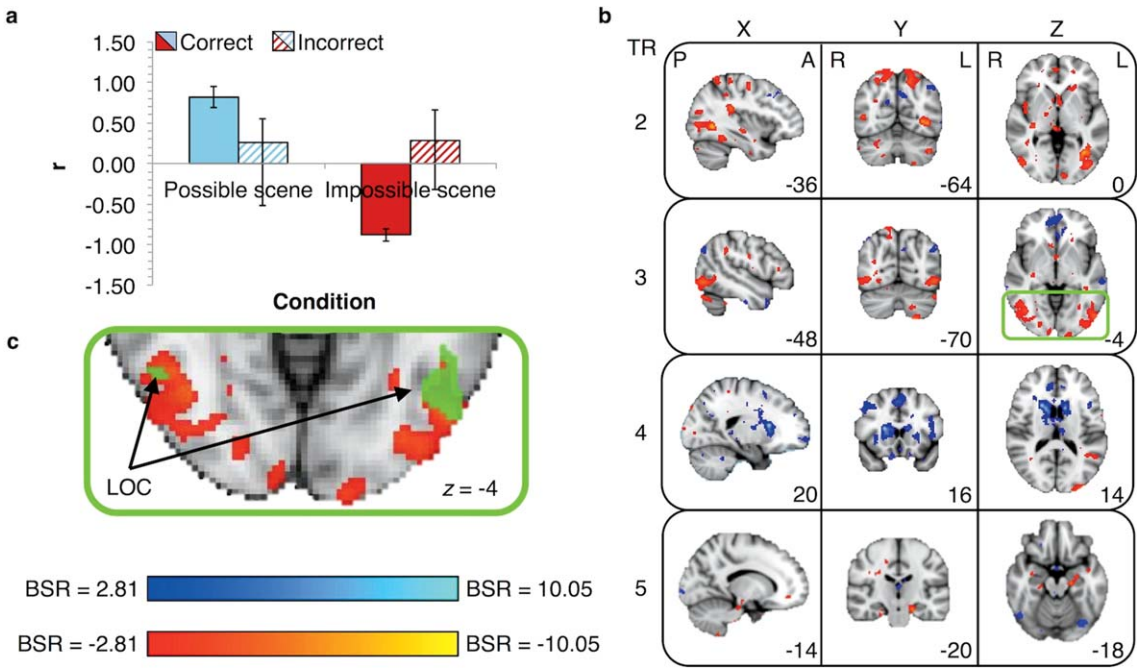

FIGURE 3. Results of the HC seed PLS contrast, depicting differences in connectivity between the HC and rest of the brain for correct possible and correct impossible scene conditions. (a) Bar graph with *Y*-axis indicating the degree to which HC seed voxel activity (26, −8, 20) correlates with the network activity for correct possible (shown in blue) and correct impossible (shown in red) conditions. Error bars denote 95% confidence intervals derived from 500 permutation tests. (b) Whole-brain patterns of connectivity across TRs 2-5 from stimulus onset, corresponding to the contrast depicted in (a). Cool colors reflect greater connectivity with HC seed for correct possible scenes, warm colors reflect greater connectivity with HC seed for correct impossible scenes. All clusters are depicted at $P < 0.005$; BSR = bootstrap ratio. (c) Inlay depicting the overlap between HC connectivity for impossible scene (red) and an overlay of LOC mask from an independent localizer (green). [Color figure can be viewed at wileyonlinelibrary.com]

Data analyses

Performance accuracy during eye-tracking (percent of trials correct), and reaction times (correct-trials only) were each analyzed with a paired-samples *t*-test in order to compare performance on the possible and impossible scenes. Eye movement data were first processed by algorithms as implemented in Tobii Studio, the full details of which are reported in previous work (Erez et al., 2013), and also available online (www.tobii.com). Briefly, gaze point data was examined with a sliding 42 ms window, and fixations within this time were processed such that any two gaze points within 35 pixels were considered a single fixation point, whereas gaze points more than 35 pixels apart were considered separate fixation points. Missing gaze data of less than 100 ms were filled in by interpolation.

To provide insight into participant viewing patterns, the spatially incoherent portion of each impossible scene was delineated as an area of interest (AOI) and each impossible scene AOI was then superimposed onto the associated possible scene counterpart in order to serve as a baseline for eye movements to the AOI. We constrained the analysis of eye movement data to trials for which there was no loss of eye-tracking signal (mean 6.89 trials excluded, pattern of results similar without filtering). Eye-tracking data from each correct trial were then interrogated with the following measures: number of fixations overall; percentage of fixations to the AOI compared to the rest of the scene; and ratio of fixations within the AOI compared to between the AOI and rest of the scene. Each of these measures was subjected to a paired-samples t-test of possible versus impossible scenes.

RESULTS

fMRI Study

Participants were significantly above chance at determining the spatial coherency of both possible ($t_{(15)} = 13.67$, $P < 0.001$) and impossible ($t_{(15)} = 13.34$, $P = < 0.001$) scenes, with no significant difference in accuracy between the two trial types ($t_{(15)} = 1.89$, $P = 0.08$; see Fig. 1d). Reaction times were, however, significantly longer for correct possible compared to correct impossible scenes ($t_{(15)} = 8.91$, $P < 0.001$; see Fig. 1e). Univariate fMRI ROI analyses of "(*correct impossible* + *correct possible*) > (*correct possible scrambled* + *correct impossible scrambled*)" revealed that all four ROIs responded more to scenes than their scrambled counterparts (see Supplementary Material for further details), implying that each of these regions is involved in processing scenes.

