

Building Mental Experiences: From Scenes to Events

Anna Mary Monk

Submitted for a PhD in Cognitive Neuroscience

January 2021



Supervised by

Professor Eleanor A. Maguire (Principal)
Professor Gareth R. Barnes (Subsidiary)

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

Signed:

Date: 26.03.2021

Abstract

Mental events are central to everyday cognition, be it our continuous perception of the world, recalling autobiographical memories, or imagining the future. Little is known about the fine-grained temporal dynamics of these processes. Given the apparent predominance of scene imagery across cognition, in this thesis I used magnetoencephalography to investigate whether and how activity in the hippocampus and ventromedial prefrontal cortex (vmPFC) supports the mental construction of scenes and the events to which they give rise.

In the first experiment, participants gradually imagined scenes and also closely matched non-scene arrays; this allowed me to assess whether any brain regions showed preferential responses to scene imagery. The anterior hippocampus and vmPFC were particularly engaged by the construction of scene imagery, with the vmPFC driving hippocampal activity. In the second experiment, I found that certain objects – those that were space-defining – preferentially engaged the vmPFC and superior temporal gyrus during scene construction, providing insight into how objects affect the creation of scene representations. The third experiment involved boundary extension during scene perception, permitting me to examine how single scenes might be prepared for inclusion into events. I observed changes in evoked responses just 12.5-58 ms after scene onset over fronto-temporal sensors, with again the vmPFC exerting a driving influence on other brain regions, including the hippocampus. In the final experiment, participants watched brief movies of events built from a series of scenes or non-scene patterns. A difference in evoked responses between the two event types emerged during the first frame of the movies, the primary source of which was shown to be the hippocampus.

The enduring theme of the results across experiments was scene-specific engagement of the hippocampus and vmPFC, with the latter being the driving influence. Overall, this thesis provides insights into the neural dynamics of how scenes are built, made ready for inclusion into unfolding mental episodes, and then linked to produce our seamless experience of the world.

Impact statement

Most of us perceive the world as a series of visual scenes that are punctuated by eye blinks and saccades. We only have to open our eyes to recognise the prominence of scenes in our daily experience of the world. Somehow, these scenes become linked together, such that we have a sense of a seamless unfolding of life and events. This can be likened to how the individual frames of a movie are imperceptible, and appear to us as though a continuous flow of experience.

Scene imagery and mental events are central to cognition – they support our ability to recall our personal past, and to simulate and plan for the future, underwriting our capacity for autonomy. The importance of these cognitive processes is starkly laid bare by patients who lose such abilities. They are often stuck in the present, with even their day- and sleep-dreaming compromised. Two brain regions have been particularly implicated in supporting scene and event processing – the hippocampus and the ventromedial prefrontal cortex. These areas are often compromised by neurological and psychiatric conditions, including dementia. If we can understand the exact roles played by these regions, how they interact, and the neural mechanisms involved in scene and event processing, then in the future we may be able to devise more principled interventions to manage or alleviate memory and associated cognitive impairments, and to test the efficacy of treatment regimens.

Such an understanding could also have impacts beyond the clinical domain. Even within the healthy population, there are marked individual differences in the ability to imagine scenes, recall the past and imagine the future, but we do not know why. Elucidating the mechanisms involved in these processes could help to explain why this variability occurs, informing in turn efforts to improve, for example, memory. Similarly, a fundamental grasp of how memory and future planning works could influence educationalists in how they tailor and maximise efforts to help students reach their full potential.

The work in this thesis contributes to establishing foundational knowledge about the brain's ability to mentally construct scene imagery and event representations. This is because there are not many magnetoencephalography (MEG) studies of these cognitive processes, resulting in a dearth of knowledge about the neural dynamics involved. I have

demonstrated here the value of using MEG to study complex processes like imagination of scenes and construction of evolving events. Clearly, much more remains to be understood. I hope this work will serve as a stepping stone for further discoveries about the neural basis of cognitive functions that are so central to our everyday mental life.

Acknowledgements

There are many people I am grateful to for their support throughout my PhD. It is sad to not be able to express this in person due to the Covid-19 pandemic; I hope the opportunity to put that right will not be delayed for too much longer.

Thank you to my various colleagues in the Maguire lab, old and new. Most notably, my PhD supervisor Eleanor Maguire. I frankly could not have asked for a more formidable guide during this PhD journey. Despite how incredibly busy you are, somehow you make time for everyone in the lab. Thank you for investing so much energy in overseeing my work, providing feedback, and encouraging me in the career I am now pursuing. My secondary supervisor, Gareth Barnes, thank you for your guidance and support as I learnt about MEG, and for your infectious enthusiasm. My dear friends, Daniel Barry, Marshall Dalton, Victoria Hotchin, and Zita Patai who, despite moving to pastures new during the course of my PhD, and separated by a global pandemic, have continued to extend their advice and insightful comments on my work in the last 4 years. David Bradbury, Daniel Bates, and everyone else in the Imaging Support team, for your uplifting personalities and good humour down in the MEG basement; your combined efforts ensured any major hiccups during scanning were averted. I am also grateful to the IT department: ‘Ric Davis, Liam Reilly, and Chris Freemantle. It has been a pleasure to work with such talented, nice people. Kam Ramkissoon and Maddy Scott, you are both wonderful, brightening the atmosphere even on darker days. I am also very grateful to Pete Zeidman, whose patience and extraordinary talent for making even the most complex ideas accessible, helped me to get to grips with DCM.

Finally, thank you to my family and boyfriend. *Ai miei genitori, Francesca e Robert, avete stimolato un’enorme curiosità in me fin’ da piccola, il desiderio di saperne di più su una cosa, e di sempre impegnarmi.* My father, Robert, deserves a special mention for agreeing to be a “model” (in the loosest sense) for a day, appearing in Figures 1 and 14. John, your patience, love, and never-ending ability to raise my spirits sustained me during those long days of writing up this thesis.

Table of contents

Abstract.....	3
Impact statement	4
Acknowledgements	6
1 General introduction	12
1.1 An introduction to scene imagery.....	12
1.2 Evidence from functional magnetic resonance imaging (fMRI)	17
1.3 Evidence from neuropsychology	32
1.4 A simple model of scene and event construction.....	43
1.5 Predictions and thesis overview	46
1.6 Publications	50
2 General methods	52
2.1 Précis	52
2.2 Participants	52
2.3 Overview of the experimental set-up	53
2.4 The origin of the MEG signal	56
2.5 Instrumentation	59
2.6 MEG data pre-processing	66
2.7 MEG data analyses.....	70
2.8 MEG summary	89
3 Building scenes	90
3.1 Précis	90
3.2 Introduction	90
3.3 Materials and methods.....	95
3.4 Results	108

3.5	Discussion.....	120
4	The influence of objects on building scenes	126
4.1	Précis	126
4.2	Introduction	126
4.3	Materials and methods.....	130
4.4	Results	140
4.5	Discussion.....	147
5	Preparing scenes for inclusion into events	152
5.1	Précis	152
5.2	Introduction	153
5.3	Materials and methods.....	157
5.4	Results	165
5.5	Discussion.....	171
6	Building events.....	177
6.1	Précis	177
6.2	Introduction	178
6.3	Materials and methods.....	183
6.4	Results	195
6.5	Discussion.....	200
7	General discussion	207
7.1	Précis	207
7.2	Summary of the main results	207
7.3	The mental construction of single scenes	209
7.4	From single scenes to unfolding events	217
7.5	Outstanding questions and future directions	231
8	References	237

List of figures

Figure 1. Our experience of the world is a succession of scenes perceived each time we open our eyes	12
Figure 2. The prevalence of visual scene imagery.....	15
Figure 3. Brain regions commonly engaged during autobiographical memory recall, imagining the future, and spatial navigation.....	19
Figure 4. The hippocampus	20
Figure 5. Examples of space-defining (SD) and space-ambiguous (SA) objects	27
Figure 6. Preference for SD objects when constructing and deconstructing scenes ...	28
Figure 7. Illustration of the boundary extension (BE) phenomenon when identical scenes are presented at study and test.....	31
Figure 8. The ventromedial prefrontal cortex	35
Figure 9. The main white matter tracks connecting the hippocampus and vmPFC....	35
Figure 10. Example of bilateral hippocampal damage	38
Figure 11. Attenuated BE in patients with hippocampal damage compared to healthy controls	39
Figure 12. Example lesions of the vmPFC	42
Figure 13. BE in vmPFC-damaged patients, hippocampal-damaged patients, and control participants, including brain-damaged control patients.....	44
Figure 14. Schematic illustration of a model of scene and event construction	45
Figure 15. Hippocampal-vmPFC interactions during the construction of novel scene imagery.....	50
Figure 16. The generation of electric and magnetic fields in the brain.....	59
Figure 17. The MEG system.....	62
Figure 18. Strengths of physiological and environmental magnetic fields	62
Figure 19. Effect of third-order gradiometry on results of source reconstruction	66
Figure 20. Anatomical regions of interest.....	77

Figure 21. Distribution of MEG sensors over the head	82
Figure 22. Example of dynamic causal models involving two regions.....	87
Figure 23. MEG data pre-processing and analysis pipeline	89
Figure 24. Structure of a trial	99
Figure 25. Eye fixations and example drawings.....	110
Figure 26. MEG ROI results	116
Figure 27. Effective connectivity between the anterior hippocampus (aHipp) and ventromedial prefrontal cortex (vmPFC) during step-by-step scene and array construction	119
Figure 28. Example stimuli and trial structure	133
Figure 29. Eye fixations	141
Figure 30. MEG results	144
Figure 31. Trial timeline.....	160
Figure 32. Behavioural and MEG results.....	168
Figure 33. An example DCM.....	171
Figure 34. Experimental paradigm.....	185
Figure 35. Eye fixations	196
Figure 36. Image 1 evoked activity	199

List of tables

Table 1. Debriefing ratings	112
Table 2. MEG whole brain power changes during scenes relative to arrays at each construction stage.....	114
Table 3. Results of the surprise post-scan object recognition memory test	142

List of abbreviations

2D	two-dimensional	PHC	parahippocampal cortex
3D	three-dimensional	ROI	region-of-interest
BE	boundary extension	RFX	random effects
BMS	Bayesian Model Selection	RSC	retrosplenial cortex
DCM	dynamic causal modelling	RSVP	rapid serial visual presentation
EEG	electroencephalography	SA	space-ambiguous (objects)
EP	exceedance probability	SD	space-defining (objects)
fMRI	functional magnetic resonance imaging	Std Dev	standard deviation
FWE	family-wise error	SEM	standard error of the mean
ITI	inter-trial interval	SNR	signal-to-noise
MEG	magnetencephalography	SPM	statistical parametric mapping
MNI	Montreal Neurological Institute	SQUIDs	Superconducting Quantum Interference Devices
MRI	magnetic resonance imaging	STG	superior temporal gyrus
MSR	magnetically shielded room	vmPFC	ventromedial prefrontal cortex
MTL	medial temporal lobe	VR	virtual reality

1 General introduction

1.1 An introduction to scene imagery

Our lived experience of the world comprises a series of scenes that are perceived between the interruptions imposed by eye blinks and saccades. Each time we blink, the scene before us disappears; when we open our eyes again a new scene appears. This is undeniably how most of us experience the world – in the form of one scene after another (Figure 1).

