bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Unrestricted eye movements strengthen causal connectivity from hippocampal to oculomotor regions during scene construction

Natalia Ladyka-Wojcik[1]*, Zhong-Xu Liu[2]*, and Jennifer D. Ryan[1,3,4]

[1] *Department of Psychology, University of Toronto, Toronto ON, CANADA*

[2] *Department of Behavioral Sciences, University of Michigan-Dearborn, Dearborn MI, USA*

[3] *Rotman Research Institute, Baycrest Health Sciences, Toronto ON, CANADA*

[4] *Department of Psychiatry, University of Toronto, Toronto ON, CANADA*

***Equal Contribution**

Corresponding Authors:

Natalia Ladyka-Wojcik
100 St. George Street, Sidney Smith Hall Room 523
Department of Psychology,
University of Toronto
Toronto, ON, CANADA, M5S 3G3
Email: natalia.ladyka.wojcik@mail.utoronto.ca

Zhong-Xu Liu
4901 Evergreen Rd.
Behavioral Sciences Department,
University of Michigan Dearborn
Dearborn, MI, 48128
Email: zhongxu@umich.edu

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Scene construction is a key component of memory recall, navigation, and future imagining, and relies on the medial temporal lobes (MTL). A parallel body of work suggests that eye movements may enable the imagination and construction of scenes, even in the absence of external visual input. There are vast structural and functional connections between regions of the MTL and those of the oculomotor system. However, the directionality of connections between the MTL and oculomotor control regions, and how it relates to scene construction, has not been studied directly in human neuroimaging. In the current study, we used dynamic causal modeling (DCM) to investigate this relationship at a mechanistic level using a scene construction task in which participants' eye movements were either restricted (*fixed-viewing*) or unrestricted (*free-viewing*). By omitting external visual input, and by contrasting free- versus fixed- viewing, the directionality of neural connectivity during scene construction could be determined. As opposed to when eye movements were restricted, allowing free viewing during construction of scenes strengthened top-down connections from the MTL to the frontal eye fields, and to lower-level cortical visual processing regions, suppressed bottom-up connections along the visual stream, and enhanced vividness of the constructed scenes. Taken together, these findings provide novel, non-invasive evidence for the causal architecture between the MTL memory system and oculomotor system associated with constructing vivid mental representations of scenes.

Keywords: scene construction, eye movements, medial temporal lobes, hippocampus, oculomotor system, dynamic causal modeling

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Our ability to form spatial representations in our mind's eye is key for supporting navigation, memory, and future thinking (Hassabis & Maguire, 2007; Robin et al., 2016). Functional magnetic resonance imaging (fMRI) studies have demonstrated engagement of the parahippocampal place area (PPA) and hippocampus (HPC) in the encoding of scenes (Epstein & Kanwisher, 1998), as well as in scene construction, the mental generation of coherent spatial contexts in the absence of visual input (Hassabis & Maguire, 2007). Together, the PPA and HPC support scene construction by binding disparate features and objects into integrated representations (Douglas et al., 2017; Maguire & Mullally, 2013).

Parallel lines of evidence from the field of vision science suggest that saccadic eye movements may play a key role in the construction of scene representations (Kowler, 2011). Contributions of the oculomotor system to scene construction have received limited investigation (see Mirza et al., 2016; Parr & Friston, 2017), largely because prior neuroimaging studies have typically instructed participants to close their eyes (Hassabis et al., 2007; Mullally et al., 2012). Saccade motor maps accurately code for locations of objects in space (Zimmermann & Lappe, 2016). Adaptively changing the targeting position of saccades (i.e., changing the required saccade amplitude) subsequently disrupts localization of objects in visual space, suggesting that scene perception and memory may rely on an oculomotor map (Bahcall & Kowler, 2000; Ryan & Shen, 2020). Conversely, saccades and corresponding gaze fixations are guided by prior knowledge regarding the expected locations of objects within scenes (Castelhano & Heaven, 2011). Patterns of gaze fixations during imagination are similar to those during perception, suggesting that oculomotor mechanisms support mental imagery (Gurtner et al., 2021) by reinstating previously encoded spatiotemporal content (Wynn et al., 2019). Individuals move their eyes across a blank screen in accordance with object positions during recall of previously studied scenes (Johansson et al., 2006) and when merely listening to

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

auditory scene descriptions (Spivey et al., 2000). Eye movements therefore support construction of mental representations even in the absence of external visual input (Conti & Irish, 2021).

Recent fMRI findings highlight the functional connectivity between the PPA and early visual regions (Baldassano et al., 2013). Computational modeling has revealed vast structural connections between the MTL and oculomotor control regions, including the frontal eye fields (FEF) (Ryan, Shen et al., 2020; Shen et al., 2016). These connections are functionally relevant; simulated stimulation of HPC subfields and parahippocampus resulted in rapid evoked responses in the FEF (Ryan, Shen et al., 2020). In humans, single-pulse transcranial magnetic stimulation of FEF created top-down activity that directly shaped responses in lower-level visual regions (Veniero et al., 2021). However, to our knowledge, no study has investigated the direction of information flow among the MTL, FEF, and early visual regions in human neuroimaging, and specifically, during scene construction.

The present study used combined eyetracking-fMRI recordings and manipulations of viewing behavior to elucidate interactions among the MTL, FEF, and visual cortex during scene construction. Participants were prompted with word labels of scenes and instructed to freely move their eyes around a blank screen (free-viewing) or to maintain fixation (fixed-viewing) while imagining the cued scene. Dynamic causal modeling (Friston et al., 2003) was used to interrogate directionality of effective couplings between regions and how the directional coupling can be modulated by the viewing manipulation. Prior work has shown that neural activity in the PPA and HPC (Liu et al., 2020) scaled with increasing gaze fixations, whereas restricting fixations reduced neural activity (Liu et al., 2017, 2020). By comparing the free-viewing versus fixed-viewing conditions, we could clearly assess directionality of information flow between regions. Compared to fixed-viewing, free-viewing during scene construction was hypothesized to strengthen top-down connections from the PPA and HPC towards the FEF and lower-level

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

visual regions. When viewing was restricted (fixed-viewing), we predicted suppression of bottom-up connections from early visual regions towards oculomotor regions. This work highlights the interaction of the MTL and oculomotor system in the active construction of scenes.

Methods

Participants

Thirty-three healthy young adults (18 female) aged 18 to 30 (age: *M* = 22.97 years, *SD* = 3.31; education: *M* = 16.29 years, *SD* = 1.93) from the University of Toronto and surrounding Toronto area community completed this experiment in exchange for monetary compensation. Thirty-one subjects had participated in a scene viewing task earlier in the same scanning session, as reported in Liu et al. (2020). All participants had normal or corrected-to-normal vision (including color vision), and none had any neurological or psychological conditions. The study was approved by the Research Ethics Board at Rotman Research Institute at Baycrest Health Sciences.

Stimuli

Stimuli consisted of 28 unique word labels for common semantic scene categories (e.g., casino, ski resort, etc.). These word labels were presented in the center of the screen followed by either a green fixation dot (free-viewing condition) or a red fixation dot (fixed-viewing condition) (Figure 1).

Procedure

Participants completed one run containing 28 trials in the scanner. Half of the trials were studied under free-viewing instructions and half were studied under restricted (i.e., fixed) viewing instructions, in a randomized order (Figure 1). At the start of each trial, participants were shown, for 2 seconds, a word label of a scene category that they were to mentally construct. The word labels were counterbalanced over the two viewing conditions across participants. Next, a fixation dot was presented at the center of the screen for 13 seconds. The color of the

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

fixation dot indicated the viewing condition for the trial; a green fixation dot indicated a free-viewing trial and a red fixation dot indicated a fixed-viewing trial. In the free-viewing condition, participants were instructed to freely explore the blank screen as they wished for the duration of the trial as they were mentally constructing a scene based on the word label cue. In the fixed-viewing condition, participants were instead required to keep their eye gaze at the location of the fixation dot while mentally constructing the cued scene. Following each scene construction trial, participants were given 2.15 seconds to respond to a vividness rating question using an MRI-compatible button box. Possible vividness ratings were: 1 (not vivid), 2 (vivid), and 3 (very vivid).

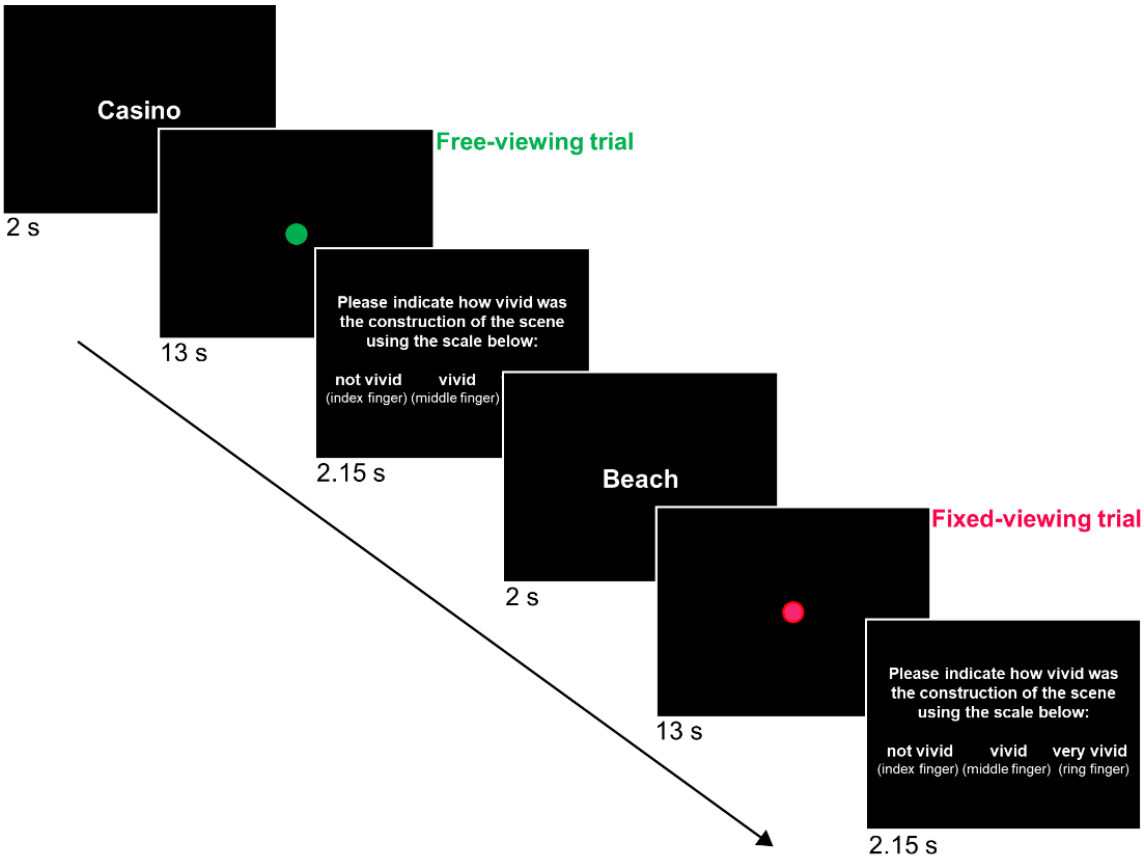

Figure 1: Scene construction task procedure. Participants were presented with a word cue of a scene and then instructed to either freely move their eye gaze across the screen (free-viewing) or to keep their eye gaze fixed on the fixation dot (fixed-viewing), as they mentally constructed the cued scene. Following the scene construction period, participants responded to the vividness of their mental construction with a 3-button response for not vivid, vivid, or very vivid.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

All stimuli were presented with Experiment Builder (Eyelink 1000; SR Research) back-projected to a screen (projector resolution: 1024x768) and viewed with a mirror mounted on the head coil.