Our univariate fMRI ROI analyses (see Supplementary Material Table 1 for whole-brain level findings) of *correct impossible > correct possible scenes* revealed significantly greater neural activity only in the HC and PHG. There were four clusters of activity in the HC, one in the right anterior HC (peak: 26, −8, −20, $P = 0.016$, family-wise error small volume corrected (FWE-svc), 96 voxels; see Fig. 2a), and three more posteriorly in the left hemisphere (peak: −28, −24, −18, $P = 0.014$, FWE-svc, 57 voxels; peak: −18, −38, 0, $P = 0.030$, FWE-svc, 15 voxels; peak: −26, −14, −24, $P = 0.046$, FWE-svc, 8 voxels). In the PHG, a significant cluster of activity was found in the medial ERC (peak: 16, −6, −28, $P = 0.0043$, FWE-svc, 7 voxels; see Fig. 2b). Notably, there were no significant clusters in either the HC or PHG associated with greater activity during *correct possible > correct impossible scenes*. Aspects of the PPA and RSC, on the other hand, revealed a contrasting pattern of activity. Significantly greater activity was found for *correct possible > correct impossible scenes* in the right PPA (peak: 26, −48, −6, $P < 0.0001$, FWE-svc, 367 voxels; see Fig. 2c) and left PPA (peak: −28, −54, −10, $P < 0.0001$, FWE-svc, 252 voxels), as well as the RSC in the right hemisphere (peak: 12, −50, 8, $P = 0.006$, FWE-svc, 51 voxels; see Fig. 2d), with no significant activity associated with *correct impossible > correct possible scenes* in either of these ROIs. Note that all peaks in the HC, PHG, PPA, and RSC survived whole-brain FWE correction ($P < 0.05$) with the exception of two of the HC peaks (peak: −18, −38, 0, peak: −26, −14, −24).

We next examined the functional connectivity of each ROI with multivariate PLS, using a contrast-based approach (i.e., bidirectional contrast between correct impossible and correct possible scenes). Signal was extracted from the peak HC and ERC voxels in each hemisphere identified with the univariate contrast *correct impossible > correct possible scenes*, and from the peak RSC and PPA voxels in each hemisphere identified with the univariate contrast *correct possible > correct impossible scenes*. Importantly, the approach used to select the seed regions is agnostic with respect to condition-dependant changes in functional connectivity with the rest of the brain.

The right HC seed region exhibited distinct patterns of connectivity during correct impossible and correct possible scene trials ($P = 0.04$). The same analyses, however, did not reveal a reliable distinction in seed connectivity between the left HC ($P = 0.40$), ERC ($P = 0.56$), PPA (left: $P = 0.76$; right: $P = 0.89$), or RSC ($P = 0.08$), and the rest of the brain. During the accurate identification of impossible scenes, we found robust connectivity between the right HC and several posterior and anterior regions of visual cortex and the temporal lobe, including the occipital pole (−18, −92, −18), intracalcarine cortex (18, −86, 8), inferior lateral occipital cortex (−32, −72, 4), superior lateral occipital cortex (−16, −60, 52), temporal occipital fusiform cortex (30, −42, −18), and PHG (−18, −20, −20; see Fig. 3b; for full details see Supplementary Material Table 2). Notably, two of these peak connectivity coordinates (44, −60, 0; 52, −62, −10) fell within an independently defined mask of object-sensitive lateral occipital complex (Malach et al., 1995; see Fig. 3c; LOC; *objects > scenes + faces*; see Functional Localizer fMRI analysis).

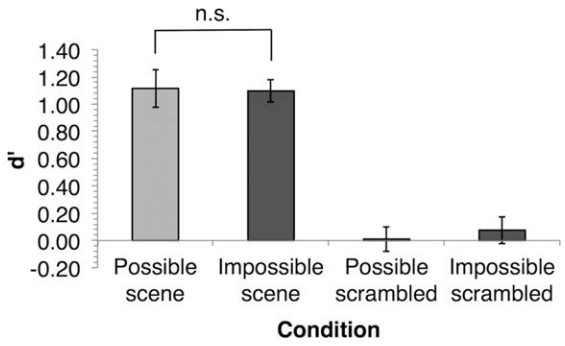

FIGURE 4. Subsequent recognition memory performance (d') after scanning for correct coherency judgement trials. Error bars represent standard error.

For the same contrast of connectivity between correct impossible and correct possible scenes, the accurate identification of possible scenes was associated with greater connectivity between the HC and the caudate (20, 16, 14), as well as several midline regions including the cingulate gyrus (0, −38, 16), paracingulate gyrus (−6, 50, 18) and precuneus cortex (14, −64, 26), and frontal regions, including the precentral gyrus (−44, 2, 36), middle frontal gyrus (36, 20, 42), superior frontal gyrus (−24, 8, 66), and frontal pole (12, 54, 30). Significant connectivity was also found between the HC and the superior lateral occipital cortex (−46, −82, 18), angular gyrus (−40, −58, 44), and inferior temporal gyrus (44, −24, −20; for full details see Supplementary Material Table 2). Notably, inaccurate trials did not contribute to this pattern of distinct connectivity (see Fig. 3a). While our data cannot distinguish the direction of the relationship between functional connectivity and behavioral performance, this finding suggests that these two factors are related.

Finally, in order to determine whether our fMRI findings could be explained by differences in long-term declarative memory processing (i.e., episodic encoding), we administered a surprise recognition memory test after scanning for the stimuli that had been presented in the scanner. Participants were significantly above chance at remembering correct possible and impossible scene images as measured by d' correct possible scenes: $t_{(15)} = 8.08$, $P < 0.001$; correct impossible scenes: $t_{(15)} = 13.40$, $P < 0.001$; see Fig. 4). Crucially, there was no significant difference in performance between these two ($t_{(15)} = 0.14$, $P = 0.89$).