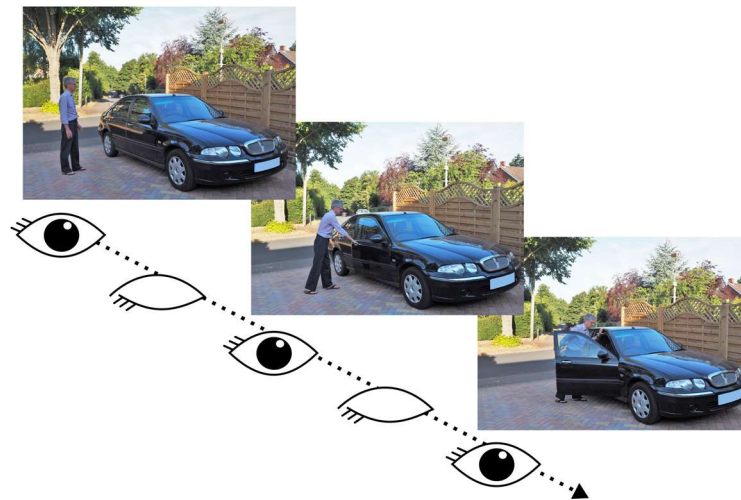


Figure 1. Our experience of the world is a succession of scenes perceived each time we open our eyes.

I define a scene as a naturalistic three-dimensional spatially-coherent representation of the world, typically populated by objects, and viewed from an egocentric perspective (Maguire and Mullally, 2013; Dalton et al., 2018). Whether they are from our ongoing experience, two-dimensional (2D) representations (such as photographs) or

three-dimensional (3D) places, scenes are contexts you could potentially step into (e.g. a forest) or operate within (e.g. a scene of the desk area in front of you). This stands in contrast to single isolated objects or landmarks against a blank background – these are not scenes and, in fact, are typically deployed in experiments as control conditions to compare against scenes (Hassabis et al., 2007a; Barry et al., 2019a).

With strong parallels in terms of how the individual frames of a movie appear to us to be continuous (Tan, 2018), somehow these separate scenes become linked together, such that we have a sense of a seamless unfolding of life and events (Cutting, 2005; Magliano and Zacks, 2011). This prominence of scenes extends beyond our “online” perception of the world, and resonates throughout cognition. For example, we re-experience these unfolding events as autobiographical memories of past episodes, we simulate future events, and we imagine navigating routes through our environment (Clark et al., 2019).

The central questions I seek to address in this thesis are: **how do we construct these individual scenes, and how do these sets of single scenes become unfolding events?**

Forming mental representations of events is fundamental to everyday life. Our experience is a series of ongoing contiguous events that allow us to construct a coherent autobiography, imbue our lives with a sense of continuity and provide the capacity to recall experiences even when the original events occurred many decades ago. This faculty contributes to our sense of self and is crucial to survival. As well as mentally travelling back to the past, we are able to project ourselves forwards into the future, where we can imagine plausible, self-relevant experiences, anticipating and rehearsing the consequences

of planned behaviours in a highly adaptive manner (Tulving, 1985, 2002; Schacter et al., 2012). This is also true of spatial navigation, whereby we can estimate our position in the environment, find our way home, and plan alternative routes when our circumstances change (O'Keefe and Dostrovsky, 1971; Teng and Squire, 1999; Rosenbaum et al., 2000). Most people can also spontaneously engage in mental imagery when envisioning a scene or event they have not experienced in real life. This natural emergence of scene imagery is clearly evident when we read stories (Speer et al., 2009), and helps us to better grasp the unfolding plots being described. In fact, for young children who have difficulty with reading comprehension, an effective intervention is to encourage greater use of scene imagery (e.g. Murdaugh et al., 2016). And, of course, the ability to visually imagine is also thought to play a key role in the creativity and problem-solving that so enriches and progresses human societies (Moulton and Kosslyn, 2011; Dietrich and Haider, 2015; Zedelius and Schooler, 2016).

However, before proceeding any further, it is necessary to ask whether the use of scene imagery is as prevalent as I have suggested above, particularly during the operation of crucial cognitive functions such as autobiographical memory recall, imagining the future, and spatial navigation. One way to probe this is to examine the explicit strategies people use when engaging in tasks that exercise these functions. There is surprisingly little systematic work in this domain, especially in relation to the use of specifically scene imagery (Andrews-Hanna et al., 2010). To address this issue, I was one of the researchers involved in a large study called *MEMO* (<https://www.fil.ion.ucl.ac.uk/memo/>) where 217 participants performed numerous cognitive tasks, including those assessing scene imagination, autobiographical memory recall, future thinking, and spatial navigation. In a separate session, once all the tasks had been completed, participants were provided with

a large number of strategies tailored to each task, with careful control such that participants were not inadvertently led to select any particular option (Clark et al., 2020). Participants recalled autobiographical memories from early childhood up to the last 12 months (Levine et al., 2002). Future thinking involved imagining and describing plausible future scenes relevant to the participant him/herself (e.g. an event next weekend; Hassabis et al., 2007b). Navigation was tested after viewing movies of navigation through an unfamiliar town, and involved tasks examining recognition memory for scenes, making proximity judgments between landmarks, placing photographs from routes in the correct order, and drawing a sketch map of the town (Woollett and Maguire, 2010). We found that the use of visual scene imagery was by far the most prevalent strategy during all of these tasks. This was true for each sub-task individually, in the case of the autobiographical memory and navigation tasks, as well as when sub-task data were combined (Figure 2).

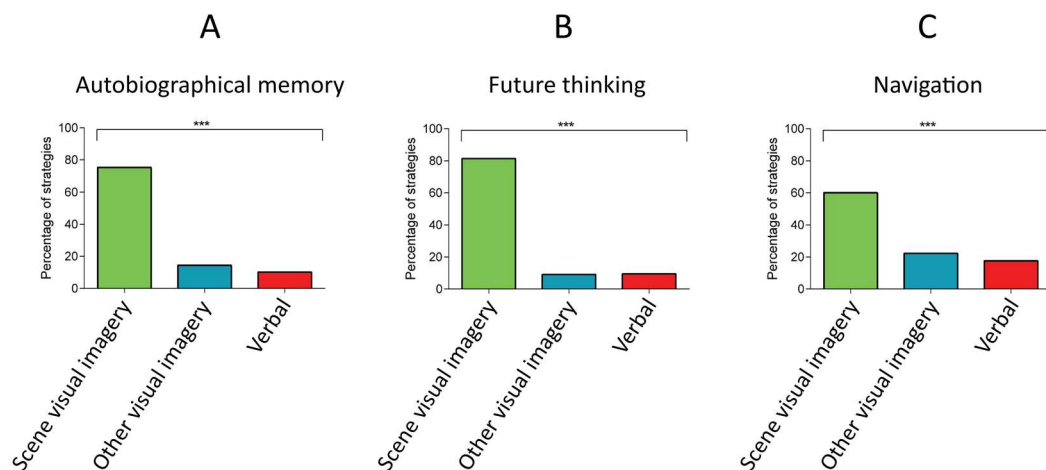


Figure 2. The prevalence of visual scene imagery. Displayed is the percentage count of each top strategy type reported by participants during tasks that assessed (A) autobiographical memory (including memories of multiple ages), (B) future thinking, and (C) navigation (a combination of several navigation sub-tasks). Deployment of scene imagery strategies (green) was compared with non-scene visual imagery (teal) and verbal strategies (red). *** $p < 0.001$. Adapted with permission from Clark et al. (2020).

This thorough interrogation of explicit strategy use in a large sample of participants clearly demonstrates that scene mental imagery dominates when people engage in recalling the past, imagining the future, and spatial navigation. This was also true in the case of learning and recalling pairs of high imagery words. In contrast, scene imagery was employed far less for tasks where visual imagery was less useful, such as the encoding and retrieval of abstract (very low imagery) word pairs, where verbal strategies instead predominated (Clark et al., 2020). Interestingly, when we examined task performance alone, we found that the ability to construct mental images of scenes fully mediated performance on tasks assessing autobiographical memory recall and the ability to imagine the future (Clark et al., 2019).

Together, these data serve to underline that people do indeed think in scenes a great deal of the time, and across numerous cognitive functions. How this might be implemented in the brain is the question I consider next.

In this chapter I will focus on those brain areas that are most germane to my specific research questions, rather than describing the functions of every brain area that has been implicated in scene and event processing. Similarly, even for those regions of particular interest, I will consider extant theories as they relate to the specific focus of this thesis, as opposed to exhaustively reviewing the numerous theoretical accounts of the functioning of a brain region. In addition, while low-level visual scene features – such as spatial frequency, contrast, luminance, and textures – are important fields of study (Berman et al., 2017; Groen et al., 2017; Dima et al., 2018; Epstein and Baker, 2019), they are beyond the scope of my investigations, which instead concentrated on higher level aspects of scene and event processing.

There are many aspects of event processing that could be considered. In particular “event segmentation”, which refers to the chunking of events into meaningful segments, has received increasing attention in recent years. It has been argued that the spontaneous segmentation of continuous experience is supported by hippocampal-cortical processing at time points where discontinuity is detected during an episode, resulting in boundaries being marked. While this aspect of events is not central to my experiments, this literature is comprehensively considered elsewhere (Zacks et al., 2007; Richmond and Zacks, 2017; Brunec et al., 2018b), and will be discussed further in Chapter 7.

1.2 Evidence from functional magnetic resonance imaging (fMRI)

Most functional neuroimaging studies of scene and event processing have involved the use of fMRI, and that is what I consider in this section. I describe the smaller number of directly relevant electroencephalography (EEG) and magnetoencephalography (MEG) studies later in the chapter.