Eyetracking

During the scene construction task, monocular eye movements were recorded inside the scanner using the EyeLink 1000 MRI-compatible remote eyetracker with a 1000Hz sampling rate (SR Research Ltd., Mississauga, Ontario, Canada). The eyetracker was placed inside the scanner bore (behind the participant's head) and detected the right pupil and corneal reflection via a mirror mounted on the head coil. To ensure successful tracking during the task, a nine-point calibration was performed at the beginning of the scanning session and online manual drift correction was performed between trials when necessary. EyeLink's default eye movement event parser was used to categorize fixations and saccades. A velocity threshold of 30°/s and an acceleration threshold of 8000°/s were used to classify saccades (saccade onset threshold = 0.15°). Events not defined as saccades or blinks were classified as fixations. The number of fixations that participants made during scene construction was calculated and exported to a MATLAB-compatible environment using the EyeLink software Data Viewer for further analyses.

MRI scan acquisition

As specified in Liu et al. (2020), a 3T Siemens MRI scanner with a standard 32-channel head coil was used to acquire structural and functional MRI images. T1-weighted high-resolution MRI images for structural scans were obtained using a standard 3D MPRAGE (magnetization-prepared rapid acquisition gradient echo) pulse sequence (176 slices, FOV = 256 x 256 mm, 256x256 matrix, 1 mm isotropic resolution, TE/TR = 2.22/2000 ms, flip angle = 9 degrees, and scan time = 280 s). For the functional scan, BOLD signal was assessed using a $T2^*$-weighted EPI acquisition protocol with TR = 2000 ms, TE = 27 ms, flip angle = 70 degrees, and FOV = 192 x 192 with a 64 x 64 matrix (3 mm x 3 mm in-place resolution; slice thickness = 3.5 mm with no gap). A total of 250 volumes were acquired for the fMRI run, with the first 5

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

discarded to allow the magnetization to stabilize to a steady state. Both structural and functional images were acquired in an oblique orientation 30° clockwise to the anterior–posterior commissure axis.

fMRI data preprocessing

The fMRI preprocessing procedure was previously reported in Liu et al. (2020) and is reproduced here. SPM12 (Statistical Parametric Mapping, Wellcome Trust Centre for Neuroimaging, University College London, UK) in the MATLAB environment (The MathWorks Inc., Natick, USA) was used to process the functional images. Following standard SPM12 preprocessing procedure, slice timing was first corrected using *sinc* interpolation with the reference slice set to the midpoint slice. Next, functional images were aligned using a linear transformation, and for each participant functional image parameters from the alignment procedure (along with global signal intensity) were checked manually using the toolbox ART (http://www.nitrc.org/projects/artifact_detect/). Anatomical images were co-registered to the aligned functional image and segmented into white matter, gray matter, cerebrospinal fluid, skull, and soft tissues using SPM12's default 6-tissue probability maps. Segmented images were then used to calculate the transformation parameters mapping from the subjects' native space to the MNI template space. The resulting transformation parameters were used to transform all functional and structural images to the MNI template. The functional images were finally resampled at 2x2x2 mm resolution and smoothed using a Gaussian kernel with an FWHM of 6 mm. The first five fMRI volumes from each run were discarded to allow the magnetization to stabilize to a steady state.

GLM fMRI analysis

We used SPM12 to conduct the first-level (i.e., individual) whole brain General Linear Model (GLM) analysis, comparing brain activation differences between the free-viewing and fixed-viewing conditions. We separately convolved the onset of trials (duration = 13 s) in the free-viewing and fixed-viewing condition with the canonical hemodynamic function (HRF) in

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

SPM12, which served to be the 2 main regressors of interest. We added 6 motion parameters obtained from the co-registration process as regressors of no interest. Default high-pass filters with a cut-off of 128 s were applied and serial correlations were removed using a first-order autoregressive model AR(1). Next, to examine differences in neural responses that were elicited by gaze fixations, we contrasted the free-viewing with the fixed viewing condition for each participant. Then, at the group-level, individual participants' above-described contrast estimates were entered into one-sample *t* tests.

For this analysis, we first focused on two *a priori* ROIs, i.e., HPC and PPA. Based on our previous findings (Liu et al., 2017; 2020), we hypothesized that both regions should show stronger activity in the free- versus fixed-viewing condition. For the HPC mask, FreeSurfer's *recon-all* function (https://surfer.nmr.mgh.harvard.edu/; Fischl, 2012) was used to extract subject-specific anatomical masks. For the PPA, group-level masks were defined functionally based on an earlier scene-scrambled picture processing task (i.e., scene versus scrambled images) reported in Liu et al. (2020), as 31 participants had previously participated in the picture processing task during the same testing session. The MNI coordinates for the peak activation in the right PPA were [32, -34, -18] and peak activation in the left PPA were [-24, -46, -12]. The left and right PPA mask contained 293 and 454 (1x1x1 mm^3) voxels, respectively. In this analysis, we used one-tailed *t* tests to test our *a priori* hypothesis.

In addition to the ROI analysis, we also examined voxel-wise whole brain results to explore regions showing different engagement in the two conditions (i.e., *free-viewing – fixed-viewing*). The statistical threshold was set to p = .005 with 10 voxel extension (uncorrected) to facilitate future meta-analyses (Lieberman & Cunningham, 2009).The automated anatomical labeling (AAL) toolbox (Tzourio-Mazoyer et al., 2002) was used to identify anatomical labels for regions that showed significant effects.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Dynamic Causal Modeling analysis

Model design

In the present study, we used dynamic causal modeling (DCM) to assess the directionality of information flow among regions of the MTL involved in scene representations and regions involved in oculomotor control, along with early visual regions. DCM for fMRI has been shown to have high scan-rescan reliability (Schuyler et al., 2010) and lends itself to modeling neural activity separately from BOLD responses (Stephan & Friston, 2010). Since DCM estimates region-specific hemodynamic responses and neural states, it is also less susceptible to issues related to HRF variations in different brain regions (Friston et al., 2014; Stephan et al., 2007). Here, we focused on three separate models to investigate the top-down influences of the MTL and the bottom-up influences of early visual regions on the oculomotor system.

First, we focused on a 3-ROI model with the PPA, FEF, and primary visual cortex (V1). This model (see Figure 2A) was selected to investigate how top-down signals from the scene processing region PPA towards the oculomotor control region FEF may be modulated when participants were allowed to freely move their eyes. Specifically, we hypothesized that allowing free eye movements during scene construction would strengthen the directional connections from the PPA to the FEF, and that this would, in turn, drive the activation of V1, typically the first stage of cortical processing of visual information.

Next, to examine how the hippocampus specifically interacts with the oculomotor control and early visual regions, we specified a 3-ROI model with the HPC, FEF, and V1 (see Figure 2B). Like the above-described model with the PPA, we were interested in investigating how top-down signals from the memory representation region HPC towards the FEF may be modulated by free-viewing versus fixed-viewing during scene construction. As for the PPA, we hypothesized that allowing free viewing would strengthen the top-down connections from the HPC to the FEF, and from the HPC and FEF to V1. Because there is already robust evidence

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

showing that the PPA and HPC closely interact with each other, in this study, we did not put the two regions in the same model to avoid complex models and to focus on our main questions of interest.

In the third model, we explored whether the top-down modulation effect arising from free- versus fixed-viewing would be evident in the earliest visual processing pathway region in the central nervous system, i.e., the lateral geniculate nucleus (LGN) in the thalamus. Therefore, we constructed a 3-ROI model with the LGN, V1, and FEF (see Figure 2C). We excluded the connection between the FEF and LGN based on low prior evidence in humans and nonhuman primates for connectivity between these two regions (Gilbert & Li, 2013; Kashihara, 2020).

Due to the increased complexity in interpretation when including inter-hemispheric connections for DCM (Stephan et al., 2010; Stephan & Friston, 2010; Zeidman et al., 2015), we designed our models for the right and left hemispheres separately.

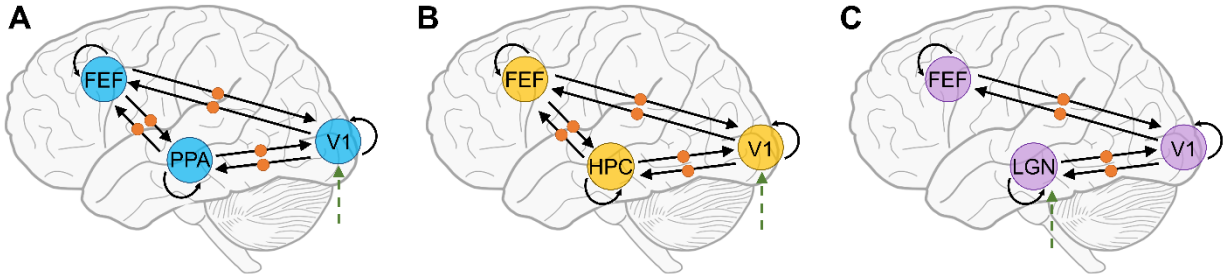

FIGURE 2: Full dynamic causal model (DCM) spaces. (A) Full model for the PPA, FEF, and V1 ROIs. (B) Full model for the HPC, FEF, and V1 ROIs. (C) Full model for the LGN, V1, and FEF ROIs. For each model, the solid straight arrows represent intrinsic connections between regions and the curved solid arrows represent intrinsic self-connections (A matrix in the state equation of DCM). The orange circles on the arrows for each model represent modulatory connections (B matrix in the state equation of DCM) and the dotted green arrows represent the driving input (C matrix in the state equation of DCM).

Time series extraction for DCM

We used the significant group-level activation clusters to extract time series for the DCM analysis, which were constrained to be within the boundaries of our ROIs. For the PPA and HPC, we used the same masks as in the GLM ROI analyses to constrain boundaries. For all

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

other ROIs, we used peak activation coordinates in the activated clusters in these regions (i.e., FEF, V1, and LGN) from the group-level GLM results, which we subsequently confirmed using Neurosynth (Yarkoni et al., 2011). Specifically, the location of the FEF was based on both the anatomical landmark (i.e., the superior frontal sulcus intersection with the precentral gyrus; (Vernet et al., 2014) and MNI coordinates found in the literature ([L: -22 -10 50; R: 20 -9 49]; Donner et al., 2000). Both V1 and LGN were located based on their anatomical landmarks on the MNI template. Then, an 8-mm radius sphere centered at peak activation was used for each participant to include voxels that showed free- vs. fixed-viewing effect (p = 0.05, no corrections) as the ROIs. Including voxels that show task modulation effects can facilitate DCM analysis (Zeidman et al., 2019). However, when the number of voxels was lower than 50 (i.e., too few voxels in the ROI) for a specific ROI of a specific participant, we relaxed the threshold to 0.1, 0.5, or without using any threshold (i.e., threshold = 1) until at least 50 voxels could be obtained for that ROI and that participant. The same threshold procedure was applied to the HPC and PPA to ensure a sufficient number of voxels included in the DCM analysis.