Modelling the fMRI data according to subsequent memory (irrespective of coherency judgement accuracy) did not reveal significantly greater activity for *remembered > forgotten scenes* (possible and impossible combined) in any of the four ROIs, although there was a sub-threshold cluster in the left PPA (peak: −18, −42, −10, $P = 0.091$, FWE-svc). Comparisons of *remembered > forgotten scenes* within each scene type also did not reveal significant effects of memory for either possible or impossible scenes. Notably, however, two clusters of activity approached significance for *remembered > forgotten possible*

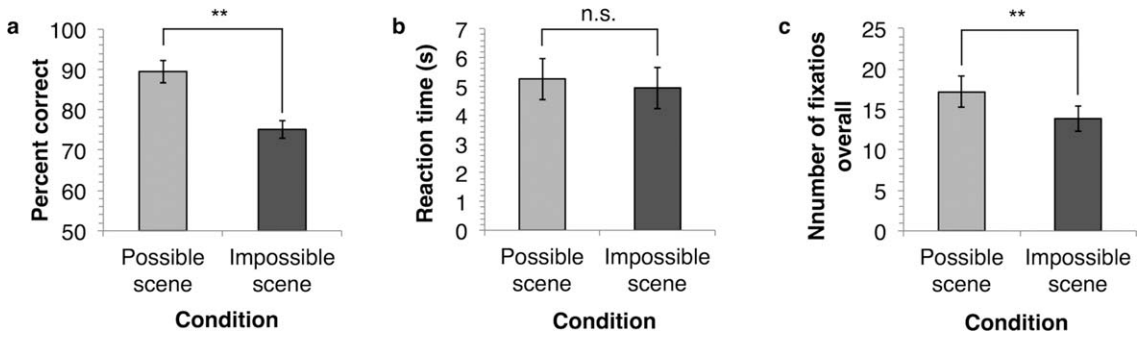

FIGURE 5. Eye-tracking perceptual judgement task performance. (a) Bar graphs depict average accuracy (percent correct). The following graphs represent performance on correct trials only: (b) reaction time (s) and, (c) number of fixations across each scene. Error bars represent standard error, two asterisks indicate significance at $P < 0.01$.

scenes: one in the right hemisphere PPA (peak: 22, −48, −14, $P = 0.054$, FWE-svc), and one in the right HC (peak: 24, −14, −16, $P = 0.062$, FWE-svc).

Eye-Tracking Study

Since HC activity has been previously shown to be sensitive to short-term spatial information load (Schon et al., 2009; Lee and Rudebeck 2010a), one potential explanation for the greater HC activity during impossible compared to possible scenes is that the former place a greater demand on holding scene information online while making coherency judgements. To address this, we recorded eye movements while participants performed self-paced spatial coherency judgments to the scenes presented in the fMRI study. Given that fixations denote attention to a scene region (Irwin and Gordon, 1998), we expected that the amount of scene information available for mental representation would increase with the number of eye movements across a trial (Van Orden et al., 2001), as the task necessitates the integration of information acquired at each fixation. We thus used number of fixations as a proxy for the amount of information held online (i.e., in short-term or working memory) across a trial. Eye-tracking also allowed us to investigate which areas of the scene participants examined while performing spatial coherency judgments. To this end, we quantified viewing patterns for each scene to assess how often participants fixated on the spatial incoherency embedded within the scene, as well as the pattern of fixations between the region of spatial incoherency and rest of the scene.

Behavioral accuracy was significantly greater for possible compared to impossible scenes ($t_{(13)} = 4.19$, $P < 0.01$; see Fig. 5a), whereas reaction times were not significantly different across conditions ($t_{(13)} = .61$, $P = 0.55$; see Fig. 5b). Analysis of eye movements revealed more total fixations for correct spatial coherency judgments of possible scenes than for correct spatial coherency judgments of impossible scenes ($t_{(13)} = 3.89$, $P < 0.01$; see Fig. 5c), and this effect remained after normalizing for time spent viewing the scene ($t_{(13)} = 2.21$, $P < 0.05$). These data imply that judging the spatial coherency of possible scenes required greater sampling of scene information than judging the spatial coherency of impossible scenes, inconsistent with the notion that impossible scenes were associated with greater maintenance of scene information as compared to possible scenes over the course of a trial.

To assess how spatial incoherencies affected eye movement patterns during impossible spatial coherency judgments, we delineated the spatially incoherent portion of each impossible scene as an area of interest (AOI; see Fig. 6a). Each impossible scene AOI was then superimposed onto the associated possible scene counterpart in order to serve as a baseline for eye movements to the impossible region (see Fig. 6b). We examined the proportion of fixations to the impossible region AOI compared to the rest of the scene (see Fig. 6c). The vast majority (81%) of fixations across the impossible scene were to the AOI, which was significantly greater than the percent of fixations to the corresponding AOI in the possible scenes (59%, $t_{(13)} = 10.51$, $P < 0.001$; see Fig. 6d); when scaled by the size of the AOI, these effects held ($t_{(13)} = 10.25$; $P < 0.001$). To further quantify how participants viewed the spatial incoherency within the scene, we compared the number of consecutive fixations within the bounds of the AOI ("within" fixations) to consecutive fixations between the AOI and the rest of the scene ("between" fixations) and calculated a within-to-between ratio of saccades for possible and impossible scenes (Gajewski and Henderson, 2005; Erez et al., 2013). This revealed more within-AOI fixations than between-AOI fixations for impossible, but not possible, scenes (impossible: $t_{(13)} = 4.19$, $P = 0.001$, possible: $t_{(13)} = 0.28$, $P = 0.79$). Likewise, the within-to-between AOI ratio was greater for impossible scenes than for possible scenes ($t_{(13)} = 6.74$, $P < 0.001$; see Fig. 6e). This pattern of viewing suggests that any given fixation to a spatial incoherency within a scene was usually immediately followed by additional fixations to other parts of the spatial incoherency, as opposed to the surrounding spatially coherent elements of the image. This pattern of viewing seems to indicate that before identifying a scene as impossible, multiple fixations are made across the spatially incoherent portion of the scene, possibly in an attempt to bind the elements of the spatial incoherency together.