The use of fMRI has consistently revealed a wide set of brain areas associated with autobiographical memory recall (e.g. Maguire, 2001; McDermott et al., 2009; Bonnici et al., 2012), imagining the future (e.g. Hassabis et al., 2007a; Schacter et al., 2012), and spatial navigation (e.g. Bohbot et al., 2004; Wolbers and Büchel, 2005; Spiers and Maguire, 2006; Svoboda et al., 2006; Wolbers et al., 2007; Spreng et al., 2009; Doeller et al., 2010; Rodriguez, 2010; Beaty et al., 2018; Herweg and Kahana, 2018). This includes the ventromedial prefrontal cortex (vmPFC), lateral and polar temporal cortex, posterior parahippocampal cortex (PHC), hippocampus, retrosplenial cortex (RSC), and posterior parietal cortex (see Figure 3). However, it is now also clear that many of these brain

regions are also engaged when people view or imagine single scenes in the absence of an event or spatial navigation (Hassabis et al., 2007a; Epstein, 2008; Zeidman et al., 2015b; Çukur et al., 2016; Dalton et al., 2018; Robin et al., 2018; Robin and Olsen, 2019). This suggests that scene imagery may be fundamental to how the brain processes unfolding events (Zeidman and Maguire, 2016) consistent with the prevalence of visual scene imagery across event cognition that I alluded to earlier (Clark et al., 2019, 2020).

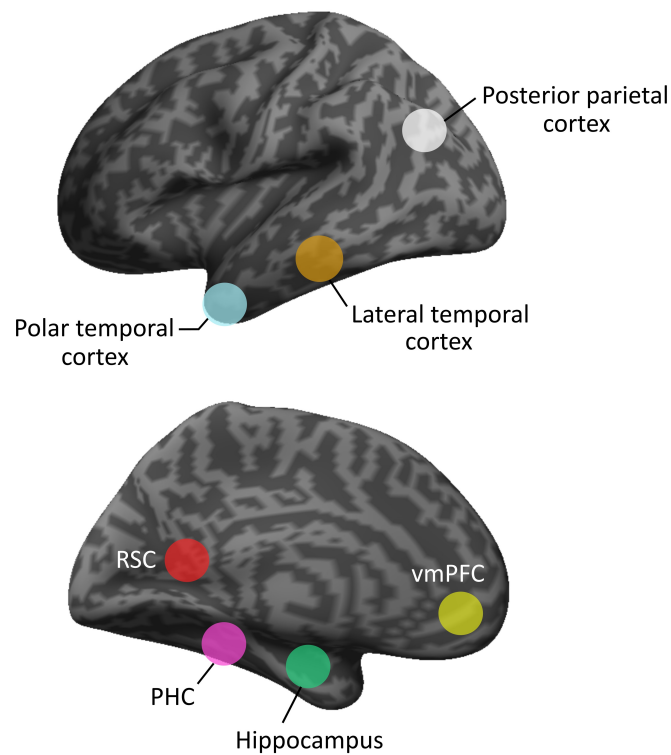


Figure 3. Brain regions commonly engaged during autobiographical memory recall, imagining the future, and spatial navigation. Key brain areas include the ventromedial prefrontal cortex (vmPFC), lateral and polar temporal cortices, posterior parahippocampal cortex (PHC), hippocampus, retrosplenial cortex (RSC), and posterior parietal cortex.

Visual scene stimuli have been deployed extensively in fMRI perceptual, associative, recognition, recall and imagination studies (e.g. Lee et al., 2005; Epstein and Higgins, 2007; Barense et al., 2010; Aly et al., 2013; Harel et al., 2013; van Kesteren et al.,

2013; Horner et al., 2015; Zeidman et al., 2015b; Zeidman and Maguire, 2016; Javadi et al., 2017; Dalton et al., 2018; Patai et al., 2019; Robin and Olsen, 2019b; Chang et al., 2019; Epstein and Baker, 2019; Faivre et al., 2019; Grande et al., 2019; Berens et al., 2020). My consideration of the fMRI literature is guided here by three questions. First, during tasks designed to evoke scene imagery, do the brain regions engaged show a preference for scene representations? Second, if a preferential response to scenes is established, how are scenes built? How, for example, do properties of the objects that populate scenes contribute to the scene construction process? Third, how might single scenes be prepared for inclusion into unfolding mental events?

I note that the PHC and RSC in particular have long been identified as “scene-selective” brain regions which seem to be more engaged by scene stimuli than single isolated objects or faces (Epstein and Kanwisher, 1998; Epstein and Higgins, 2007; Epstein 2008; Park and Chun 2009; Vann et al., 2009; Summerfield et al., 2010; Epstein and Baker, 2019; Berens et al., 2020). However, the specific aspects of scenes to which these regions respond have been debated (Ranganath and Ritchey, 2012). For example, an examination of certain properties of single objects, when scenes were not involved, has also implicated these regions (Mullally and Maguire, 2011; Auger et al., 2012), a point I will return to later.

The hippocampus (Figure 4), situated medially in each temporal lobe, has historically been the subject of intense study in the domains of spatial (O’Keefe and Dostrovsky, 1971; O’Keefe and Nadel, 1978) and autobiographical (Scoville and Milner, 1957; Clark and Maguire, 2016) memory. However, in the last decade in particular, the remit of the hippocampus has been shown to extend beyond space and memory to also include scene processing (Lee et al., 2005; Hassabis and Maguire, 2007; Maguire and

Mullally, 2013; McCormick et al., 2017; Epstein and Baker, 2019). Furthermore, functional differentiation down the long axis of the hippocampus has attracted considerable interest, where the anterior-posterior split is typically defined at the uncus apex (Poppenk et al., 2013), although the exact roles played by each portion in humans remains unclear (Poppenk et al., 2013; Strange et al., 2014; Ritchey et al., 2015; Zeidman and Maguire, 2016; Brunec et al., 2018a).

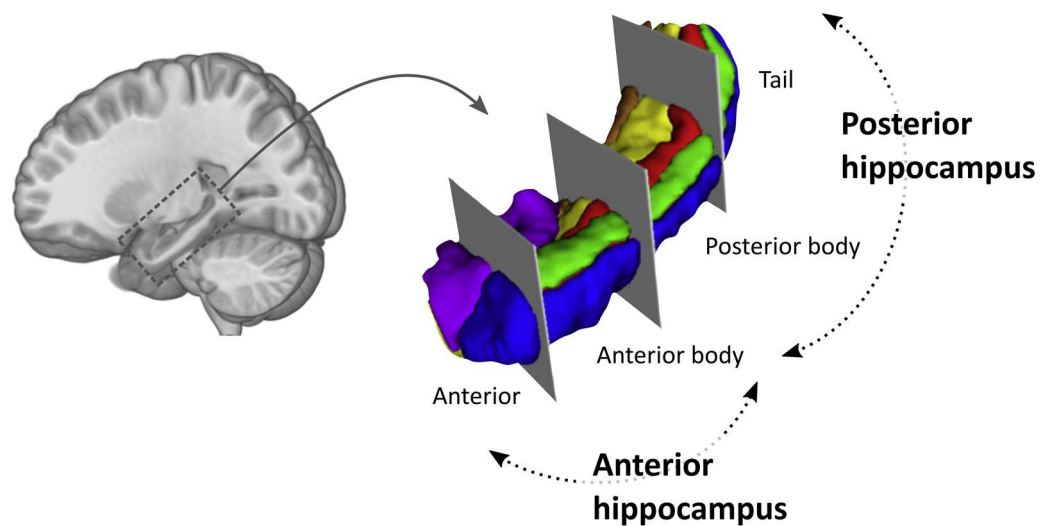


Figure 4. The hippocampus. A 3D rendering of the hippocampus shows the location of the anterior hippocampus and the posterior hippocampus. The uncus (purple) can be considered the border between these sub-regions. The other colours represent different hippocampal sub-fields. Adapted with permission from Dalton et al. (2019).

In relation to scenes, fMRI studies have demonstrated an anterior-posterior hippocampal distinction between imagination and visual perception (Zeidman et al., 2015a, 2015b; Hodgetts et al., 2016, 2017). In particular, Zeidman et al. (2015b) found that the anterior medial portion of the hippocampus was engaged in common during both the perception of scenes – an “online” scene construction process whilst viewing an externally presented scene – and the active imagination of scenes – an “offline” process

where representations of scenes were internally generated (Zeidman et al., 2015b). Connectivity of the anterior hippocampus was strongest and more widespread during imagination. The posterior hippocampus was engaged specifically during visual scene perception. In this thesis, I focus on the anterior hippocampus given its involvement in constructing mental scene representations (Hassabis et al., 2007b; Graham et al., 2010; Zeidman et al., 2015a, 2015b; Hodgetts et al., 2016; Zeidman and Maguire, 2016).

1.2.1 Preferential brain responses when building single scenes

While fMRI studies have pinpointed the anterior hippocampus as being engaged during scene perception and imagination, including when compared to simpler stimuli such as single objects (Henderson et al., 2008; Barense et al., 2010; MacEvoy and Epstein, 2011; Zeidman et al., 2015b; Hodgetts et al., 2016, 2017; Brandman and Peelen, 2017; Dalton et al., 2018; Chang et al., 2019; Pitcher et al., 2019; Josephs and Konkle, 2020), it remains unclear whether scenes, rather than other types of multi-feature stimuli, preferentially engage this region. One account, the scene construction theory, posits that scenes are preferentially processed by the hippocampus (Hassabis and Maguire, 2007; Maguire and Mullally, 2013). By contrast, another perspective, the relational theory, argues that it is the associating of multiple elements that requires hippocampal input, irrespective of whether scenes are the outcome of such processing (Cohen and Eichenbaum, 1993; Lee et al., 2005; Eichenbaum, 2006; Aly et al., 2013; Eichenbaum and Cohen, 2014; Erez et al., 2016; also see Yonelinas et al., 2019 for a related view).

Several fMRI studies have explicitly addressed this issue. McCormick et al. (2021) compared the viewing of naturalistic scenes with scrambled versions of the same scenes. The scrambled images possessed the same spatial frequency and colour scheme as the

original scenes, but their phase was randomised such that any meaning was removed from the image. With task requirements and image complexity matched across both image types, they found that the anterior hippocampus was more engaged by scenes compared to the scrambled images. Other brain areas showed a similar scene preference, including the vmPFC, lateral temporal cortex, PHC and parietal regions.