After ROIs were defined, SPM12 was used to extract the time series. Specifically, BOLD signals at each voxel of an ROI were extracted after effects of non-interest were partialed out. Then, the first principal component of the time series data from all voxels in the ROI was computed and used for the DCM analysis.

Individual-level DCM analysis

DCM implemented in SPM12 was used for effective connectivity analysis. DCM for fMRI models the dynamics of the neural states underlying the BOLD response by a differential state equation that describes how these responses change with the current neural states, contextual conditions, and driving inputs (Friston et al., 2003). The advantage of DCM is that it not only provides a measure of endogenous effective connectivity between regions, but also a measure of how these directed connections are modulated by task demands (Stephan & Friston, 2010).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The state equation for DCM is:

$$\dot{z} = \left(A + \Sigma u_j B^j\right)z + Cu$$

The intrinsic interactions between the neuronal states are endogenous connections and quantified by A parameters. These interactions are mediated by anatomical connections and are irrespective of task condition. The influences of the task conditions on connectivity between ROIs are modulations and quantified by B parameters, i.e., connections altered on free-viewing trials over fixed-viewing trials. The influences of driving inputs are quantified by C parameters (in this case, caused by all trials). In this study, we used a mean-centered driving input so that the matrix A parameters represented an average effective connectivity across experimental conditions and matrix B modulatory parameters add or subtract from this average (i.e., strengthening or weakening the average intrinsic connectivity). By using this mean-centered approach, we were able to focus on how the free- and fixed-viewing conditions modulate the directionality of functional underlying couplings.

As shown in Figure 2, the PPA and HPC models are configured to have bidirectional intrinsic, or underlying, connections between connected ROIs (A matrix), which is consistent with the literature on the anatomical and functional connectivity between these regions. Within-ROI auto-connections in the A matrix were also added by default (Zeidman et al., 2019). Since we were interested in how the eye movement conditions modulate effective connectivity between regions, we included all possible modulatory connections between the combinations of ROIs (B matrix) for the free-viewing condition. For all of these model designs, the driving inputs (C matrix) were set to enter the earliest ROI in the visual stream (i.e., V1 in the PPA – FEF – V1 and HPC – FEF – V1 models, and LGN in LGN – V1 – FEF model).

Group-level DCM analysis

In line with previous work using DCM (Zeidman et al., 2019), we used a Parametric Empirical Bayes (PEB) approach to evaluate group effects (i.e., commonality among participants) on connectivity parameters. With different modulatory parameters switched on or

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

off, the model space for PPA – FEF – V1 and HPC – FEF – V1 consisted of 64 possible models, and the model space for LGN – V1 – FEF had 16 possible models. This PEB process was implemented using Bayesian Model Reduction and then averaging the parameters from the best reduced models with Bayesian Model Averaging. Specifically, the winning model was selected on the basis of offering the best fit to the data with the highest exceedance probability, which denotes the probability that this model is more likely than any other in the given dataset. This analysis produced weighted model parameters for the winning model, which we report along with connectivity matrices. Since this approach relies on simultaneous estimation of nested models with Bayesian inference, we did not need to correct for multiple comparisons (Gelman & Tuerlinckx, 2000; Stephan et al., 2007). As suggested by Kass and Raftery (1995), we reported model parameters with posterior probabilities (Pp) above 95% which corresponds to strong evidence in favour of a model.

Results

Eye movements during fMRI scanning

To confirm the effect of the eye movement manipulation, we compared the average number of gaze fixations (Figure 3A) and the average saccade amplitude (Figure 3B) that participants made in the free-viewing and fixed-viewing conditions. Paired *t*-tests revealed that participants made a greater number of gaze fixations ($t(32) = 7.01$, $p < .001$) in the free-viewing than the fixed-viewing condition, and that participants had a larger saccade amplitude measured in degrees of visual angle ($t(32) = 7.00$, $p < .001$) in the free-viewing than the fixed-viewing condition. The fixation frequency across participants for the free- vs. fixed-viewing conditions based on the location within the blank screen (1024x768 pixels) is displayed in Figure 3D.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

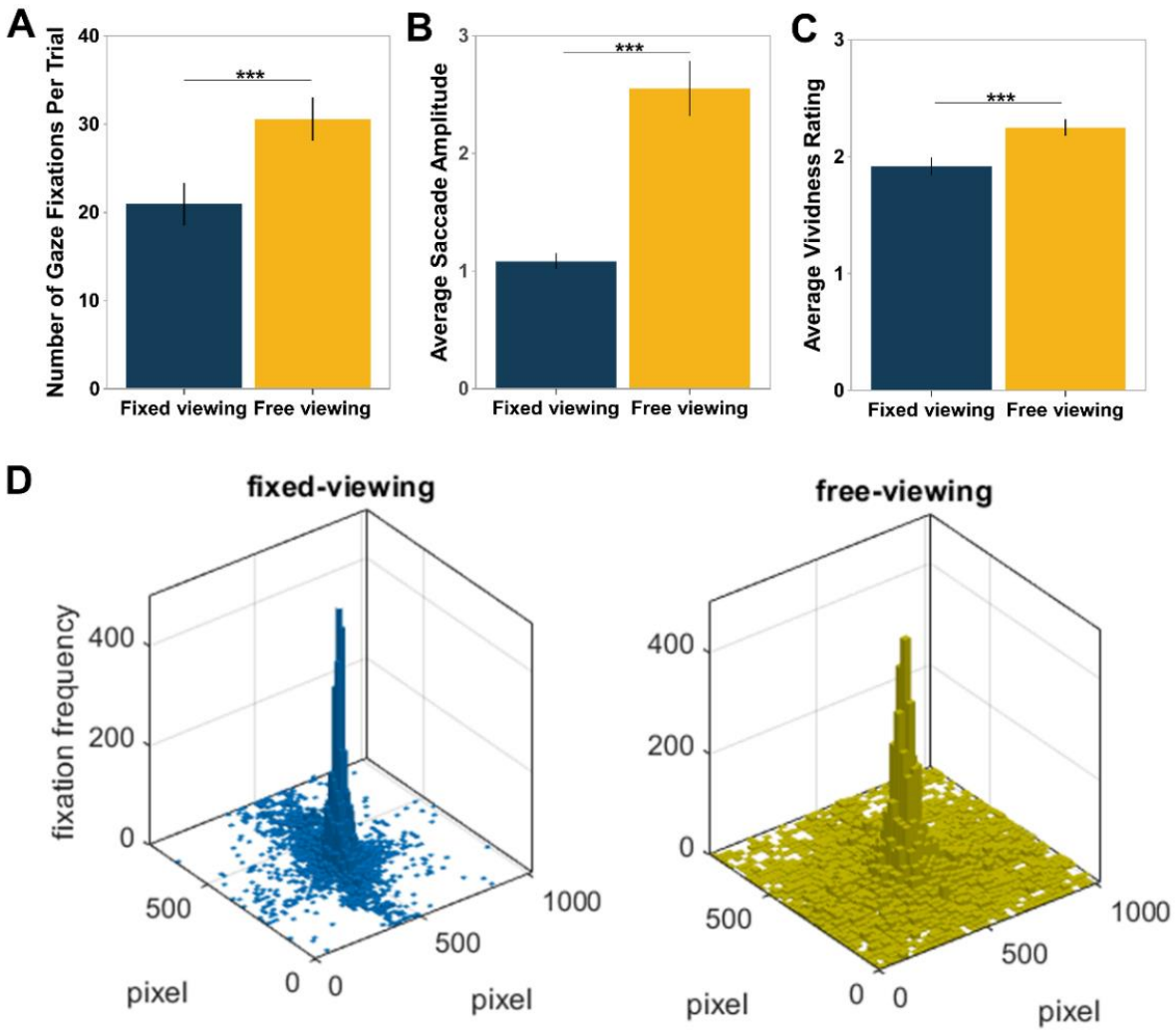

FIGURE 3: Eye-tracking results for free- vs. fixed-viewing. (A) Average number of gaze fixations per trial in the fixed-viewing and free-viewing conditions. More fixations were elicited in the free-viewing condition. (B) Average saccade amplitude in the fixed- and free-viewing conditions. Greater saccade amplitude was elicited during the free-viewing condition. (C) Average vividness rating in the fixed- and free-viewing conditions. Constructed scenes were rated more vivid during the free-viewing condition. (D) Fixation frequency based on location within the blank screen (1024x768 pixels) during fixed-viewing (left) and free-viewing (right) trials across participations. Error bars are ± SEM. ***p <.001

Vividness Ratings

At the end of each trial of the scene construction task, participants were asked to provide a vividness rating ranging across 1 (not vivid), 2 (vivid), and 3 (very vivid). We compared the average vividness rating for the fixed- and free-viewing conditions. A paired *t*-test revealed that participants rated vividness for trials in the free-viewing condition (M = 2.25, SD = 0.39) higher

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

than for trials in the fixed-viewing condition (M = 1.92, SD = 0.42; $t(32) = 4.13$, $p < .001$; Figure 3C).