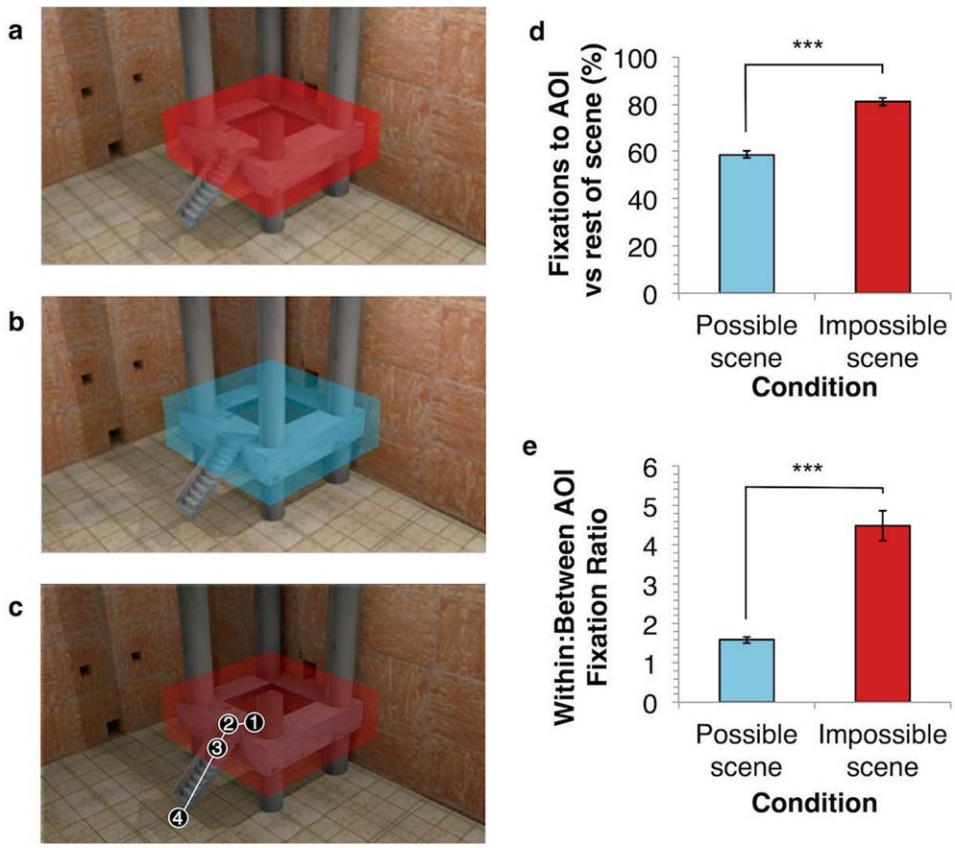

FIGURE 6. Area of Interest (AOI) analyses for the eye-tracking perceptual judgement task. Examples of AOI over (a) the structural incongruency within an impossible scene (in red) and (b) the corresponding area of the possible scene counterpart (in blue). (c) Example of fixations over a single trial, with a within to between ratio of 3:1. Fixations are numbered first ("1") to last ("4"), saccades are represented by white lines connecting fixations. The saccade between fixations 1 and 2 constitutes a "within" fixation, whereas the saccade between fixations 3 and 4 constitutes a "between" fixation. The following graphs represent performance on correct trials only: (d) average percent fixations to the AOI compared to the entire scene, and (e) average ratio of fixations within the AOI compared to between the AOI and rest of the scene. Error bars represent standard error; three asterisks indicate significance at $P < 0.001$. [Color figure can be viewed at wileyonlinelibrary.com]

Finally, when we analyzed eye movements for the subset of trials in which participants responded within 5000 ms (the time constraint in the fMRI task), we found a similar pattern of behavioral results to those reported for the fMRI task (no difference in accuracy, significant difference in reaction time), whereas the eye-movement results were nearly identical to those reported above (see Supplementary Material).

DISCUSSION

To date, few studies have focused on how the brain supports spatial coherency, a crucial component of normal scene processing in which components of a scenes are bound into a coherent percept. Here we examined the neural correlates of spatial coherency in scene processing regions using a novel paradigm: spatial incoherencies were embedded into virtual reality scenes, and participants judged whether these scenes were possible or impossible in the real word. The present work provides, to our knowledge, the first demonstration that activity in the HC, ERC, RSC, and PPA is sensitive to violations of spatial coherency. Univariate fMRI analyses revealed that these brain regions contributed differentially, with greater HC and medial ERC activity during the detection of spatial incoherency, and greater PPA and RSC activity during the detection of spatial coherency. Further, seed-based multivariate functional connectivity analyses revealed reliable changes in HC functional connectivity when accurately assessing spatially incoherent as compared to coherent scenes, which were not present for the PPA, ERC, or RSC.

Given the well-established role of the HC in mnemonic processing (Squire, 1982; Moscovitch et al., 2005), it is important to note that our observation of increased hippocampal activity for impossible, relative to possible, scenes was not driven by differences in memory demands across conditions. In short, we found no evidence for greater incidental encoding for

the impossible scene condition. There were no differences in the proportion of remembered to forgotten scenes between possible and impossible scenes, and a direct analysis of remembered versus forgotten scenes did not reveal any significant changes in activity. These findings also undermine the argument that the increased HC activity during impossible scenes was simply the result of these scenes being more attention-grabbing than the possible scenes, since one would expect increased attention to a stimulus to enhance subsequent memory. There was, however, a trend for greater activity in subsequently remembered versus forgotten possible scenes in the HC and PPA, which was not present for impossible scenes. This trend is consistent with previous reports of greater activity in the HC for old versus new possible objects, but not impossible objects (Schacter et al., 1995), suggesting that spatial coherency may affect how the MTL encodes stimuli. Thus, given that subsequent memory was not different between conditions, and direct fMRI contrasts of remembered and forgotten stimuli did not reveal significant activity in any ROI, it seems unlikely that greater activity in the HC to impossible scenes can be explained by a greater memory demand in this condition.