Summerfield et al. (2010) adopted a different approach to addressing the issue of preferential responses to scene processing. They “slowed-down” the scene construction process through the use of auditorily presented phrases describing single objects in a serial manner. Participants were required to integrate these objects together in their imagination to form a scene. This was compared to another condition where participants heard a series of short “jargon” phrases, one after the other. These phrases comprised word combinations designed to elicit minimal imagery and mnemonic associations. This condition controlled for basic auditory stimulation, language, attention, working memory, difficulty, and effort. When the two tasks were compared, the anterior hippocampus, vmPFC and other areas including lateral temporal cortex, PHC, RSC and parietal regions, were more engaged by the scene construction process.

While the McCormick et al. (2021) and Summerfield et al. (2010) studies provide some support for the hippocampus and other areas showing a scene preference, in order to compare the scene construction and relational theories more precisely, an even closer matching of tasks is required. Dalton et al. (2018) devised just such an fMRI paradigm. Participants gradually built scene imagery from three successive auditorily-presented object descriptions and an imagined 3D space. This was contrasted with constructing mental images of non-scene arrays that were composed of three objects and an imagined 2D space. The scene and array stimuli were, therefore, highly matched in terms of content

and the associative (relational) and constructive processes they evoked, with the resultant representations differing only in terms of whether or not they evoked a scene. Moreover, the objects in each triplet were not contextually related. Two other conditions involved imagination of either a 2D or 3D space alone (devoid of objects).

Relational theory would predict similar processing of the scenes and arrays, since both stimuli are representations of the same number of elements (objects) bound together within space. However, Dalton et al. (2018) found that the anterior medial hippocampus was engaged preferentially during the construction of scenes relative to arrays, consistent with previous findings (Zeidman et al., 2015b; Hodgetts et al., 2016; Zeidman and Maguire, 2016) and the scene construction theory (Hassabis and Maguire, 2007; Maguire and Mullally, 2013). Of note, when the imagined 3D and 2D spaces alone were examined, neither was associated with increased hippocampal activity (see Zeidman et al., 2012 for a similar result). It was also evident that different regions within the hippocampal region were engaged by the construction of arrays – the posterior hippocampus and the entorhinal cortex. It has been suggested that the posterior hippocampus may govern the construction of more abstracted visuospatial imagery (Constantinescu et al., 2016), and may explain why patients with hippocampal damage retain the capacity for associative processing on some tasks (e.g. Mayes et al., 2004; Kumaran et al., 2007; and for a review see Clark and Maguire, 2016). These distinct portions of the hippocampal region were implicated despite the close matching of stimuli, suggesting that scene and array representations tap into separable mental construction processes, served by different circuitry within the hippocampus. This speaks against the view that scenes are merely exemplars of associative processing. The anterior hippocampus, in particular, appears to

be especially attuned to the construction of scenes, rather than to multi-component representations in general.

The role of the hippocampus in supporting verbal paired associates provides further potential evidence for preferential scene processing in the anterior hippocampus. Typically, patients with bilateral hippocampal damage are impaired at recalling previously presented word pairs. This could be because the hippocampus is responsible for binding information together, as argued by the relational theory. However, it has been noted (e.g. Clark and Maguire, 2016; Clark et al., 2018) that standard verbal paired associates tests typically use highly imageable concrete words, often a mixture of object words (e.g. cat) and scene words (e.g. forest). Consequently the use of imagery and item binding are conflated.

Clark et al. (2018) sought to separate these task features using fMRI where healthy participants had to encode word pairs. The pairs could contain highly imageable scene words, highly imageable object words, or very low imagery abstract words (e.g. opinion). An increase in anterior hippocampal activity was observed during the encoding of highly imageable word pairs. Moreover, the anterior hippocampus was more engaged for pairs of scene words compared to pairs of object words, illustrating the importance of not only visual imagery, but of scenes specifically, in hippocampal processing. Abstract word pairs did not provoke any increase in hippocampal activity, even when subsequent memory for the word pairs was matched to the scene and object word pairs. This suggests that associative binding alone, in the absence of high imagery, may not preferentially recruit the hippocampus. Overall, it seems that the relational theory cannot easily account for these findings and those of Dalton et al.'s (2018) scenes and non-scene arrays, and instead the findings accord better with the scene construction theory.

Another prominent hippocampal theory has relevance here – the cognitive map theory (O’Keefe and Nadel, 1978) and its later developments (e.g. Byrne et al., 2007). According to this account, the hippocampus specifically supports flexible, allocentric (independent of the observer’s perspective) representations of spatial relationships. Within this perspective, scenes are regarded as spatial contexts, and it is this spatial component that is held to drive hippocampal responses to scenes. However our experience of scenes is fundamentally egocentric (person-centred). Moreover, as noted previously in the Dalton et al. (2018) study, neither 3D nor 2D space alone was associated with increased hippocampal activity, echoing a previous finding where the depiction of a 3D space alone (again, without objects) did not recruit the hippocampus (Zeidman et al., 2012). Space *per se* therefore seems insufficient to engage the hippocampus. I will consider the cognitive map theory and space further in Section 1.4. However, from the pertinent fMRI literature, it seems that representations that combine objects specifically in the context of a 3D space, i.e. scenes, consistently engage the anterior hippocampus, rather than associative or spatial processing *per se*.

Having noted the brain areas that seem to show preferential responses during scene construction, an issue I will return to in Chapter 3, I next consider whether particular properties of objects might influence the scene construction process.

1.2.2 Properties of objects in scenes

Although there is a large body of literature on higher-order object properties that appear to modulate responses in scene-selective cortex (for a review, see Epstein and Baker, 2019), one object attribute discovered by Mullally and Maguire (2011) holds particular relevance for constructing scenes. Participants reported that certain objects, when viewed

or imagined in isolation (and therefore in the absence of a scene), evoked a strong sense of 3D space surrounding them – objects the authors described as space-defining (SD). In other words, these objects appeared to anchor themselves in the surrounding space. Conversely, space-ambiguous (SA) objects do not impose themselves on the surrounding space in the same way (for examples, see Figure 5). Instead, many participants spontaneously described SA objects as “floating” in space, emphasising the detachment of these objects from an explicit sense of surrounding space (Mullally and Maguire, 2011).

As I have already noted, the PHC is typically regarded as a “scene-selective” brain region. However, Mullally and Maguire (2011) found that PHC was engaged to a greater extent when imagining or viewing isolated SD objects, relative to SA objects, despite the absence of a scene context. Object size and contextual associations did not account for the PHC-mediated SD effect. This preferential response of the PHC to single SD objects aligns with findings from several subsequent studies (Kravitz et al., 2011; Harel et al., 2013). The PHC has also been shown to be particularly sensitive to the spatial frequency of scenes (Kauffmann et al., 2015). In addition, it is even engaged by the simple depiction of a 3D space using simple high spatial frequency dot field stimuli, devoid of both a scene layout and objects (Zeidman et al., 2012). Altogether these findings suggest that the PHC may not be responsive to scenes *per se*, but perhaps to the presence of 3D space, which is of course contained in scenes.

Mullally and Maguire (2011) also found that SD objects tended to be less portable, maintaining a permanent location. Notably, however, a considerable number of SA objects also had reasonably high permanence, suggesting that, while certainly related to SD, permanence is not the sole basis of the SD effect. Indeed, object permanence has

been linked to the RSC rather than the PHC (Auger et al., 2012, 2015, 2017; Auger and Maguire, 2013, 2018; Troiani et al., 2014).

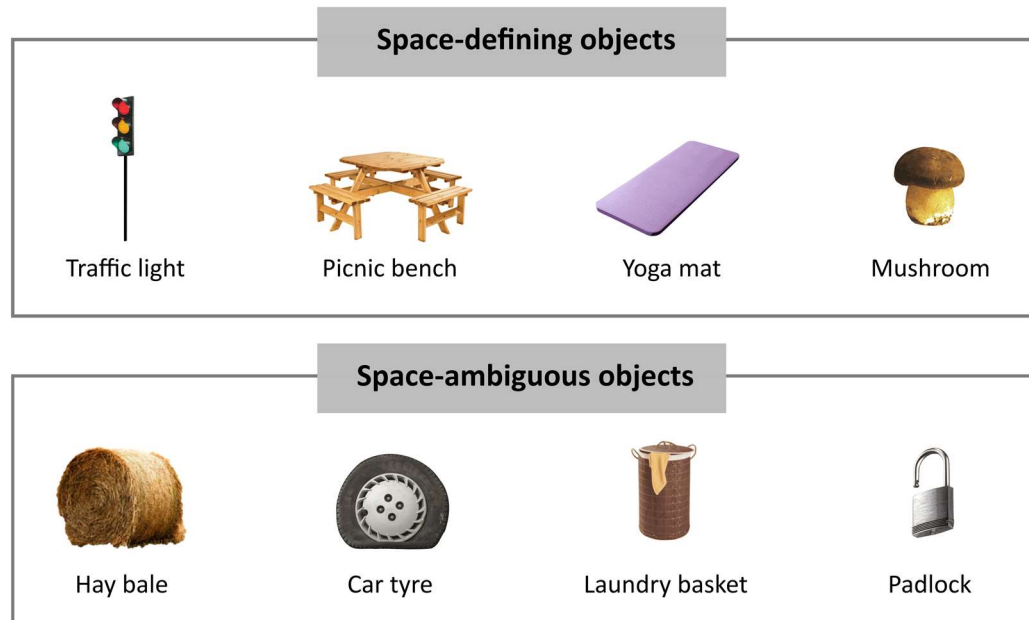
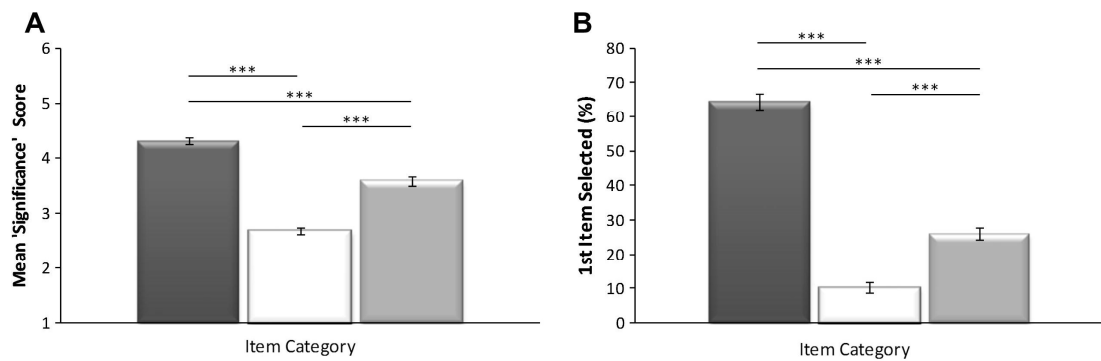


Figure 5. Examples of space-defining (SD) and space-ambiguous (SA) objects. SD objects evoke a strong sense of 3D space surrounding them when viewed or imagined in isolation, while SA objects do not. The examples illustrate how objects in either category can be large (e.g. traffic light and hay bale), heavy (e.g. picnic bench and car tyre), portable (e.g. yoga mat and laundry basket), or small (e.g. mushroom and padlock).