GLM fMRI Results

We first examined the brain activation contrast between free- and fixed-viewing in the PPA and HPC to confirm whether activity in our *a priori* ROIs of the PPA and HPC were modulated by viewing condition during scene construction. As hypothesized, the left and right PPA showed stronger activation in the free-viewing compared to the fixed-viewing condition (left: $t(32) = 4.65$, $p < .0001$; right: $t(32) = 3.62$, $p < .001$; one-tailed). Additionally, the left HPC showed stronger activation in the free-viewing compared to the fixed-viewing condition ($t(32) = 1.70$, $p < .05$; one-tailed). Our ROI analysis of the right HPC did not yield significant effects ($t(32) = 0.54$, $p > 0.05$; one-tailed; Figure 4A).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

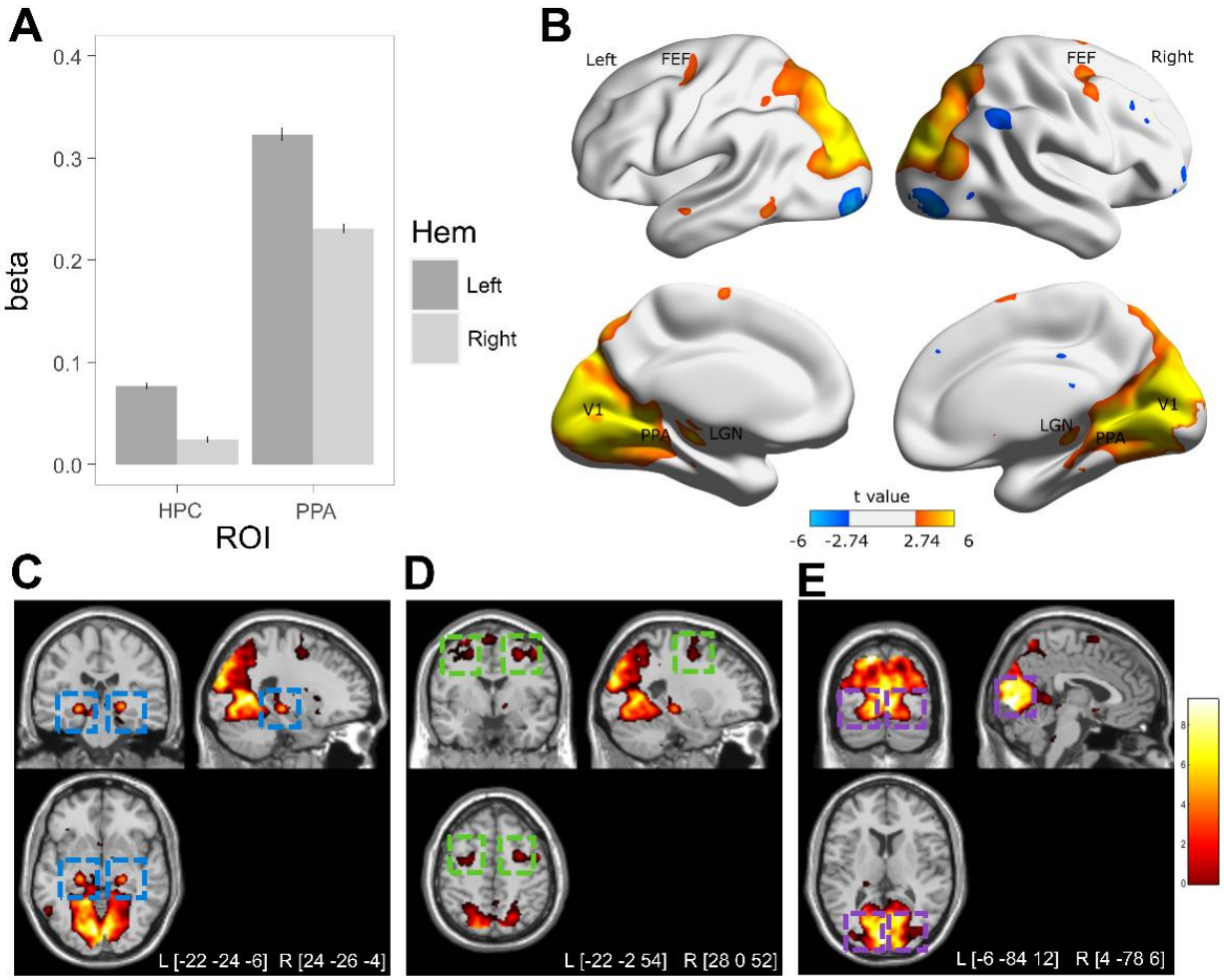

FIGURE 4: GLM fMRI results. (A) ROI analysis revealed a significant effect of free-viewing over fixed-viewing in bilateral PPA (left p <.0001; right p <.001; one-tailed), left HPC (p <.05; one-tailed), but not in right HPC (p > 0.05; one-tailed). Error bars are ± SEM. (B) Brain surface plots showing voxel-wise activation differences between the free-viewing and fixed-viewing conditions. (C) Voxel-wise whole brain results showing stronger activation in the left and right LGN in the free- vs. fixed-viewing condition (highlighted in blue boxes; *ps* < .00001). (D) Voxel-wise whole brain results showing stronger activation in the left and right FEF in the free- vs. fixed-viewing condition (highlighted in green boxes; p = .001 and .0003). (E) Voxel-wise whole brain results showing stronger activation in the left and right V1 in the free- vs. fixed-viewing condition (highlighted in purple boxes; *ps* < .00001). For (B), (C), (D), and (E), brain images are thresholded at p < .005, 10 voxel extension (no corrections) for illustration purposes. Clusters showed free-viewing > fixed-viewing effects at FEF, PPA, HPC, V1 and LGN are indicated.

The results of the voxel-wise whole brain analysis that showed increased and decreased neural activity during free-viewing compared to fixed-viewing are illustrated in Figure 4B and listed in Table 1 in Appendix. As can be seen in Figure 4C, 4D, and 4E, both the ventral and

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

dorsal visual processing pathway regions, including in the bilateral LGN (Figure 4C), bilateral FEF (Figure 4D), and bilateral V1 (Figure 4E), showed stronger activation when participants were allowed to freely move their eyes. These results allowed us to then examine the directional connectivity among these ROIs in the free- vs. fixed-viewing condition using DCM.

Table 1. Brain regions that showed stronger and weaker activation in the free-viewing than fixed-viewing condition.

Region	Cluster size	*t* value	*p* value	MNI coordinates		
				x	y	z
Free-viewing > fixed-viewing						
Hippocampus_L	95	6.474554	1.38E-07	-22	-24	-8
Vermis_4_5	142	5.630919	1.58E-06	6	-64	-4
Precuneus_L	874	8.541123	4.62E-10	-10	-80	48
Cerebelum_4_5_L	207	6.210346	2.96E-07	-8	-44	0
Parietal_Sup_L	909	9.285465	6.73E-11	-12	-82	46
Cerebelum_Crus1_L	34	4.57421	3.41E-05	-8	-84	-16
Fusiform_L	531	7.036311	2.81E-08	-22	-76	-12
Cerebelum_Crus1_R	23	4.181899	0.000105	20	-84	-18
Thalamus_R	56	5.945185	6.36E-07	22	-26	-2
ParaHippocampal_L	127	4.891206	1.36E-05	-30	-44	-8
Temporal_Mid_R	22	3.803112	0.000303	44	-74	22
Hippocampus_R	49	5.009638	9.66E-06	24	-28	-6
ParaHippocampal_R	190	4.751655	2.04E-05	20	-42	-10
Vermis_6	84	5.574678	1.87E-06	4	-76	-10
Cerebelum_6_L	446	6.719711	6.87E-08	-10	-80	-14
Calcarine_L	1727	8.572356	4.26E-10	-2	-86	12
Cerebelum_4_5_R	157	6.261166	2.55E-07	8	-44	0
Cerebelum_6_R	361	5.429782	2.84E-06	14	-82	-16
Fusiform_R	723	5.860109	8.14E-07	24	-72	-12
Parietal_Inf_L	219	5.319814	3.92E-06	-28	-78	42
Thalamus_L	37	4.017025	0.000167	-20	-28	-2
Occipital_Inf_L	31	4.821281	1.67E-05	-28	-78	-10
Precuneus_R	654	6.624448	9.02E-08	10	-78	48
Parietal_Sup_R	589	5.494176	2.36E-06	16	-80	48
Occipital_Mid_L	1169	7.514287	7.41E-09	-18	-94	16
Occipital_Sup_R	979	7.963541	2.17E-09	18	-84	22
Cuneus_R	1199	8.17195	1.23E-09	18	-84	24
Occipital_Mid_R	843	6.131649	3.71E-07	34	-82	26
Lingual_L	1786	8.766879	2.56E-10	-18	-60	2
Cuneus_L	1145	9.048055	1.24E-10	-2	-88	20
Calcarine_R	1546	8.499772	5.16E-10	4	-78	4
Lingual_R	1830	8.820668	2.22E-10	10	-60	2
Occipital_Sup_L	1025	8.872018	1.95E-10	-14	-82	44
Precentral_R	123	4.12053	0.000125	30	0	50
Frontal_Mid_R	135	4.373616	6.07E-05	46	0	56
Frontal_Sup_R	111	4.006374	0.000172	28	2	54

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Precentral_L	230	3.988691	0.000181	-28	-6	48
Frontal_Mid_L	72	4.03974	0.000157	-30	-4	64
Frontal_Sup_L	147	4.208813	9.71E-05	-30	-4	66
Temporal_Mid_L	32	3.603136	0.000526	-56	-56	-4
Temporal_Inf_L	53	4.053346	0.000151	-56	-58	-8
Supp_Motor_Area_L	134	3.76635	0.000336	-6	-4	66
Supp_Motor_Area_R	150	3.975854	0.000187	12	4	70
Temporal_Sup_L	29	3.622045	0.0005	-46	-12	-12
Temporal_Mid_L	26	3.854475	0.000263	-46	-12	-14
Frontal_Mid_L	11	3.357805	0.00102	-46	40	24
Frontal_Inf_Tri_L	11	3.670721	0.000437	-50	34	24
Putamen_L	27	3.288356	0.001226	-20	12	6
Putamen_L	15	3.149294	0.001766	-18	16	-10
ParaHippocampal_L	12	3.193907	0.001572	-28	-26	-22
Fixed-viewing > free-viewing						
Occipital_Mid_L	81	6.07279	4.4E-07	-30	-96	-6
Occipital_Inf_L	105	6.833785	4.97E-08	-28	-94	-10
Occipital_Mid_R	46	4.263794	8.3E-05	34	-94	0
Occipital_Inf_R	205	6.103413	4.02E-07	38	-88	-10
Precuneus_R	21	4.971969	1.08E-05	22	-44	10
Cingulum_Post_R	11	3.737836	0.000363	12	-38	14
Parietal_Inf_R	185	5.23807	4.97E-06	50	-58	50
Angular_R	145	4.952644	1.14E-05	48	-60	50
Parietal_Inf_L	26	3.832842	0.000279	-52	-54	48
Angular_L	16	3.75408	0.000348	-52	-64	36
Frontal_Mid_Orb_R	15	3.343355	0.00106	38	42	-8
Frontal_Inf_Orb_R	17	3.441299	0.000815	42	44	-10
Frontal_Mid_R	10	3.153018	0.001749	40	52	0
Frontal_Mid_Orb_R	28	3.372491	0.000981	38	54	-4
Cingulum_Mid_R	15	3.282947	0.001244	2	-24	36
Frontal_Mid_R	50	3.256237	0.001335	44	22	40
Temporal_Mid_R	14	3.205742	0.001524	44	-70	0

Note: All clusters survived the threshold of p <.005, with 10 voxel extension (no correction). The names of the anatomical areas in the table were obtained using the AAL toolbox for SPM12 and follow the automated anatomical labeling (AAL) template naming convention (Tzourio-Mazoyer et al., 2002). R/L = right/left hemisphere.