A separate eye-tracking experiment to examine scene information load showed that, when making correct spatial coherency judgments, a greater number of fixations were made to possible compared to impossible scenes. This was true for both number of total fixations as well as number of fixations per second, suggesting that correctly identifying possible scenes may entail holding more scene-features online compared to correctly identifying impossible scenes. We also observed a shift in fixation strategy when participants viewed impossible scenes such that multiple fixations were made across the spatially incoherent portion of the scene. This qualitative difference in eye-movements likely reflects an attempt to bind the incoherent portions of impossible scenes into a coherent whole. It is important to highlight that the number of fixations cannot account for the increased activity for impossible scenes in the HC and ERC, although this may have contributed to the greater activity for possible scenes in the PPA and RSC (see later for discussion). A similar argument could also be made for the longer reaction times in the possible scene condition: if greater or protracted cognitive effort is driving the increase in neural response, this would only increase activity associated with coherency judgments of possible scenes. In the HC, we found the exactly opposite pattern: greater activity for impossible scenes, which were associated with fewer fixations and shorter reaction times. We argue, therefore, that our data suggest that the HC is sensitive to the perceptual properties of scenes, and speak to a current debate in the literature as to whether the HC is involved in scene processing beyond the mnemonic domain (Kim et al., 2011; Lee et al., 2012).

We propose that the observed increase in HC activity to spatially incoherent scenes reflects the representational content of this brain region, supporting the view that a critical determinant of HC involvement in scene processing is the extent to which a given task taxes complex conjunctions of spatial features, irrespective of the cognitive process involved (e.g., memory or perception; Lee et al., 2012). According to this viewpoint, the greater HC activity during impossible scenes can be interpreted as the increased demand that these stimuli place on binding the different spatial features within each image to create conjunctive representations. It is important to note that since a non-spatial coherency comparison (e.g., possible vs. impossible objects) was not included in the current study, the present data cannot, on their own, be interpreted as evidence for a specialized role for the HC in processing scenes. Crucially, however, it is the convergence of the current work with other studies, which suggests that our findings reflect a role for the HC in representing conjunctions of spatial features. Neuropsychological studies of populations with HC damage have reported visual discrimination deficits for complex scenes, but not for objects or faces (Lee et al., 2005a; b), including intact perception of Escher-like impossible objects (Lee and Rudebeck, 2010b). Similarly, neuroimaging studies of neurologically intact individuals found increased HC activity when participants discriminated scenes on the basis of conjunctions of spatial features as opposed to single features and/or objects (Lee et al., 2008; Barense et al., 2010). Notably, the tasks in each of these studies required the simultaneous comparison of multiple scene stimuli, complicating the interpretation of their results with respect to perception and/or mnemonic processing. Here, we report evidence for HC involvement in a perceptual scene task involving a unique, single stimulus on each trial, thus eliminating the need to remember information across trials or maintain multiple distinct scenes within each trial. Amidst a growing literature indicating HC involvement in spatial cognition beyond long-term memory, we believe that these data provide the strongest evidence to date that the role of the HC in spatial cognition is not limited to the domain of long-term or working memory, but extends to perception.

In keeping with the idea that the HC processes complex conjunctions of spatial features and that viewing impossible scenes taxes the binding of scene elements, our functional connectivity analysis revealed differential HC connectivity between correct possible and correct impossible scene conditions. Most notably, correct impossible trials (compared to correct possible trials) were associated with increased connectivity between the hippocampus and posterior visual regions, including the LOC. Although one interpretation is that impossible scenes were perceived as more object-like, leading to increased recruitment of the LOC, we prefer an alternate explanation. We argue that our functional connectivity findings suggest that the hippocampus recruits object regions in the service of scene representation in order to resolve violations of spatial coherency. This argument is supported by past research highlighting HC involvement in scene perception, but not object perception (Lee et al., 2008; Lee and Rudebeck, 2010b). More specifically, we believe that a greater demand on scene feature binding during the perception of impossible scenes may place a greater emphasis on local feature integration, resulting in a greater interplay between LOC and the HC. This conclusion is further supported by our eye-tracking findings, in which we found that participants fixate sequentially within the impossible section of

a scene, rather than between this section and the rest of the scene. Not inconsistent with this suggestion, recent work has suggested a role for object regions, such as the LOC, in scene processing. For example, MacEvoy and Epstein (2011) recently demonstrated that the LOC is sensitive to scene-category information associated with particular objects. Similarly, Park et al. (2011) reported that scene classification based on patterns of activity in the LOC was more likely to confuse scenes with similar content (i.e., two natural scenes) than two scenes with similar geometric properties (i.e., two images of open fields), suggesting that the LOC does carry scene-relevant information. While these studies were aimed at exploring scene-recognition based on semantic information carried by scene-embedded objects, the current findings hint that LOC-dependent perceptual information may also contribute to scene cognition.

In contrast, accurate possible trials were associated with greater connectivity between the HC and a number of brain regions including the caudate, and multiple midline, temporal and frontal regions (see Supplementary Material Table 2 for full details). With no *a priori* hypotheses regarding the involvement of these regions, further work will be required to fully understand the contributions of these areas to the identification of possible scenes. Nevertheless, given the dominance of frontal regions in our observed results, a comment on their involvement is in order. It is possible that the involvement of the observed frontal areas may be explained by their role in maintaining top-down representations of stimulus class. For example, the dorsal medial prefrontal cortex has been observed to be significantly active when participants are anticipating the presentation of a specific class of images, regardless of incoming sensory information (Summerfield et al., 2006). In contrast, greater orbitofrontal cortex activity has been found when there is a match between anticipated and incoming stimuli (Summerfield and Koechlin, 2008). In the current paradigm, it is therefore conceivable that co-activation between HC and these frontal regions reflects subjects' prior knowledge of scenes, which is more closely matched to the presented possible compared to impossible scenes.