In a subsequent behavioural study, Mullally and Maguire (2013) found that SD objects may enjoy a privileged role in building scenes. Participants were instructed to gradually construct and deconstruct scenes in their imagination element-by-element using SD, SA, and background items (such as floors, ceilings). The order of object choices in both of these tasks indicated the significance of each type of item in the formation of scenes. During scene construction, participants showed a strong preference for SD objects, even when comparatively larger and more permanent SA objects were available.

When asked to deconstruct scenes, participants retained significantly more SD than SA objects, with the last object remaining most likely to be an SD object (Figure 6).

Scene Construction



Scene Deconstruction

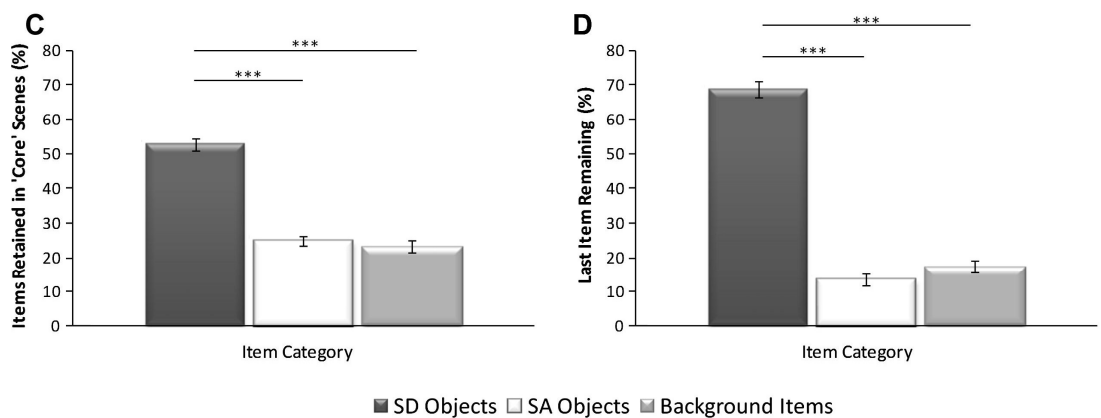


Figure 6. Preference for SD objects when constructing and deconstructing scenes. (A) The higher mean “significance” score attained by SD objects reflects this category of item was selected earlier in the scene construction process. The mean significance score of SD objects was significantly greater than that of both SA objects and background items. (B) SD objects were significantly more likely to be selected as the first item when constructing scenes than both SA objects and background items. (C) A significantly larger percentage of SD objects were retained when deconstructing scenes, compared to SA objects and background items. The “core” scene was the scene that remained after removing items the participant considered dispensable. (D) Participants were then instructed to remove the core items one at a time until only one was left. Displayed is the percentage of SD, SA, and background items retained as the final item. A clear preference was shown for SD objects over both SA objects and background items. *** $p < 0.001$. Reproduced with permission from Mullally and Maguire (2013).

These findings seem to place SD objects at the heart of constructing scenes. Mullally and Maguire (2011) assessed objects in isolation during fMRI. Understanding how SD objects influence the brain areas engaged while scenes are being built remains unknown – an issue I return to in Chapter 4.

1.2.3 Preparing scenes for inclusion in events

Until now I have focussed on single scenes. However, our daily experience is not characterised by a series of self-contained scenes. Despite the ambiguity of our visual input – caused by physical occlusions and eye blinks – our impression is that of a stable, continuous stream of scene-like imagery. What might facilitate the transition from one scene to the next, taking us a step closer to an unfolding event?

An important feature of perception, and arguably crucial to survival, is the ability to predict what lies beyond the immediate view (Gregory, 1980; Friston, 2010). This has clear relevance across cognition, particularly in relation to future thinking and decision-making. Without this faculty, a coherent integration of different views of our surroundings would be lacking, the ability to plan our actions would be compromised, and so our behaviour would be limited to the influence of what is immediately in front of us.

Our ability to imagine what might be just beyond the view in a scene is exemplified by a cognitive phenomenon known as boundary extension (BE). This describes an implicit, anticipatory form of scene construction where, upon viewing a scene, we construct a mental representation of the scene that extends beyond its observed borders (Intraub and Richardson, 1989). The result of this extrapolation is that when the

scene is no longer in view, people incorrectly recall more of the scene than was originally displayed. Importantly, this anticipatory error only occurs in response to scenes, and not to single isolated objects (Gottesman and Intraub, 2002). BE has been widely replicated in a range of populations, using many different scene stimuli (Intraub and Richardson, 1989; Intraub et al., 1992, 1996; Seamon et al., 2002; Chadwick et al., 2013; Czigler et al., 2013; Spanò et al., 2017). Interestingly, it is not limited to the visual modality, with BE also being observed when tested haptically (Intraub et al., 2015).

BE is thought to comprise two phases (Figure 7). In the first phase, when we view a novel scene for the first time, there is an automatic extrapolation beyond the boundaries of the scene, whereby an extended mental representation of the scene is constructed. This phase is the initial “BE effect” (Figure 7, left panel). This implicitly-constructed representation is maintained, even when the scene is no longer present. After a delay, a second phase occurs at retrieval, where the “BE error” is made (Figure 7, right panel). When presented with the exact same picture, and asked whether it is closer-up, the same (the correct answer) or farther away, people often erroneously say the second picture is closer-up than the first. This is because people tend to perceive the second picture as possessing less background when compared to their extended internal representation of the scene. The BE effect is greatest for scene views that are close-cropped with less background, since this provokes the need and opportunity for greater extension (Intraub et al., 1992; see also Intraub, 2020 for a further explanation of BE in response to Bainbridge and Baker, 2020). BE may help to smooth the transitions between the individual scenes that we perceive, perhaps contributing to our sense of a continuous unfolding of life and events (Hubbard et al., 2010).

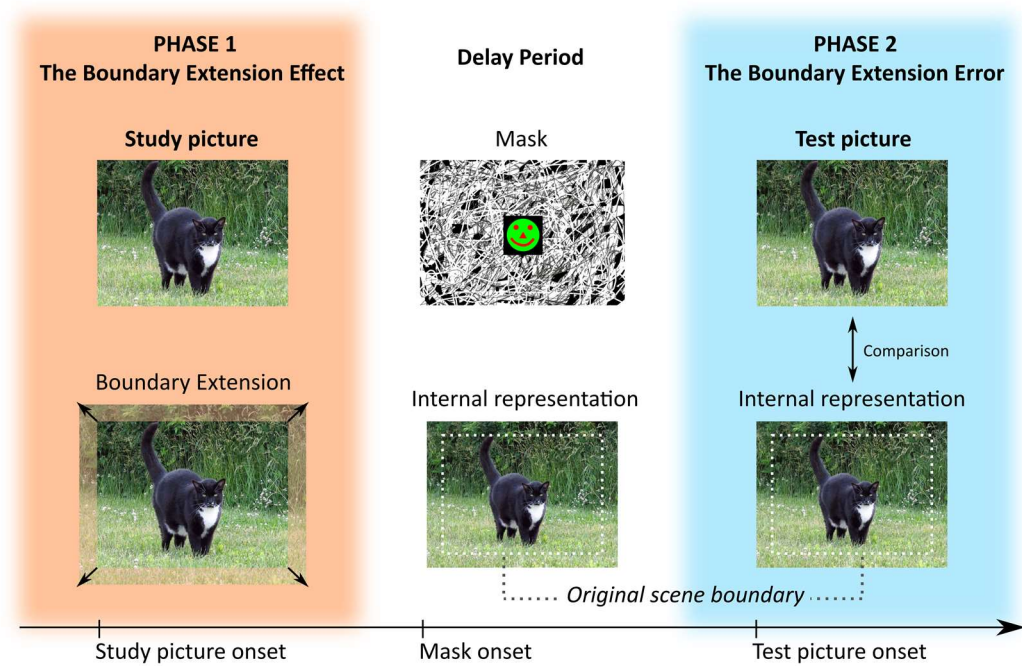


Figure 7. Illustration of the boundary extension (BE) phenomenon when identical scenes are presented at study and test. When we view a scene (**phase 1, left panel**), we automatically extrapolate beyond the original view, called the “BE effect”, forming an extended mental representation of the scene, retained in memory even after the scene is no longer present (**delay period, centre panel**). The white dashed rectangle indicates the boundaries of original stimulus. When the exact same scene picture is presented at test, the extended mental representation is compared to the test picture, leading to the perception that this second scene is “closer-up” than the first, even though they are identical (**phase 2, right panel**). This is the “BE error”.

To the best of my knowledge, there are only two fMRI studies that have investigated BE. The studies differ in a fundamental way. Park et al. (2007) examined the neural response to the second (test) scene, and found that the PHC and RSC were sensitive to the BE error (Park et al., 2007). By contrast, Chadwick et al. (2013) focussed on the initial point of scene extrapolation (the BE effect) during presentation of the study picture. They used a classic BE paradigm known as the rapid serial visual presentation (RSVP) task (Intraub et al., 1996; Intraub and Dickinson, 2008; Mullally et al., 2012b), where a photograph of a scene is presented for a very brief period (250 ms), followed by a brief interval where a visual noise mask is shown (see Figure 7). The same scene picture

is shown a second time, at which point the participant judges whether the test image is the same as the first (the correct answer), closer-up, or farther away (incorrect answers). The BE effect was associated with an increase in activity in the PHC and the hippocampus, indicating these regions supported the active extrapolation of scenes. Moreover, the hippocampus was found to drive activity in early visual cortices – regions traditionally thought of as upstream during visual perception. These findings suggest a central role for the hippocampus in the automatic construction of unseen scenes (Chadwick et al., 2013). I will consider BE further in Section 1.3, and return to it again in Chapter 5.

fMRI studies have been invaluable in revealing the wide set of brain regions that are involved in scene processing. The anterior hippocampus emerges as an area of interest because it seems to be particularly attuned to scene representations. SD objects appear to be influential in constructing scenes, while BE may facilitate the transitions between the multiple scenes that comprise unfolding events. Complementary insights can be gleaned from neuropsychological studies of scene and event processing, which I consider next.