DCM Results

Model 1: PPA – FEF – V1. As shown in Figures 5A and 5C (i.e., the *A* matrix parameter results), our DCM analyses found bidirectional excitatory intrinsic connectivity between the scene processing region PPA and early visual region V1, and between the oculomotor control region FEF and V1, during scene construction when the two viewing conditions were averaged. There was also an excitatory connection directed from FEF to PPA (in the left hemisphere) and

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

an inhibitory connection from PPA to FEF (in the right hemisphere), but no excitatory effect from PPA to FEF. However, when participants were allowed to freely move their eyes during scene construction (i.e., the free-viewing condition), the PPA showed enhanced excitatory connectivity to FEF ($B_{PPA\rightarrow FEF}$ = 0.876 and 0.582 for the left and right hemisphere, *Pp* > 95%; see green arrows in Figure 5B and 5D). Furthermore, both PPA and FEF showed enhanced excitatory effects on the early visual region V1, especially in the left hemisphere ($B_{PPA\rightarrow V1}$ = 1.039 and $B_{FEF\rightarrow V1}$ = 0.840, *Pp* > 95%). Taken together, these excitatory connections indicate an enhanced top-down information flow from the scene processing region PPA to the oculomotor control region FEF, and to the early visual region V1. In the right hemisphere, the connectivity from FEF to PPA was also enhanced when participants were allowed to freely move their eyes ($B_{FEF\rightarrow PPA}$ = 0.410, *Pp* > 95%), whereas the bottom-up effect (i.e., the directional connectivity from V1 to PPA and FEF) was weakened during the free- versus fixed-viewing condition ($B_{V1\rightarrow PPA}$ = -0.426/-0.610, $B_{V1\rightarrow FEF}$ = -0.904/-0.758 for the left/right hemisphere, *Pp* > 95%; see red arrows in Figure 5B and 5D).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

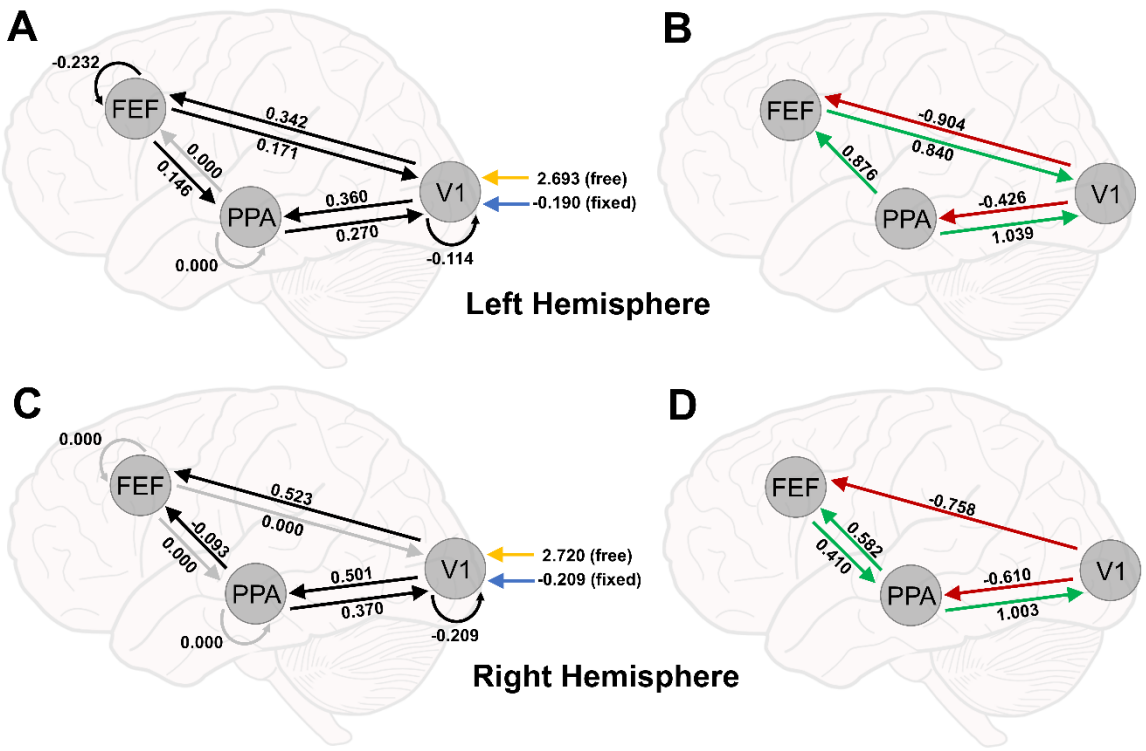

FIGURE 5: DCM results for PPA – FEF – V1. (A) Intrinsic connections for the PPA – FEF – V1 model in the left hemisphere. (B) Modulatory connections for the PPA – FEF – V1 model in the left hemisphere. (C) Intrinsic connections for the PPA – FEF – V1 model in the right hemisphere. (D) Modulatory connections for the PPA – FEF – V1 model in the right hemisphere. For (A) and (C) solid black arrows with a straight line indicate intrinsic connections that exceeded posterior probabilities (Pp) of 95%, while grey arrows indicate connections that did not exceed a 95% Pp. Straight arrows represent intrinsic connections between ROIs, and curved arrows represent the strength of self-connections within ROIs, with smaller (i.e., more negative) values indicating less self-inhibition within the ROI (i.e., long activation sustainment). The model inputs (*C* matrix) are indicated as "free" and "fixed" to V1. For (B) and (D) red arrows indicate significant negative modulatory effects and green arrows indicate positive modulatory effects (Pp >95%).

Model 2: HPC – FEF – V1. As shown in Figures 6A and 6C (i.e., the *A* matrix parameter results), we found bidirectional underlying excitatory connectivity in both hemispheres between the memory region HPC and early visual region V1, and between the oculomotor control region FEF and V1, during scene construction when the two viewing conditions were averaged. In the right hemisphere, the DCM analysis found an excitatory intrinsic connection from HPC to FEF, and an inhibitory connection from FEF to HPC. However, in the left hemisphere, there was no excitatory effect between the HPC and FEF (Figure 6A). When free-viewing was contrasted with

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the fixed-viewing condition, the HPC showed enhanced excitatory connectivity to FEF in both hemispheres ($B_{HPC \rightarrow FEF}$ = 0.646 and 0.627 for the left and right hemisphere, *Pp* > 95%; Figure 6B and 6D). Additionally, in the left hemisphere, we found an enhanced excitatory effect from FEF to HPC ($B_{FEF \rightarrow HPC}$ = 0.490, *Pp* > 95%). Although there were no enhanced excitatory effects from HPC towards V1 bilaterally, in the right hemisphere we found an enhanced excitatory connection from FEF to V1 ($B_{FEF \rightarrow V1}$ = 0.735, *Pp* > 95%). These excitatory connections indicate an enhanced top-down information flow from the memory region HPC to the oculomotor control region FEF, with enhanced information also flowing from FEF back to HPC, during free-viewing. Finally, the bottom-up effect from V1 to HPC and FEF was weakened during the free- versus fixed-viewing condition ($B_{V1 \rightarrow HPC}$ = -0.532/-0.383, $B_{V1 \rightarrow FEF}$ = -0.751/-0.834 for the left/right hemisphere, *Pp* > 95%; see red arrows in Figure 6B and 6D), consistent with the results found in the PPA – FEF – V1 model.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

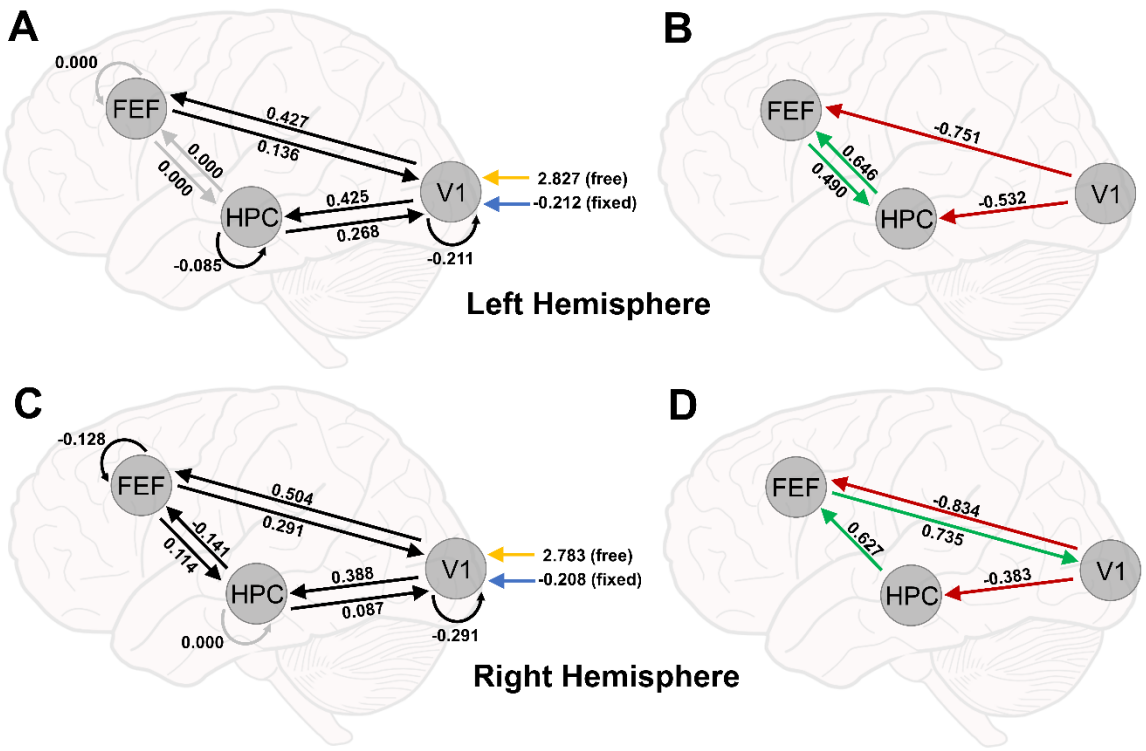

FIGURE 6: DCM results for HPC – FEF – V1. (A) Intrinsic connections for the HPC – FEF – V1 model in the left hemisphere. (B) Modulatory connections for the HPC – FEF – V1 model in the left hemisphere. (C) Intrinsic connections for the HPC – FEF – V1 model in the right hemisphere. (D) Modulatory connections for the HPC – FEF – V1 model in the right hemisphere. For (A) and (C) solid black arrows with a straight line indicate intrinsic connections that exceeded posterior probabilities (Pp) of 95%, while grey arrows indicate connections that did not exceed a 95% Pp. Straight arrows represent intrinsic connections between ROIs, and curved arrows represent the strength of self-connections within ROIs, with smaller (i.e., more negative) values indicating less self-inhibition within the ROI (i.e., long activation sustainment). The model inputs (C matrix) are indicated as “free” and “fixed” to V1. For (B) and (D) red arrows indicate significant negative modulatory effects and green arrows indicate positive modulatory effects (Pp >95%).