Our observation that the HC is involved in our scene perception task is consistent with several theories of HC function. Informational accounts posit that the HC binds relational associations between items, and between items and contexts (Davachi, 2006; Ranganath, 2010). Some have suggested an extension of the role of HC in relational binding to include relationships between items held in working memory (Olsen et al., 2012), and, more recently, some forms of perception (Aly et al., 2013). Here we argue that the constituent elements of spatially impossible scenes may be bound in relation to one another by the HC, an assertion not incompatible with more recent interpretations of relational theory. Note, however, that relational accounts do not assign any special status to scene stimuli per se, but rather view scenes as one instance of a relational percept. In contrast, the scene construction hypothesis posits that scenes are the central concern of the HC, and that the HC will be implicated in any cognitive function that relies on the construction of a scene (Maguire and Mullally, 2013). Particularly, scene construction theory has focused on the need for the HC in "top-down" creation of scene representations beyond the bounds of incoming visual information, for example, the extension of scene boundaries in memory beyond the veridical limits of those presented scenes (Mullally et al., 2012). Comparatively, the representational hierarchical model (Saksida and Bussey, 2010) emphasizes the hierarchical organization of incoming sensory information, with complex scenes constituting the greatest level of conjunctive representations in the ventral visual stream. Nonetheless, both the representational hierarchical model and the scene construction model place scene representation as the critical foundation of HC function and are, therefore, in our opinion complementary viewpoints.

In contrast to the HC, the PPA demonstrated preferential involvement during the perception of spatially coherent scenes. Since response times and the number of fixations was greater for possible compared to impossible scenes, we cannot discount entirely that longer processing of the possible scenes may have contributed to this finding. Our preferred interpretation, however, is that this response pattern reflects the well-established role of the PPA in processing the perceptual layout of spatial scenes (Epstein et al., 1999). Although relatively little research has investigated how the PPA processes spatial geometry, our findings echo early reports of PPA activity, showing that this region is more active to non-contiguous, coherently arranged scene fragments in comparison to separated and incoherently rearranged scene fragments (Epstein and Kanwisher, 1998). Similarly, Zeidman et al. (2012) reported that the PPA responds preferentially to very simple "horizons" containing depth cues compared to nearly identical stimuli without such cues. This demonstration of PPA response to bare horizons implies that three-dimensional space is an important driver of activity in this region. Additionally, scene classification based on pattern analysis of PPA response have shown that the PPA is more likely to confuse scenes with similar geometric properties than scenes with similar content, implying that another important driver of PPA response is the geometric layout of a scene (Park et al., 2011). Our study adds an important component to this growing literature by showing that the PPA is sensitive to quite subtle manipulations of overall scene structure, such that neural activity in this region is diminished for scenes with slight violations of 3D spatial coherency. In sum, we believe that violations of 3D spatial coherency in the impossible scenes make these stimuli less "scene-like" compared to spatially coherent scenes, thus disrupting normal scene processing in the brain. This is in keeping with the idea that the PPA plays an important role in processing the perceptual layout of spatial scenes (Epstein et al., 1999), and that both the implication of 3D space and spatial layout modulate PPA activity. Notably, unlike the HC, we did not find differential connectivity between the PPA and the rest of the brain for the different scene conditions. Although null results can be difficult to interpret, we note that the PPA is thought to represent details of a scene, but this role is not thought to extend to binding these elements together in the service of scene representation in the same manner as the HC (Epstein, 2008). Further research will

be necessary in order to understand fully how the HC and PPA interact during the processing of spatial coherency.

Though the response to impossible scenes found in the HC strongly suggests a role of this region in binding disparate scene elements into a coherent whole, limitations in our design preclude conclusions regarding the specific nature of structural coherencies that drive the observed effect. In our impossible scene stimuli, spatial coherency was violated in one of a number of ways, including inconsistent occlusion of scene elements and violations of perspective (see Fig. 1). Moreover, the number of scene elements implicated in the violation varied across stimuli. Because of the limited number of scenes employed, a systematic investigation of HC response related to the degree or kind of spatial incoherency would be severely underpowered in the current paradigm. Importantly, though, we carefully controlled for any potential confounds between coherency and scene complexity by matching the possible and impossible scenes as closely as possible by creating pairs of possible and impossible scenes that were identical except for a region of incoherency in the latter. Future investigations might contrast the kind of spatial violations embedded in impossible scenes or parametrically vary the number of elements of a scene implicated in spatial coherency.

Finally, although our primary focus was on the HC and PPA, we also observed sensitivity to manipulations of spatial coherency in the ERC and RSC. A small cluster of activity in the medial ERC showed a similar response pattern to the HC. There was greater activity during the impossible compared to possible scene condition, although, importantly, there was no significant change in functional connectivity associated with the medial ERC across conditions. Given the rich anatomical connectivity between this region and the HC (Amaral and Witter, 1989), this finding may reflect initial processing of converging spatial information, destined for further processing in the HC. In contrast, response patterns in the RSC were similar to those in the PPA, with higher activity for possible compared to impossible scenes. In the light of the role of the RSC in global spatial navigation (Epstein and Higgins, 2007; Park and Chun, 2009), as opposed to feature binding, this response could reflect a preference for an intact spatial environment similar to that of the PPA.

The current study provides novel insight into the neural correlates underlying spatial coherency processing, and highlights the differential contributions of the HC and PPA to scene cognition. Here, for the first time, we demonstrate that activity in the HC is sensitive to a scene perception task in which scene stimuli were presented in isolation. In line with an important role for the HC in scene perception and processing complex conjunctions of spatial features, we found greater activity in this structure for impossible compared to possible scenes. Further, increased functional connectivity between the HC and posterior visual regions when viewing spatially incoherent scenes included the LOC. This, along with our eye-tracking data, suggests that the identification of impossible scenes places a greater emphasis on binding the local elements of the spatial incoherency together. Taken together, we propose that our results are consistent with a central role of the HC in the perceptual processing of scenes.

ACKNOWLEDGMENTS

Authors thank Jennifer Ryan for helpful discussion, staff at the Centre for Addiction and Mental Health, Toronto for assisting with imaging data collection, and our participants for their time. The authors have no conflict of interest to declare.

REFERENCES

Aguirre GK, D'Esposito M. 1999. Topographical disorientation: A synthesis and taxonomy. Brain 122:1613–1628.

Aly M, Ranganath C, Yonelinas AP. 2013. Detecting changes in scenes: The hippocampus is critical for strength-based perception. Neuron 78:1127–1137.