1.3 Evidence from neuropsychology

Of all the scene-responsive brain areas revealed by fMRI, two in particular wreak havoc when damaged (McCormick et al., 2018a; Aly, 2020). The first is the hippocampus, which I have already noted is often divided into its anterior and posterior portions (Figure 4). This is, however, a complex region that can be further subdivided. As will become clear, my focus is not on these specific hippocampal subfields, but I acknowledge the growing body of evidence showing that each subfield has a unique pattern of connectivity, and

may therefore play a distinct role in cognition (Wisse et al., 2012; Zeidman et al., 2015a; Dalton et al., 2017, 2019; Olsen et al., 2019).

The second area I will focus on is the vmPFC. While this region is consistently activated during fMRI studies of event processing (Maguire, 2001; Svoboda et al., 2006; McDermott et al., 2009; Spreng et al., 2009; Roy et al., 2012; D'Argembeau, 2013; Benoit et al., 2014; Bonnici and Maguire, 2018; Lieberman et al., 2019), its engagement during scene processing is often overshadowed by the PHC, RSC, and latterly the hippocampus. Yet the vmPFC has been implicated in tasks involving scenes, including during imagination (Hassabis et al., 2007a; Bertossi et al., 2016a; McCormick et al., 2018a; Barry et al., 2019a, 2019b; McCormick et al., 2020). The vmPFC is located medially in the prefrontal cortex and, just like the hippocampus, is not a homogeneous structure. It comprises subdivisions that are cytoarchitecturally distinct, specifically Brodmann areas 10, 14, 25, as well as parts of 32, 11, 12 and 13 (Catani et al., 2012; McCormick et al., 2018a; also see Figure 8), which likely play distinct functional roles (Bush et al., 2000; Northoff and Bermpohl, 2004; Amodio and Frith, 2006; Schoenbaum et al., 2011; see Roy et al., 2012 for a review).

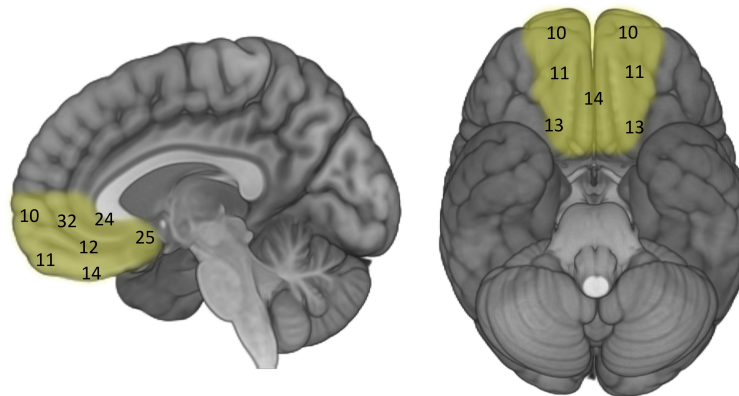


Figure 8. The ventromedial prefrontal cortex. Illustration of the parcellation of the ventromedial (**left panel**) and orbital (**right panel**) surfaces into Brodmann areas. The Brodmann areas 10, 11, 12, 13, 14, 25, 32 together form the region of interest used in my experiments.

Of particular interest here, the vmPFC is known to have strong anatomical links with the anterior hippocampus (Catani et al., 2012, 2013; Adnan et al., 2016). They are connected via three main reciprocal pathways – the fornix, the cingulum bundle, and the uncinate fasciculus (Concha et al., 2005; Catani et al., 2013; see Figure 9). The fornix connects the hippocampi in the two hemispheres, and projects from the hippocampus and entorhinal cortex to the vmPFC. The cingulum bundle is an extensive network of white matter tracts and includes connections between the anterior hippocampus, PHC, and vmPFC. Finally, the uncinate fasciculus connects anterior temporal lobe structures with the frontal cortex, including the hippocampus and vmPFC. A fourth pathway, via mammillothalamic and anterior thalamic projections, provides indirect connections between the hippocampus and vmPFC.

Many neurological and psychiatric pathologies impinge upon the hippocampus and vmPFC, including dementia and schizophrenia. I will focus on the deficits associated with non-progressive focal bilateral damage to these areas, rather than diffuse conditions.

These more selective lesions facilitate clearer inferences about the functions of these two regions. Naturally, no brain region operates in isolation, and I acknowledge that functions like scene and event processing are supported by many brain regions, meaning that damage to one brain area will likely influence the function of others to which it is connected (see Hayes et al., 2012; Rudebeck et al., 2013; Henson et al., 2016; Argyropoulos et al., 2019; Miller et al., 2020).

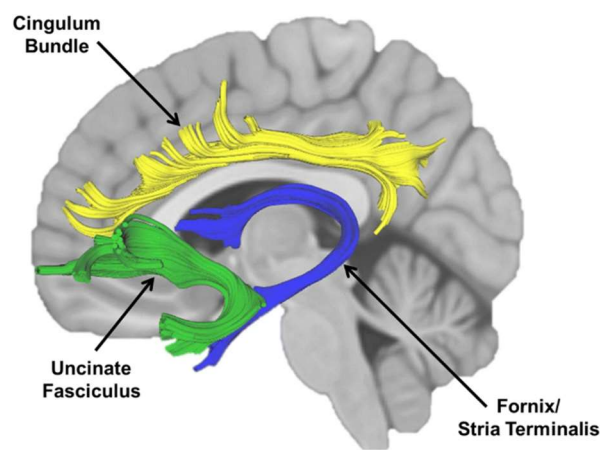


Figure 9. The main white matter tracks connecting the hippocampus and vmPFC. There are three main pathways: the fornix (blue), the cingulum bundle (yellow), and the uncinate fasciculus (green). Reproduced with permission from Harnett et al. (2018).

1.3.1 Effects of hippocampal damage

Bilateral hippocampal damage (see Figure 10) has classically been linked to a deficit in forming new autobiographical memories of past experiences (Scoville and Milner, 1957; Squire, 1992), as well as the ability to vividly retrieve them from memory (Viskontas et al., 2000; Addis et al., 2007; Rosenbaum et al., 2008, 2009; St-Laurent et al., 2011). The most famous patient with hippocampal amnesia to illustrate these deficits was “HM” in the 1950’s (Scoville and Milner, 1957; Corkin et al., 1997). As noted previously, patients with hippocampal damage are also impaired at recalling word pairs. The tests used

typically involve imageable words, and Clark et al. (2020) have shown that this task involves the use of scene imagery and activates the hippocampus (Clark et al., 2018). Therefore, the patients may be impaired on word pair associate tasks because a difficulty in visualising scenes.

This idea resonates beyond the inability to recall memories, because hippocampal-damaged patients are also impaired at imagining plausible future scenarios (Klein and Loftus, 2002; Hassabis et al., 2007b; Addis and Schacter, 2012; Kurczek et al., 2015; but also see Squire et al., 2010). Interestingly, these patients also struggle to imagine single fictitious atemporal scenes (not rooted in the past or future). This suggests their problem is not related to mental “time-travel” *per se*, but rather may be centred on a problem constructing scene imagery that underlies these cognitive abilities (Hassabis et al., 2007b). They are, however, still able to imagine single isolated objects (Hassabis et al., 2007b), so the deficit pertains specifically to scenes and not imagination more generally.

This apparent scene-related deficit in the context of bilateral hippocampal damage is further evidenced by difficulties in discriminating between different scene views, even when all images are visible (Lee et al., 2005, 2013; Barense et al., 2010; also see Kim et al., 2011 for a different perspective). This not only suggests an impairment related to scene perception, but may also reflect the need for intact hippocampi in order to form internal models of scenes, in order for scenes to be mentally rotated and compared to solve the task (Zeidman and Maguire, 2016). Importantly, patients can detect differences between single objects or faces. Furthermore, this difficulty relating to scene representations seems to concern specifically their construction – hippocampal-damaged patients can describe, from their semantic knowledge, elements that would be within a particular scene (Hassabis et al., 2007b; Maguire and Mullally, 2013; McCormick et al., 2018a), and

perform comparably to healthy controls when detecting semantic violations in scenes, but fail to detect constructive violations (McCormick et al., 2017).

Healthy control participant



Patient with bilateral hippocampal damage



Figure 10. Example of bilateral hippocampal damage. T2-weighted coronal MRI images of the hippocampi of a healthy control (**top**) and a patient with selective bilateral hippocampal damage (**bottom**) is shown, with the hippocampal region highlighted in green. Reproduced with permission from McCormick et al. (2016).

Another aspect of cognition that hippocampal damage appears to affect is mind-wandering. Patients with hippocampal lesions report the same frequency of mind-wandering as control participants, but what they think about in their idle moments differs. McCormick et al. (2018c) showed that these patients thought more in semantic terms rather than the more episodic thinking of control participants. Moreover, the form of mind-wandering also diverged in a fundamental way – controls mind-wandered in terms of vivid scenes, resonating with reports of the prevalence of scene imagery observed by Clark et al. (2020). By contrast, patients were much less likely to use or experience scene imagery while day-dreaming. Interestingly, Spanò et al. (2020) recently reported that bilateral hippocampal damage resulted in vastly reduced frequency of sleep dreaming, and

the small number of dreams the patients had during the night were impoverished in terms of episodic content. This suggests the hippocampus may be needed for normal dreaming, which is typically episodic in nature.

Bilateral hippocampal damage has also been shown to adversely affect visual perception in a range of tasks (e.g. Lee et al., 2005, 2012, 2013; Graham et al., 2010; Warren et al., 2012; Aly et al., 2013; Yonelinas, 2013; Epstein and Baker, 2019). Of most relevance to this thesis, Mullally et al. (2012) found that patients had attenuated BE relative to controls in the RSVP task (Figure 11) – described earlier in Section 1.2.3. They correctly identified significantly more study-target scene pairs as “the same”, while controls responded incorrectly with “closer-up” more often. Therefore, despite being amnesic, paradoxically the patients performed better on this memory test because they could not imagine what was beyond the view in scenes.

To summarise, focal bilateral hippocampal lesions have been linked to a constellation of scene-related perceptual, imagination and mnemonic deficits. Considering this evidence, it seems reasonable to conclude that the impairments of hippocampal-damaged patients extend beyond memory. Instead, these patients seem to be limited to what they can see in front of their eyes and have a fundamental deficit in constructing scene imagery. This prevents them from imagining single scenes, potentially preparing scenes for inclusion into unfolding mental events, and then building events from sets of scene images during recall or future thinking. On face value, lesions to another brain region, the vmPFC, seem to result in similar deficits, but as I outline next, there are, in fact, distinct differences between the resultant cognitive profiles.