Model 3: LGN – V1 – FEF. When the two viewing conditions were averaged during scene construction, we found bidirectional intrinsic connectivity between V1 and LGN in both hemispheres (Figure 7A and 7C, i.e., the *A* matrix parameter results). Here, the connection from LGN to V1 was excitatory, whereas the connection from V1 to LGN was inhibitory. Additionally, we found an inhibitory intrinsic connection from FEF to V1 in both hemispheres, as well as an excitatory connection from V1 to FEF in the left hemisphere. When free-viewing was contrasted

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

with the fixed-viewing condition, the bottom-up influence from the LGN, to V1, and to FEF was weakened ($B_{LGN\rightarrow V1}$ = -1.309/-1.380, $B_{V1\rightarrow FEF}$ = -1.001/-1.176 for the left/right hemisphere, *Pp* > 95%; see red arrows in Figure 7B and 7D). In both hemispheres, the top-down influence from the FEF to V1 was strengthened ($B_{FEF\rightarrow V1}$ = 1.057 and 1.136 for the left and right hemisphere, *Pp* > 95%), as found in the previous models. Interestingly, the modulation effect from V1 to LGN was inhibitory ($B_{V1\rightarrow LGN}$ = -0.777 and -0.816 for the left and right hemisphere, *Pp* > 95%), and the input (mainly from the fixed-viewing trials) had a negative influence on LGN (the blue upward arrow in Figure 7A and 7C, i.e., the C matrix). Taken together, these results indicate that when participants could freely move their eyes during scene construction, the bottom-up information flow was inhibited, and although the enhanced top-down influence occurred, it stopped at the earliest cortical region V1, i.e., did not extend to the thalamic region LGN.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

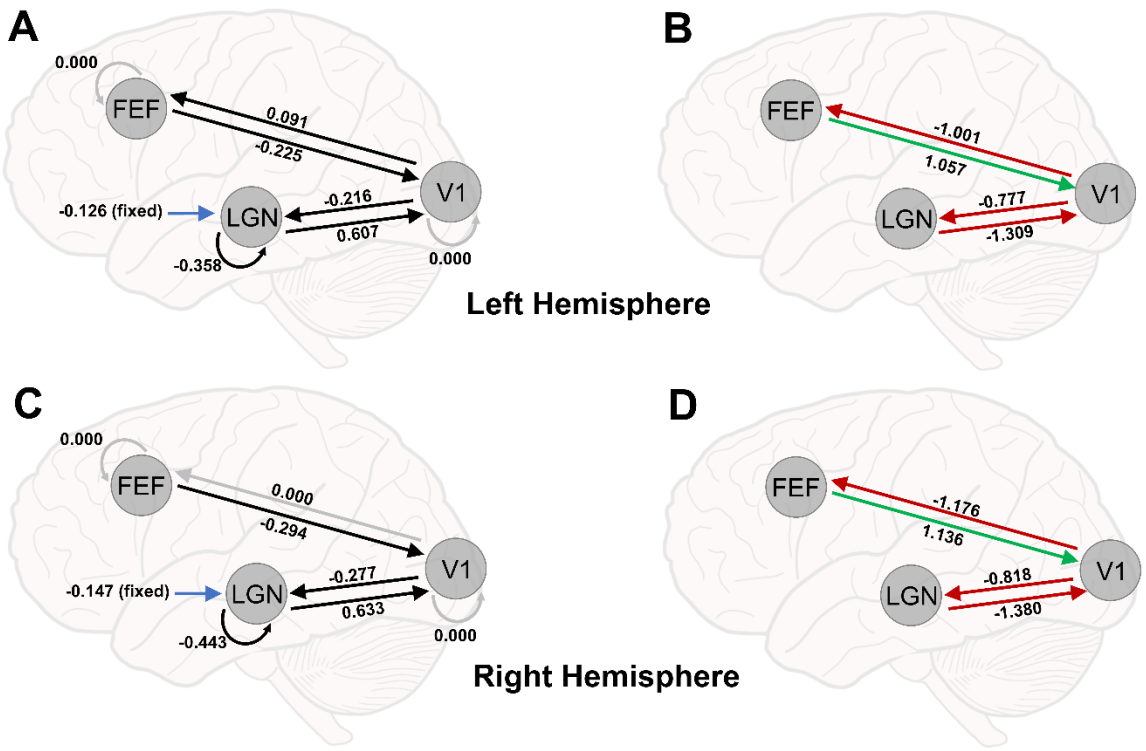

FIGURE 7: DCM results for LGN – V1 – FEF. (A) Intrinsic connections for the LGN – V1 – FEF model in the left hemisphere. (B) Modulatory connections for the LGN – V1 – FEF model in the left hemisphere. (C) Intrinsic connections for the LGN – V1 – FEF model in the right hemisphere. (D) Modulatory connections for the LGN – V1 – FEF model in the right hemisphere. For (A) and (C) solid black arrows with a straight line indicate intrinsic connections that exceeded posterior probabilities (Pp) of 95%, while grey arrows indicate connections that did not exceed a 95% Pp. Straight arrows represent intrinsic connections between ROIs, and curved arrows represent the strength of self-connections within ROIs, with smaller (i.e., more negative) values indicating less self-inhibition within the ROI (i.e., long activation sustainment). The significant model input (*C* matrix) for fixed-viewing is indicated as "fixed" to LGN. For (B) and (D) red arrows indicate significant negative modulatory effects and green arrows indicate positive modulatory effects (Pp >95%).

Discussion

Separate lines of evidence have pointed to a role for the MTL memory system as well the oculomotor system in the mental construction of scenes (Mirza et al., 2016; Mullally et al., 2012; Parr & Friston, 2017; Pearson, 2019). Recent computational modeling evidence has shown, using simulated stimulation, that information may rapidly flow from the MTL to regions of the oculomotor system (Ryan, Shen et al., 2020); thereby highlighting an intimate relationship

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

between the two systems (see Ryan, Shen, & Liu, 2020). However, to date, the information flow between the MTL and oculomotor system, and with early visual cortex, for the mental construction of scenes has not been investigated. The current study investigated this relationship at a mechanistic level using a scene construction task in which participants' eye movements were either restricted or unrestricted. Since our task did not include external visual input, we were able to determine the directionality of neural connectivity using DCM between key regions of these systems while participants constructed scene representations. As opposed to when eye movements were restricted (*fixed-viewing*), allowing free eye movements during construction of novel scenes strengthened top-down connections from the MTL to oculomotor regions, and to lower-level cortical visual processing regions, and suppressed bottom-up connections along the visual stream from LGN towards V1 and FEF. Moreover, vividness of imagined scenes was rated higher during free-viewing over fixed-viewing. Taken together, these findings provide novel, non-invasive evidence for the causal architecture between the MTL memory and oculomotor systems associated with constructing vivid mental representations of scenes.

The PPA and HPC have well-established roles in scene processing (Bird et al., 2010; Boccia et al., 2017; Douglas et al., 2017). Given that previous work has demonstrated a positive association between MTL activity and visual exploration (Liu et al., 2017, 2020), and has outlined considerable anatomical connectivity between the MTL memory and oculomotor systems (Ryan, Shen, Kacollja, et al., 2020; Shen et al., 2016), we predicted that neural activity in the PPA and HPC would be stronger with unrestricted over restricted eye movements. Indeed, both left and right PPA, and left HPC, were more strongly engaged when participants imagined novel scenes in the free-viewing versus fixed-viewing trials. Extensive research has confirmed functional coupling of the PPA and HPC (Baldassano et al., 2013; Sulpizio et al., 2016); here, the current findings highlight the importance of incorporating functional connectivity with oculomotor regions into models of scene construction alongside the MTL. Critically, we

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

revealed a positive modulatory (i.e., excitatory) effect in the free-viewing condition from the MTL towards the FEF. In both hemispheres, the modulatory connections from the HPC and PPA towards FEF were strengthened during free-viewing trials. Conversely, our DCM results revealed that the HPC and PPA received excitatory effects from the FEF when eye movements were unrestricted, which may explain stronger engagement of these MTL regions in the free-viewing condition. Together, these results are highly compatible with the notion that the MTL and oculomotor systems interact in a reciprocal manner (Ryan, Shen, & Liu, 2020), such that information from memory may guide oculomotor behavior (Meister & Buffalo, 2016; Voss et al., 2017), and eye movements may support updating of ongoing scene construction, even in the absence of external visual input (Ringo et al., 1994).

Similar to the HPC and PPA, the FEF showed stronger activation in the free-viewing versus fixed-viewing condition. Extensive evidence has linked the FEF with the cognitive control of eye movements, in both nonhuman animals and humans (Hanes et al., 1998; Mirpour et al., 2018; Robinson & Fuchs, 1969; Schiller et al., 1979; Selvanayagam et al., 2019). The current results add empirical support to the idea that eye movements may have a critical role in the construction of spatiotemporal content, and may promote vivid (re)experiencing (Ryan, Shen, & Liu, 2020; Wynn et al., 2019). Although previous modeling studies described the anatomical connections between the HPC and FEF in primate models (Shen et al., 2016), and simulated information flow between the regions (Ryan, Shen, et al., 2020), the current study provides the first functional evidence in humans that MTL regions directly “talk” to oculomotor control regions, and specifically, here, to facilitate vivid scene construction.

Our DCM results provide novel evidence for the directionality of information flow along the visual stream during scene construction, namely top-down flow from PPA towards V1. Functional connectivity within the scene processing network between PPA and peripheral V1 has been shown to develop in humans as early as 27 days old (Kamps et al., 2020) and may reflect maintained retinotopic organization along the visual stream (Huang & Sereno, 2013).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Previous DCM studies have shown strengthened top-down modulatory effects during imagination compared to perception from fronto-parietal regions to early visual regions (Dentico et al., 2014; Dijkstra et al., 2017), but, to our knowledge, no study to date has extended these findings to include MTL memory regions. Moreover, our findings demonstrate that top-down information flow during scene construction may be similar to information flow during mnemonic reconstructive processes (i.e., from HPC towards lower-level visual regions), in reverse of perceptual processes where information flow is thought to originate in early visual regions towards the MTL (Linde-Domingo et al., 2019). Consistent with this interpretation, we found an inhibitory influence in the free-viewing condition from V1 to FEF and PPA/HPC in both hemispheres, reflecting suppression in intrinsic baseline connections from early visual regions towards the MTL. This suggests that MTL activation is internally initiated and maintained during imagination (Campbell et al., 2018), and inhibition of bottom-up information flow may help to avoid external visual distractions, facilitating scene construction (Benedek et al., 2016; Daselaar et al., 2010).

Although there was limited visual input involved in our task, regions in the early visual pathway (V1 and LGN) showed stronger activation in the free-viewing versus fixed-viewing condition. While V1 activation in the absence of external input is not an intuitive finding given this region's role in active vision (Hubel, 1982), it is certainly not a new one (Kosslyn et al., 1995; Miyashita, 1995). In a large-scale meta-analysis of studies involving visual imagery with human participants, Winlove et al. (2018) found consistent activation of FEF and V1; our DCM results demonstrate that this directional relationship is strengthened with unrestricted viewing behaviors. Thus, the FEF, likely in tandem with other oculomotor control regions like the parietal lobe (Rafal, 2006), plays a key role in translating viewing-relevant information of mental representations to eye movement behavior. Additionally, top-down activation from higher-level memory regions to lower-level perceptual regions during mnemonic retrieval has been found in the memory literature (Linde-Domingo et al., 2019; Naya et al., 2001). In rodents, visual regions

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

as early as V1 may contain spatial information associated with that supported by the HPC (Fournier et al., 2020). Our DCM results further suggest that V1 activation can be driven by two top-down pathways: one, along the ventral visual pathway (from PPA to V1) which likely originates from the HPC, and two, through the interaction between the MTL and dorsal oculomotor control system (from HPC through FEF to V1). The specific function of the two influence pathways should be interrogated in future investigations.