Amaral DG, Witter MP. 1989. The three-dimensional organization of the hippocampal formation: A review of anatomical data. Neuroscience 31:571–591.

Barense MD, Henson RNA, Lee ACH, Graham KS. 2010. Medial temporal lobe activity during complex discrimination of faces, objects, and scenes: Effects of viewpoint. Hippocampus 20:389–401.

Bird CM, Burgess N. 2008. The hippocampus and memory: Insights from spatial processing. Nat Rev Neurosci 9:182–194.

Burgess N. 2002. The hippocampus, space, and viewpoints in episodic memory. Q J Exp Psychol a 55:1057–1080.

Davachi L. 2006. Item, context and relational episodic encoding in humans. Curr Opin Neurobiol 16:693–700.

Ekstrom AD, Kahana MJ, Caplan JB, Fields TA, Isham EA, Newman EL, Fried I. 2003. Cellular networks underlying human spatial navigation. Nature 425:184–188.

Epstein R, Graham KS, Downing PE. 2003. Viewpoint-specific scene representations in human parahippocampal cortex. Neuron 37: 865–876.

Epstein R, Harris A, Stanley D, Kanwisher N. 1999. The parahippocampal place area: recognition, navigation, or encoding?. Neuron 23:115–125.

Epstein R, Kanwisher N. 1998. A cortical representation of the local visual environment. Nature 392:598–601.

Epstein RA, Higgins JS. 2007. Differential parahippocampal and retrosplenial involvement in three types of visual scene recognition. Cerebral Cortex 17:1680–1693.

Epstein RA. 2008. Parahippocampal and retrosplenial contributions to human spatial navigation. Trends Cogn Sci 12:388–396.

Erez J, Lee ACH, Barense MD. 2013. It does not look odd to me: perceptual impairments and eye movements in amnesic patients with medial temporal lobe damage. Neuropsychologia 51:168–180.

Freud E, Rosenthal G, Ganel T, Avidan G. 2015. Sensitivity to object impossibility in the human visual cortex: evidence from functional connectivity. J Cogn Neurosci 27:1029–1043.

Fyhn M, Molden S, Witter MP, Moser EI, Moser M-B. 2004. Spatial representation in the entorhinal cortex. Science 305:1258–1264.

Gajewski D, Henderson J. 2005. Minimal use of working memory in a scene comparison task. Visual Cognition 12:979–1002.

Glover GH. 2012. Spiral imaging in fMRI. NeuroImage 62:706–712.

Greve DN, Fischl B. 2009. Accurate and robust brain image alignment using boundary-based registration. NeuroImage 48:63–72.

Harel A, Kravitz DJ, Baker CI. 2013. Deconstructing visual scenes in cortex: gradients of object and spatial layout information. Cerebral Cortex 23:947–957.

Hassabis D, Kumaran D, Vann SD, Maguire EA. 2007. Patients with hippocampal amnesia cannot imagine new experiences. Proc Natl Acad Sci USA 104:1726–1731.

Hori E, Tabuchi E, Matsumura N, Tamura R, Eifuku S, Endo S, Nishijo H, Ono T. 2003. Representation of place by monkey hippocampal neurons in real and virtual translocation. Hippocampus 13:190–196.

Iglói K, Doeller CF, Berthoz A, Rondi-Reig L, Burgess N. 2010. Lateralized human hippocampal activity predicts navigation based on sequence or place memory. Proc Natl Acad Sci USA 107:14466–14471.

Irwin DE, Gordon RD. 1998. Eye Movements, Attention and Transsaccadic Memory. Visual Cognition 5:127–155.

Jacobs J, Weidemann CT, Miller JF, Solway A, Burke JF, Wei X-X, Suthana N, Sperling MR, Sharan AD, Fried I, Kahana MJ. 2013. Direct recordings of grid-like neuronal activity in human spatial navigation. Nat Neurosci 16:1188–1190.

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 17:825–841.

Kim S, Dede AJO, Hopkins RO, Squire LR. 2015. Memory, scene construction, and the human hippocampus. Proc Natl Acad Sci USA 112:4767–4772.

Kim S, Jeneson A, van der Horst AS, Frascino JC, Hopkins RO, Squire LR. 2011. Memory, visual discrimination performance, and the human hippocampus. J Neurosci 31:2624–2629.

King JA, Burgess N, Hartley T, Vargha-Khadem F, O'Keefe J. 2002. Human hippocampus and viewpoint dependence in spatial memory. Hippocampus 12:811–820.

Krishnan A, Williams LJ, McIntosh AR. 2011. Partial Least Squares (PLS) methods for neuroimaging: A tutorial and review. NeuroImage 56:455–475.

Lee ACH, Buckley MJ, Pegman SJ, Spiers HJ, Scahill VL, Gaffan D, Bussey TJ, Davies RR, Kapur N, Hodges JR, Graham KS. 2005a. Specialization in the medial temporal lobe for processing of objects and scenes. Hippocampus 15:782–797.

Lee ACH, Bussey TJ, Murray EA, Saksida LM, Epstein RA, Kapur N, Hodges JR, Graham KS. 2005b. Perceptual deficits in amnesia: challenging the medial temporal lobe "mnemonic" view. Neuropsychologia 43:1–11.

Lee ACH, Rudebeck SR. 2010a. Investigating the interaction between spatial perception and working memory in the human medial temporal lobe. J Cogn Neurosci 22:2823–2835.

Lee ACH, Rudebeck SR. 2010b. Human Medial Temporal Lobe Damage Can Disrupt the Perception of Single Objects. J Neurosci 30:6588–6594.

Lee ACH, Scahill VL, Graham KS. 2008. Activating the medial temporal lobe during oddity judgment for faces and scenes. Cerebral Cortex 18:683–696.

Lee ACH, Yeung L-K, Barense MD. 2012. The hippocampus and visual perception. Front Hum Neurosci 6:1–17.