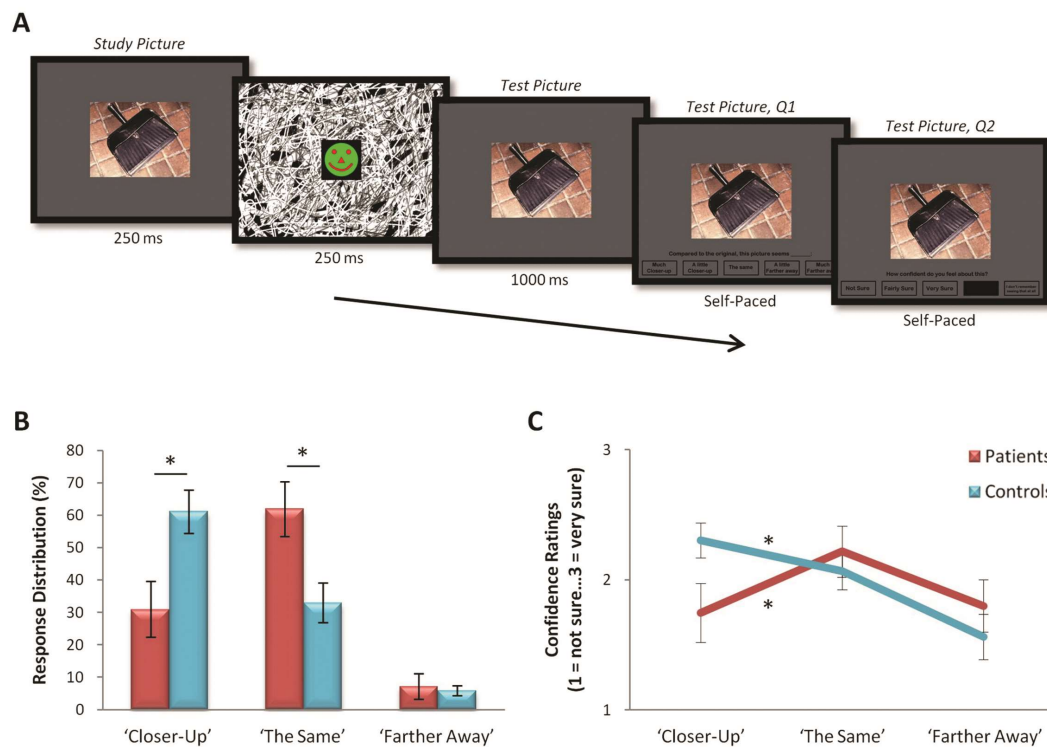


Figure 11. Attenuated BE in patients with hippocampal damage compared to healthy controls. (A) The structure of trials for the RSVP task. (B) The proportion of trials which patients (red) and controls (blue) rated as “closer-up”, “the same” (correct), or “farther away”. This is expressed as a percentage of responses in each category relative to the total number of responses. Controls made significantly more BE errors (closer-up) than patients, who conversely made significantly more correct responses (“the same”). (C) Mean confidence ratings, on a scale of 1 “not sure”, 2 “fairly sure”, and 3 “very sure”. * $p < 0.05$. Error bars are ± 1 SEM. Reproduced with permission from Mullally et al. (2012).

1.3.2 Effects of vmPFC damage

Patients with bilateral damage to the vmPFC (see Figure 12) display deficits in autobiographical memory and imagining the future that appear similar to hippocampal-lesioned patients (reviewed in McCormick et al., 2018a). Compared to healthy controls, vmPFC-damaged patients generate fewer autobiographical episodes (Kopelman et al., 1999), and provide impoverished accounts of both past memories (Bertossi et al., 2016a, 2016b, 2017) and simulated future experiences (Bertossi et al., 2016a) when tested with relatively unconstrained cues. However, unlike hippocampal-damaged patients, those

with vmPFC lesions are able to describe single snapshot scenes in as much detail as controls, whether drawn from real past events or fictitious future events, as long as they are heavily cued (Kurczek et al., 2015). This suggests their basic ability to generate scene imagery is intact, as long as they are directed to exactly what they are required to (re-) imagine. What vmPFC-damaged patients seem unable to do, however, is to go beyond a single scene and describe an event unfolding over a longer time scale (Bertossi et al., 2016a). Moreover, they are impaired at situating autobiographical memories on a timeline of their lives (Tranel and Jones, 2006). This highlights a potential key difference between the involvement of the hippocampus and vmPFC in generating scene imagery – while the former seems to be involved in the construction of individual scenes, the latter may initiate scene construction, as well as support the sequencing of scenes that comprise an unfolding mental event.

Another significant feature of autobiographical memory recall in vmPFC-damaged patients is confabulation. This refers to the unintentional production of false autobiographical memories that a patient believes to be true – often termed “honest lying” (Moscovitch, 1989; Moscovitch and Melo, 1997). This may indicate a problem with selecting appropriate memory representations and inhibiting those that are not relevant (Moscovitch and Melo, 1997; Gilboa et al., 2006; Gilboa and Marlatte, 2017). In this regard, schemas emerge as a relevant constructs – adaptive knowledge structures that comprise information that is acquired over multiple episodes (Ghosh and Gilboa, 2013). In essence, schemas provide prior knowledge with which to evaluate incoming information, and may explain why patients with vmPFC damage are impaired when deciding between schema-relevant and schema-irrelevant information (Ghosh et al., 2014). It may also be the case that schemas provide the necessary structure around which

extended mental events are (re-)built. As vmPFC patients are unable to access such information, this might explain why they are impaired at extending beyond single scenes and situating memories correctly in time.

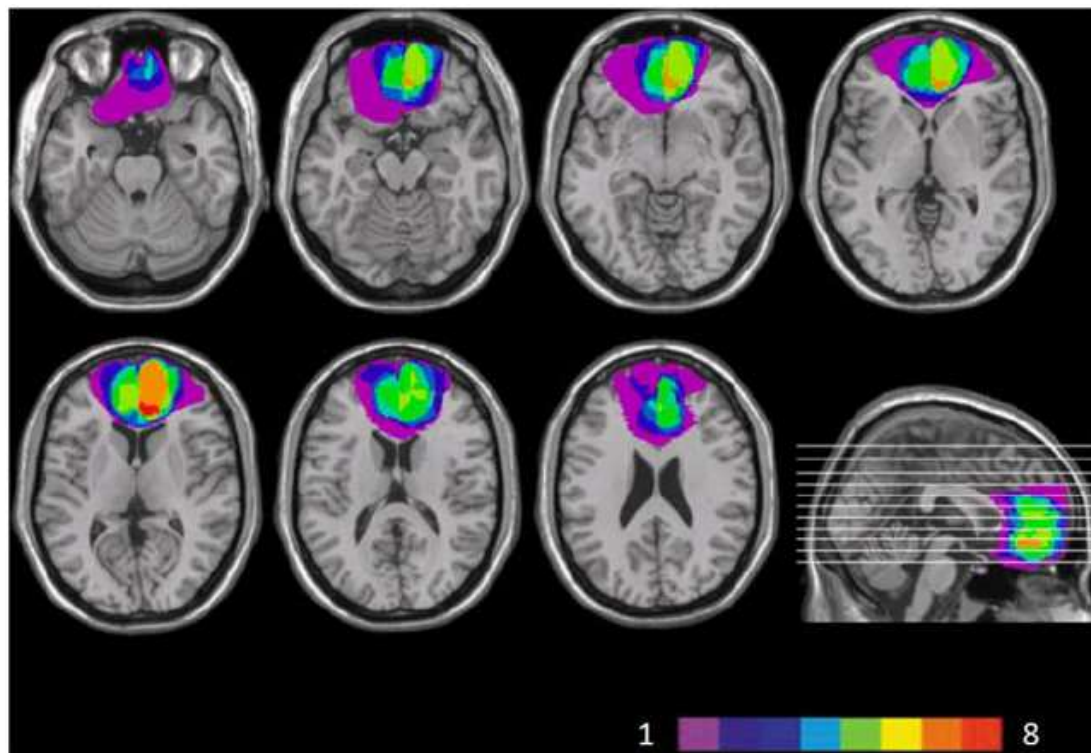


Figure 12. Example lesions of the vmPFC. Location of brain lesions within the vmPFC in eight patients. The colour bar shows the number of overlapping lesions across patients, which encompass portions of Brodman areas 10, 11, 12, 13, 14, 25, 32. Reproduced with permission from Ciaramelli et al. (2012).

vmPFC-damaged patients have other cognitive deficits that seem to coalesce around a reduced ability to initiate endogenous processing and introspection. For example, their mind-wandering – a form of spontaneous mental activity – is much reduced in frequency (Bertossi and Ciaramelli, 2016). These patients tend to make remarks instigated by features in the immediate external environment, rather than resulting from internally-generated thoughts and emotions (Bechara, 2004; Leopold et al., 2012; Bertossi and Ciaramelli, 2016; McCormick et al., 2018a).

The vmPFC has also been implicated in BE, as revealed by a recent study that compared the performance of patients with vmPFC damage, healthy controls, and brain-damaged control patients (De Luca et al., 2018). Patients with vmPFC damage performed like the previously-reported hippocampal-lesioned patients (Mullally et al., 2012b), displaying attenuated BE. Of note, the brain-damaged control patients performed like healthy controls, displaying the classic BE effect where they extended their mental representation of scenes beyond the view (Figure 13). This shows that disruption to normal BE is not the result of brain damage *per se*.

These BE results suggest that the vmPFC and hippocampus are equally involved in rapid, implicit scene construction. In the same two studies, clues about the possible individual contributions of the hippocampus (Mullally et al., 2012b) and vmPFC (De Luca et al., 2018) in scene imagery emerged when patients were further probed with an explicit scene construction task. They were shown a photograph of a scene and asked to imagine what lies beyond the view. The hippocampal-damaged patients reported being unable to imagine what was beyond the view, but could provide appropriate semantic content for what was likely to be there. The patients with vmPFC lesions, on the other hand, struggled to produce scene content and were poor at predicting what objects might be typically found in the scene (De Luca et al., 2018). This suggests a more general role for the vmPFC in constructing scenes, such as the selection of appropriate scene elements.