The role of the LGN in unrestricted eye movements was of particular interest in the present study, as there is considerable debate over this region's role in the construction of scene representations, and how this region may be connected to later visual regions in service of visual imagery (Lesica et al., 2006; Tadmor & Tolhurst, 2000). Previous work suggests that the LGN is primarily related to saccadic control; namely, that saccades in darkness lead to enhanced activity in the LGN, whereas saccades made during strong visual stimulation suppress activity (Sylvester et al., 2005). Similar to modulation effects from V1 to oculomotor control regions, our DCM results showed that the bottom-up influence from LGN towards V1 was inhibited. More importantly, top-down enhancement from the FEF to V1 was not extended to LGN. These results suggest that, in contrast to V1, stronger involvement of LGN in the free- versus fixed-viewing condition was not due to top-down excitations when eye movements were unrestricted. Instead, it was due to a suppression effect from the fixed-viewing condition. Therefore, different LGN involvement in the two conditions may be primarily related to the eye movement manipulation during the task *per se*, rather than mental construction of scenes.

We note that because participants were asked in the fixed-viewing condition to focus on a specific spot while trying to construct the scenes, it could be argued that the present results were simply due to instructions that may have led to increased working memory or executive control demands. However, we think this is unlikely for the following reasons. First, previous work has shown that maintaining central fixation does not appear to increase working memory demands (Armson et al., 2019), perhaps because, even under such instructions, participants

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

still move their eyes, albeit with fewer fixations and smaller saccades. Second, here, the fixed-viewing condition did not show obvious differential engagement of the brain regions that support executive control (except for FEF), compared to the free-viewing condition (see Figure 4B). Moreover, the finding of lower vividness ratings for the mental construction of scenes during fixed-viewing compared to free-viewing, are consistent with the neural data showing that the top-down connectivity from HPC/PPA/FEF to lower visual regions was enhanced and the reversed bottom-up connectivity was weakened. Therefore, maintaining fixations likely negatively impacted scene construction due to reduced ability to translate mental representations of novel scenes to viewing-relevant behaviour, rather than due to an increased demand on working memory capacity.

To summarize, in the present study, we successfully applied DCM to investigate the causal interactions between the MTL memory and oculomotor systems in support of scene construction. Our findings provide strong support for a top-down influence from the MTL to oculomotor control region FEF and to early cortical, but not subcortical, visual regions, and an inhibitory bottom-up modulatory effect of visual exploration from LGN to V1 and FEF when a mental scene representation was constructed. More generally, this work demonstrates how the MTL may guide eye movements to support vivid, experiential phenomena during imagination and recollection. Eye movements, as such, may be a natural effector system for memory (Ryan & Shen, 2020), critical in mental imagery of scenes.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acknowledgements: We would like to thank Ryan Aloysius, Arber Kacollja, Ling Li, and Mandy Ding for their help at different stages of this research project. All processed data associated with the reported results and analysis scripts are available from NLW and ZXL upon reasonable requests. This work was supported by a Vision: Science to Applications (VISTA) postdoctoral fellowship awarded to ZXL, National Sciences and Engineering Research Council - Alexander Graham Bell Alexander Graham Bell Canada Graduate Scholarship (524813) to NLW, and funding from the Natural Sciences and Engineering Research Council of Canada awarded to JDR (RGPIN-2018-06399).

Conflict of interest: The authors declare no competing financial interest.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

Armson, M. J., Ryan, J. D., & Levine, B. (2019). Maintaining fixation does not increase demands on working memory relative to free viewing. *PeerJ*, *7*, e6839. https://doi.org/10.7717/peerj.6839

Bahcall, D. O., & Kowler, E. (2000). The control of saccadic adaptation: Implications for the scanning of natural visual scenes. *Vision Research*, *40*(20), 2779–2796. https://doi.org/10.1016/S0042-6989(00)00117-6

Baldassano, C., Beck, D. M., & Fei-Fei, L. (2013). Differential connectivity within the Parahippocampal Place Area. *NeuroImage*, *75*, 228–237. https://doi.org/10.1016/j.neuroimage.2013.02.073

Benedek, M., Jauk, E., Beaty, R. E., Fink, A., Koschutnig, K., & Neubauer, A. C. (2016). Brain mechanisms associated with internally directed attention and self-generated thought. *Scientific Reports*, *6*(1), 22959. https://doi.org/10.1038/srep22959

Bird, C. M., Capponi, C., King, J. A., Doeller, C. F., & Burgess, N. (2010). Establishing the Boundaries: The Hippocampal Contribution to Imagining Scenes. *Journal of Neuroscience*, *30*(35), 11688–11695. https://doi.org/10.1523/JNEUROSCI.0723-10.2010

Boccia, M., Sulpizio, V., Palermo, L., Piccardi, L., Guariglia, C., & Galati, G. (2017). I can see where you would be: Patterns of fMRI activity reveal imagined landmarks. *NeuroImage*, *144*, 174–182. https://doi.org/10.1016/j.neuroimage.2016.08.034

Campbell, K. L., Madore, K. P., Benoit, R. G., Thakral, P. P., & Schacter, D. L. (2018). Increased hippocampus to ventromedial prefrontal connectivity during the construction of episodic future events. *Hippocampus*, *28*(2), 76–80. https://doi.org/10.1002/hipo.22812

Castelhano, M. S., & Heaven, C. (2011). Scene context influences without scene gist: Eye movements guided by spatial associations in visual search. *Psychonomic Bulletin & Review*, *18*(5), 890–896. https://doi.org/10.3758/s13423-011-0107-8

Conti, F., & Irish, M. (2021). Harnessing Visual Imagery and Oculomotor Behaviour to

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Understand Prospection. *Trends in Cognitive Sciences*, *25*(4), 272–283. https://doi.org/10.1016/j.tics.2021.01.009

Daselaar, S. M., Porat, Y., Huijbers, W., & Pennartz, C. M. A. (2010). Modality-specific and modality-independent components of the human imagery system. *NeuroImage*, *52*(2), 677–685. https://doi.org/10.1016/j.neuroimage.2010.04.239

Dentico, D., Cheung, B. L., Chang, J.-Y., Guokas, J., Boly, M., Tononi, G., & Van Veen, B. (2014). Reversal of cortical information flow during visual imagery as compared to visual perception. *NeuroImage*, *100*, 237–243. https://doi.org/10.1016/j.neuroimage.2014.05.081

Dijkstra, N., Zeidman, P., Ondobaka, S., van Gerven, M. A. J., & Friston, K. (2017). Distinct Top-down and Bottom-up Brain Connectivity During Visual Perception and Imagery. *Scientific Reports*, *7*(1), 5677. https://doi.org/10.1038/s41598-017-05888-8

Donner, T., Kettermann, A., Diesch, E., Ostendorf, F., Villringer, A., & Brandt, S. A. (2000). Involvement of the human frontal eye field and multiple parietal areas in covert visual selection during conjunction search. *European Journal of Neuroscience*, *12*(9), 3407–3414. https://doi.org/10.1046/j.1460-9568.2000.00223.x

Douglas, D., Thavabalasingam, S., Chorghay, Z., O'Neil, E. B., Barense, M. D., & Lee, A. C. H. (2017). Perception of Impossible Scenes Reveals Differential Hippocampal and Parahippocampal Place Area Contributions to Spatial Coherency. *Hippocampus*, *27*(1), 61–76. https://doi.org/10.1002/hipo.22673

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

Fischl, B. (2012). FreeSurfer. *NeuroImage*, *62*(2), 774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021

Fournier, J., Saleem, A. B., Diamanti, E. M., Wells, M. J., Harris, K. D., & Carandini, M. (2020). Mouse Visual Cortex Is Modulated by Distance Traveled and by Theta Oscillations.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Current Biology, *30*(19), 3811-3817.e6. https://doi.org/10.1016/j.cub.2020.07.006

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, *19*(4), 1273–1302. https://doi.org/10.1016/s1053-8119(03)00202-7

Friston, K. J., Kahan, J., Biswal, B., & Razi, A. (2014). A DCM for resting state fMRI. *NeuroImage*, *94*, 396–407. https://doi.org/10.1016/j.neuroimage.2013.12.009

Gelman, A., & Tuerlinckx, F. (2000). Type S error rates for classical and Bayesian single and multiple comparison procedures. *Computational Statistics*, *15*(3), 373–390. https://doi.org/10.1007/s001800000040

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*, *14*(5), 350–363. https://doi.org/10.1038/nrn3476

Gurtner, L. M., Hartmann, M., & Mast, F. W. (2021). Eye movements during visual imagery and perception show spatial correspondence but have unique temporal signatures. *Cognition*, *210*, 104597. https://doi.org/10.1016/j.cognition.2021.104597

Hanes, D. P., Patterson, W. F., & Schall, J. D. (1998). Role of Frontal Eye Fields in Countermanding Saccades: Visual, Movement, and Fixation Activity. *Journal of Neurophysiology*, *79*(2), 817–834. https://doi.org/10.1152/jn.1998.79.2.817

Hassabis, D., Kumaran, D., & Maguire, E. A. (2007). Using Imagination to Understand the Neural Basis of Episodic Memory. *Journal of Neuroscience*, *27*(52), 14365–14374. https://doi.org/10.1523/JNEUROSCI.4549-07.2007

Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in Cognitive Sciences*, *11*(7), 299–306. https://doi.org/10.1016/j.tics.2007.05.001

Huang, R.-S., & Sereno, M. I. (2013). Bottom-up Retinotopic Organization Supports Top-down Mental Imagery. *The Open Neuroimaging Journal*, *7*, 58–67. https://doi.org/10.2174/1874440001307010058

Hubel, D. H. (1982). Exploration of the primary visual cortex, 1955–78. *Nature*, *299*(5883), 515–524. https://doi.org/10.1038/299515a0

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Johansson, R., Holsanova, J., & Holmqvist, K. (2006). Pictures and Spoken Descriptions Elicit Similar Eye Movements During Mental Imagery, Both in Light and in Complete Darkness. *Cognitive Science*, *30*(6), 1053–1079. https://doi.org/10.1207/s15516709cog0000_86

Kamps, F. S., Hendrix, C. L., Brennan, P. A., & Dilks, D. D. (2020). Connectivity at the origins of domain specificity in the cortical face and place networks. *Proceedings of the National Academy of Sciences*, *117*(11), 6163–6169. https://doi.org/10.1073/pnas.1911359117

Kashihara, K. (2020). Microsaccadic modulation evoked by emotional events. *Journal of Physiological Anthropology*, *39*(1), 26. https://doi.org/10.1186/s40101-020-00238-6