Lever C, Burton S, Jeewajee A, O'Keefe J, Burgess N. 2009. Boundary vector cells in the subiculum of the hippocampal formation. J Neurosci 29:9771–9777.

MacEvoy SP, Epstein RA. 2011. Constructing scenes from objects in human occipitotemporal cortex. Nat Neurosci 14:1–9.

Macmillan NA, Kaplan HL. 1985. Detection theory analysis of group data: estimating sensitivity from average hit and false-alarm rates. Psychol Bull 98:185–199.

Maguire EA, Frackowiak RSJ, Frith CD. 1997. Recalling routes around London: activation of the right hippocampus in taxi drivers. J Neurosci 17:7103–7110.

Maguire EA, Mullally SL. 2013. The hippocampus: a manifesto for change. J Exp Psychol: Gen 142:1180–1189.

Maguire EA, Nannery R, Spiers HJ. 2006. Navigation around London by a taxi driver with bilateral hippocampal lesions. Brain 129: 2894–2907.

Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB. 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc Natl Acad Sci USA 92:8135–8139.

Marchette SA, Vass LK, Ryan J, Epstein RA. 2015. Outside Looking In: Landmark Generalization in the Human Navigational System. J Neurosci 35:14896–14908.

McIntosh AR, Bookstein FL, Haxby JV, Grady CL. 1996. Spatial pattern analysis of functional brain images using partial least squares. NeuroImage 3:143–157.

McIntosh AR, Lobaugh NJ. 2004. Partial least squares analysis of neuroimaging data: applications and advances. NeuroImage 23 Suppl 1:S250–63.

Moscovitch M, Rosenbaum RS, Gilboa A, Addis DR, Westmacott R, Grady C, McAndrews MP, Levine B, Black S, Winocur G, Nadel L. 2005. Functional neuroanatomy of remote episodic, semantic and spatial memory: a unified account based on multiple trace theory. J Anat 207:35–66.

Mullally SL, Intraub H, Maguire EA. 2012. Attenuated Boundary Extension Produces a Paradoxical Memory Advantage in Amnesic Patients. Curr Biol 22:261–268.

O'Keefe J, Dostrovsky J. 1971. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Res 34:171–175.

O'Keefe J, Nadel L. 1978. The hippocampus as a cognitive map. Oxford: The Clarendon Press.

Olsen RK, Moses SN, Riggs L, Ryan JD. 2012. The hippocampus supports multiple cognitive processes through relational binding and comparison. Front Hum Neurosci 6:146-

Park S, Brady TF, Greene MR, Oliva A. 2011. Disentangling scene content from spatial boundary: complementary roles for the parahippocampal place area and lateral occipital complex in representing real-world scenes. J Neurosci 31:1333–1340.

Park S, Chun MM. 2009. Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. NeuroImage 47:1747–1756.

Ranganath C. 2010. A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. Hippocampus 20:1263–1290.

Rolls ET, Robertson RG, Georges-François P. 1997. Spatial View Cells in the Primate Hippocampus. Eur J Neurosci 9:1789–1794.

Saksida LM, Bussey TJ. 2010. The representational-hierarchical view of amnesia: translation from animal to human. Neuropsychologia 48:2370–2384.

Schacter DL, Reiman E, Uecker A, Polster MR, Yun LS, Cooper LA. 1995. Brain regions associated with retrieval of structurally coherent visual information. Nature 376:587–590.

Schon K, Quiroz YT, Hasselmo ME, Stern CE. 2009. Greater Working Memory Load Results in Greater Medial Temporal Activity at Retrieval. Cerebral Cortex 19:2561–2571.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady M, Matthews PM. 2004. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23 Suppl 1:S208–19.

Smith SM, Nichols TE. 2009. Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. NeuroImage [Internet] 44:83–98. Available from: http://www.sciencedirect.com/science/article/pii/S1053811908002978

Smith SM. 2002. Fast robust automated brain extraction. Hum Brain Mapp 17:143–155.

Solstad T, Boccara CN, Kropff E, Moser M-B, Moser EI. 2008. Representation of geometric borders in the entorhinal cortex. Science 322:1865–1868.

Spiers HJ, Burgess N, Maguire EA, Baxendale SA, Hartley T, Thompson PJ, O'Keefe J. 2001. Unilateral temporal lobectomy patients show lateralized topographical and episodic memory deficits in a virtual town. Brain 124:2476–2489.

Spiers HJ, Maguire EA. 2007. A navigational guidance system in the human brain. Hippocampus 17:618–626.

Squire LR. 1982. The neuropsychology of human memory. Annu Rev Neurosci 5:241–273.

Summerfield C, Egner T, Greene M, Koechlin E, Mangels J, Hirsch J. 2006. Predictive codes for forthcoming perception in the frontal cortex. Science 314:1311–1314.

Summerfield C, Koechlin E. 2008. A Neural Representation of Prior Information during Perceptual Inference. Neuron 59:336–347.

Van Orden KF, Limbert W, Makeig S, Jung TP. 2001. Eye activity correlates of workload during a visuospatial memory task. Hum Factors 43:111–121.

Vann SD, Aggleton JP, Maguire EA. 2009. What does the retrosplenial cortex do?. Nat Rev Neurosci 10:792–802.

Vass LK, Epstein RA. 2013. Abstract representations of location and facing direction in the human brain. J Neurosci 33:6133–6142.

Williams P, Tarr MJ. 1997. Structural processing and implicit memory for possible and impossible figures. J Exp Psychol: Learn, Memory, Cogn 23:1344–1361.

Witter MP, Moser EI. 2006. Spatial representation and the architecture of the entorhinal cortex. Trends Neurosci 29:671–678.

Yonelinas AP. 2013. The hippocampus supports high-resolution binding in the service of perception, working memory and long-term memory. Behavioural Brain Res 254:34–44.

Zeidman P, Mullally SL, Schwarzkopf DS, Maguire EA. 2012. Exploring the parahippocampal cortex response to high and low spatial frequency spaces. Neuroreport 23:503–507.