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.2307/2291091

Kosslyn, S. M., Thompson, W. L., Klm, I. J., & Alpert, N. M. (1995). Topographical representations of mental images in primary visual cortex. *Nature*, *378*(6556), 496–498. https://doi.org/10.1038/378496a0

Kowler, E. (2011). Eye movements: The past 25 years. *Vision Research*, *51*(13), 1457–1483. https://doi.org/10.1016/j.visres.2010.12.014

Lesica, N. A., Weng, C., Jin, J., Yeh, C.-I., Alonso, J.-M., & Stanley, G. B. (2006). Dynamic Encoding of Natural Luminance Sequences by LGN Bursts. *PLOS Biology*, *4*(7), e209. https://doi.org/10.1371/journal.pbio.0040209

Lieberman, M. D., & Cunningham, W. A. (2009). Type I and Type II error concerns in fMRI research: Re-balancing the scale. *Social Cognitive and Affective Neuroscience*, *4*(4), 423–428. https://doi.org/10.1093/scan/nsp052

Linde-Domingo, J., Treder, M. S., Kerrén, C., & Wimber, M. (2019). Evidence that neural information flow is reversed between object perception and object reconstruction from memory. *Nature Communications*, *10*(1), 179. https://doi.org/10.1038/s41467-018-08080-2

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Liu, Z.-X., Rosenbaum, R. S., & Ryan, J. D. (2020). Restricting Visual Exploration Directly Impedes Neural Activity, Functional Connectivity, and Memory. *Cerebral Cortex Communications*, *1*(tgaa054). https://doi.org/10.1093/texcom/tgaa054

Liu, Z.-X., Shen, K., Olsen, R. K., & Ryan, J. D. (2017). Visual Sampling Predicts Hippocampal Activity. *Journal of Neuroscience*, *37*(3), 599–609. https://doi.org/10.1523/JNEUROSCI.2610-16.2016

Maguire, E. A., & Mullally, S. L. (2013). The Hippocampus: A Manifesto for Change. *Journal of Experimental Psychology. General*, *142*(4), 1180–1189. https://doi.org/10.1037/a0033650

Meister, M. L. R., & Buffalo, E. A. (2016). Getting directions from the hippocampus: The neural connection between looking and memory. *Neurobiology of Learning and Memory*, *134*, 135–144. https://doi.org/10.1016/j.nlm.2015.12.004

Mirpour, K., Bolandnazar, Z., & Bisley, J. W. (2018). Suppression of frontal eye field neuronal responses with maintained fixation. *Proceedings of the National Academy of Sciences*, *115*(4), 804–809. https://doi.org/10.1073/pnas.1716315115

Mirza, M. B., Adams, R. A., Mathys, C. D., & Friston, K. J. (2016). Scene Construction, Visual Foraging, and Active Inference. *Frontiers in Computational Neuroscience*, *10*. https://doi.org/10.3389/fncom.2016.00056

Miyashita, Y. (1995). How the brain creates imagery: Projection to primary visual cortex. *Science*, *268*(5218), 1719–1721.

Mullally, S. L., Hassabis, D., & Maguire, E. A. (2012). Scene Construction in Amnesia: An fMRI Study. *The Journal of Neuroscience*, *32*(16), 5646–5653. https://doi.org/10.1523/JNEUROSCI.5522-11.2012

Naya, Y., Yoshida, M., & Miyashita, Y. (2001). Backward Spreading of Memory-Retrieval Signal in the Primate Temporal Cortex. *Science*, *291*(5504), 661–664. https://doi.org/10.1126/science.291.5504.661

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Parr, T., & Friston, K. J. (2017). The active construction of the visual world. *Neuropsychologia*, *104*, 92–101. https://doi.org/10.1016/j.neuropsychologia.2017.08.003

Pearson, J. (2019). The human imagination: The cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience*, *20*(10), 624–634. https://doi.org/10.1038/s41583-019-0202-9

Rafal, R. D. (2006). Oculomotor Functions of the Parietal Lobe: Effects of Chronic Lesions in Humans. *Cortex*, *42*(5), 730–739. https://doi.org/10.1016/S0010-9452(08)70411-8

Ringo, J. L., Sobotka, S., Diltz, M. D., & Bunce, C. M. (1994). Eye movements modulate activity in hippocampal, parahippocampal, and inferotemporal neurons. *Journal of Neurophysiology*, *71*(3), 1285–1288. https://doi.org/10.1152/jn.1994.71.3.1285

Robin, J. H., Wynn, J. S., & Moscovitch, M. (2016). The spatial scaffold: The effects of spatial context on memory for events. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *42*(2), 308–315. https://doi.org/10.1037/xlm0000167

Robinson, D. A., & Fuchs, A. F. (1969). Eye movements evoked by stimulation of frontal eye fields. *Journal of Neurophysiology*, *32*(5), 637–648. https://doi.org/10.1152/jn.1969.32.5.637

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

Ryan, J. D., Shen, K., Kacollja, A., Tian, H., Griffiths, J., Bezgin, G., & McIntosh, A. R. (2020). Modeling the influence of the hippocampal memory system on the oculomotor system. *Network Neuroscience*, *4*(1), 217–233. https://doi.org/10.1162/netn_a_00120

Ryan, J. D., Shen, K., & Liu, Z. (2020). The intersection between the oculomotor and hippocampal memory systems: Empirical developments and clinical implications. *Annals of the New York Academy of Sciences*, *1464*(1), 115–141. https://doi.org/10.1111/nyas.14256

Schiller, P. H., True, S. D., & Conway, J. L. (1979). Effects of frontal eye field and superior

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

colliculus ablations on eye movements. *Science*, *206*(4418), 590–592. https://doi.org/10.1126/science.115091

Schuyler, B., Ollinger, J. M., Oakes, T. R., Johnstone, T., & Davidson, R. J. (2010). Dynamic Causal Modeling applied to fMRI data shows high reliability. *NeuroImage*, *49*(1), 603–611. https://doi.org/10.1016/j.neuroimage.2009.07.015

Selvanayagam, J., Johnston, K. D., Schaeffer, D. J., Hayrynen, L. K., & Everling, S. (2019). Functional Localization of the Frontal Eye Fields in the Common Marmoset Using Microstimulation. *Journal of Neuroscience*, *39*(46), 9197–9206. https://doi.org/10.1523/JNEUROSCI.1786-19.2019

Shen, K., Bezgin, G., Selvam, R., McIntosh, A. R., & Ryan, J. D. (2016). An Anatomical Interface between Memory and Oculomotor Systems. *Journal of Cognitive Neuroscience*, *28*(11), 1772–1783. https://doi.org/10.1162/jocn_a_01007

Spivey, M. J., Richardson, D. C., Tyler, M. J., & Young, E. E. (2000). *Eye Movements During Comprehension of Spoken Scene Descriptions*. 6.

Stephan, K. E., & Friston, K. J. (2010). Analyzing effective connectivity with fMRI. *Wiley Interdisciplinary Reviews. Cognitive Science*, *1*(3), 446–459. https://doi.org/10.1002/wcs.58

Stephan, K. E., Harrison, L. M., Kiebel, S. J., David, O., Penny, W. D., & Friston, K. J. (2007). Dynamic causal models of neural system dynamics: Current state and future extensions. *Journal of Biosciences*, *32*(1), 129–144. https://doi.org/10.1007/s12038-007-0012-5

Stephan, K. E., Penny, W. D., Moran, R. J., den Ouden, H. E. M., Daunizeau, J., & Friston, K. J. (2010). Ten simple rules for dynamic causal modeling. *Neuroimage*, *49*(4), 3099–3109. https://doi.org/10.1016/j.neuroimage.2009.11.015

Sulpizio, V., Boccia, M., Guariglia, C., & Galati, G. (2016). Functional connectivity between posterior hippocampus and retrosplenial complex predicts individual differences in navigational ability. *Hippocampus*, *26*(7), 841–847. https://doi.org/10.1002/hipo.22592

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Sylvester, R., Haynes, J.-D., & Rees, G. (2005). Saccades Differentially Modulate Human LGN and V1 Responses in the Presence and Absence of Visual Stimulation. *Current Biology*, *15*(1), 37–41. https://doi.org/10.1016/j.cub.2004.12.061

Tadmor, Y., & Tolhurst, D. J. (2000). Calculating the contrasts that retinal ganglion cells and LGN neurones encounter in natural scenes. *Vision Research*, *40*(22), 3145–3157. https://doi.org/10.1016/S0042-6989(00)00166-8

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *NeuroImage*, *15*(1), 273–289. https://doi.org/10.1006/nimg.2001.0978

Veniero, D., Gross, J., Morand, S., Duecker, F., Sack, A. T., & Thut, G. (2021). Top-down control of visual cortex by the frontal eye fields through oscillatory realignment. *Nature Communications*, *12*(1), 1757. https://doi.org/10.1038/s41467-021-21979-7

Vernet, M., Quentin, R., Chanes, L., Mitsumasu, A., & Valero-Cabré, A. (2014). Frontal eye field, where art thou? Anatomy, function, and non-invasive manipulation of frontal regions involved in eye movements and associated cognitive operations. *Frontiers in Integrative Neuroscience*, *8*. https://doi.org/10.3389/fnint.2014.00066

Voss, J. L., Bridge, D. J., Cohen, N. J., & Walker, J. A. (2017). A Closer Look at the Hippocampus and Memory. *Trends in Cognitive Sciences*, *21*(8), 577–588. https://doi.org/10.1016/j.tics.2017.05.008

Winlove, C. I. P., Milton, F., Ranson, J., Fulford, J., MacKisack, M., Macpherson, F., & Zeman, A. (2018). The neural correlates of visual imagery: A co-ordinate-based meta-analysis. *Cortex*, *105*, 4–25. https://doi.org/10.1016/j.cortex.2017.12.014

Wynn, J. S., Shen, K., & Ryan, J. D. (2019). Eye Movements Actively Reinstate Spatiotemporal Mnemonic Content. *Vision*, *3*(2), 21. https://doi.org/10.3390/vision3020021

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.23.461606; this version posted September 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

scale automated synthesis of human functional neuroimaging data. *Nature Methods*, *8*(8), 665–670. https://doi.org/10.1038/nmeth.1635

Zeidman, P., Jafarian, A., Seghier, M. L., Litvak, V., Cagnan, H., Price, C. J., & Friston, K. J. (2019). A guide to group effective connectivity analysis, part 2: Second level analysis with PEB. *NeuroImage*, *200*, 12–25. https://doi.org/10.1016/j.neuroimage.2019.06.032

Zeidman, P., Mullally, S. L., & Maguire, E. A. (2015). Constructing, Perceiving, and Maintaining Scenes: Hippocampal Activity and Connectivity. *Cerebral Cortex*, *25*(10), 3836–3855. https://doi.org/10.1093/cercor/bhu266

Zimmermann, E., & Lappe, M. (2016). Visual Space Constructed by Saccade Motor Maps. *Frontiers in Human Neuroscience*, *10*. https://doi.org/10.3389/fnhum.2016.00225