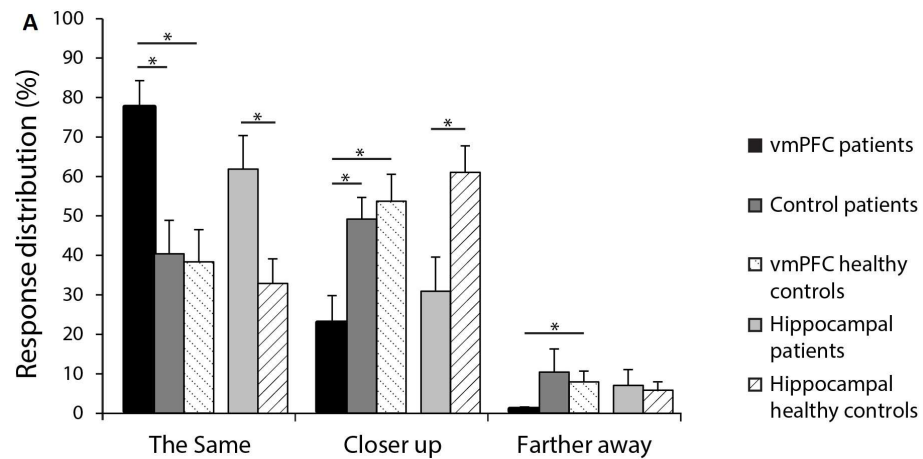


Figure 13. BE in vmPFC-damaged patients, hippocampal-damaged patients, and control participants, including brain-damaged control patients. Patients with vmPFC or hippocampal damage showed attenuated BE. $*p < 0.05$. Adapted from De Luca et al. (2018).

Damage to the vmPFC results in numerous other deficits beyond those I have considered here (see McCormick et al., 2018a for a full review). However, it is worth noting that the patients often display issues with moral decision-making (e.g. Levens et al., 2014), emotion regulation (e.g. Koenigs et al., 2007), social interactions and theory of mind (e.g. Beer et al., 2006). Overall, neuropsychological studies of vmPFC-damaged patients seem to indicate that they have difficulty focusing internally, have problems with monitoring, often suffer failures of initiation, and struggle with inhibition.

1.4 A simple model of scene and event construction

The fMRI and neuropsychological findings outlined above have been influential in informing a simple model of how the hippocampus and vmPFC may interact to mentally represent scenes and unfolding events (McCormick et al., 2018a; Ciaramelli et al., 2019). In this model, the vmPFC initiates the selection of relevant scene elements stored in the neocortex (e.g. objects typical of a residential street; Figure 14), inhibiting those irrelevant to the current scene (van Kesteren et al., 2010a, 2012; Ghosh et al., 2014). These elements

are then made available to the hippocampus that uses them to construct a spatially coherent scene snapshot from the event (Hassabis and Maguire, 2007; Mullally and Maguire, 2013a). Through continual feedback loops between vmPFC, neocortex and hippocampus, the vmPFC co-ordinates the generation and then linking of a stream of scenes that gives rise to a seamless evolving event.

This model offers a simple explanation of how multiple individual scenes results in extended mental events despite the fact that we selectively attend to our complex environment, shifting our gaze, and regularly blinking (Cutting, 2005). It is helpful to reprise an analogy introduced at the start of this chapter; while a movie is composed of individual frames (scenes), when run sequentially, these frames give us the impression of being seamless and unfolding. I suggest that the brain might manage our scene and event construction in a similar manner.

This model represents a departure from other accounts of hippocampal and memory processing. Crucially, scenes are central to this model, regarding them as the basic unit of events over and above other types of representation (Hassabis and Maguire, 2007; Maguire and Mullally, 2013). This contrasts with the view that the hippocampus supports associative processing in general, which assigns scenes to a broader category of stimuli that requires the binding of multiple elements together, such as objects and space (Lee et al., 2005; Eichenbaum, 2006; Aly et al., 2013; Erez et al., 2016; Ekstrom and Ranganath, 2017; Yonelinas et al., 2019). The fMRI studies I outlined earlier provide evidence that the hippocampus and vmPFC are preferentially engaged by scene processing, yet this remains an active area of debate. In Chapter 3 I sought to provide further empirical evidence to help adjudicate between these two perspectives.

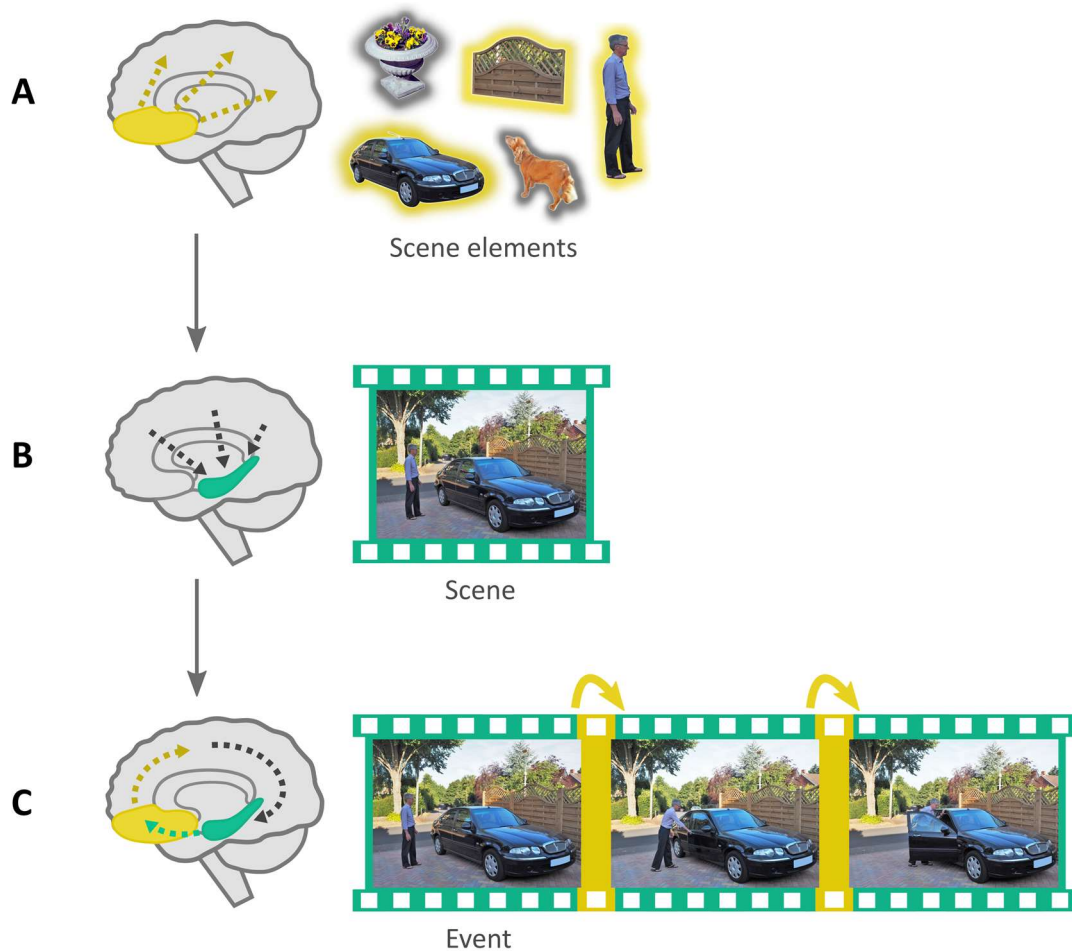


Figure 14. Schematic illustration of a model of scene and event construction. This account places the vmPFC (yellow) as the co-ordinator of scene and, in turn, event construction. **(A)** The vmPFC initiates the selection of relevant scene elements stored in neocortical areas, which are then **(B)** conveyed to the hippocampus (green), where they are constructed into a spatially-coherent scene. **(C)** Iterative feedback loops between the vmPFC, neocortex, and hippocampus help to produce and integrate the successive scenes that comprise an unfolding mental event (McCormick et al., 2018a; Ciaramelli et al., 2019).

This simple model is also distinct from the cognitive map theory, which suggests the primary role of the hippocampus is to support allocentric spatial processing (O'Keefe and Dostrovsky, 1971; O'Keefe and Nadel, 1978; Byrne et al., 2007; Epstein et al., 2017; Spiers, 2020). Scene construction, by contrast, emphasises the egocentric perspective (Hassabis and Maguire, 2007; Maguire and Mullally, 2013; Clark and Maguire, 2016). Some patients with bilateral hippocampal damage were found to be able to navigate

effectively through environments learned long before the onset of amnesia, including describing possible detours. This was especially the case when the layout of environments was regular or major roads could be used (Teng and Squire, 1999; Rosenbaum et al., 2000). Particularly striking is the fact that these patients were able to accurately complete other spatial tasks such as the estimation of distance and direction in these familiar settings, which rely on allocentric processing. Therefore, it appears that allocentric computations can be performed without a functioning hippocampus. This seems difficult to reconcile with the cognitive map theory. Notably, the patient reported by Maguire et al. (2006) was not able to visualise in advance where to turn off from one road to another whilst navigating egocentrically, nor was he able to imagine scenes (Hassabis et al., 2007b).

1.5 Predictions and thesis overview

The evidence assembled in this chapter and the simple scene and event construction model I outlined (McCormick et al., 2018a), motivated the predictions in this thesis. First, I predicted that the hippocampus and vmPFC would display preferential responses to scene processing. Second, I hypothesised that the vmPFC would initiate scene and event processing. As a consequence, the vmPFC should respond earlier than the hippocampus. Third, given the role of the vmPFC in orchestrating scene and event construction, it should drive hippocampal activity. These predictions call for a technique that is able to temporally resolve the activity of the hippocampus and vmPFC. The poor temporal resolution of fMRI makes it difficult to draw any conclusions about the relative timing of responses in different brain regions. By contrast, the millisecond temporal resolution afforded by MEG, combined with its reasonable spatial resolution and whole-brain coverage, makes it particularly useful to non-invasively study both the neural dynamics of scene and event construction, and the effective connectivity between brain regions.

Relative to EEG data, source reconstruction of MEG data is more feasible, permitting the measurement of direct neural activity and connectivity even in deep structures (e.g. Poch et al., 2011; Kaplan et al., 2012, 2017; Dalal et al., 2013; Staudigl and Hanslmayr, 2013; Backus et al., 2016; Meyer et al., 2017). This point is discussed further in Chapter 2.

There are relatively few MEG and EEG studies on scene and event processing. To date, electrophysiological investigations have typically used movies as event stimuli to test the feasibility of MEG or EEG to reveal consistent patterns of neural activity across participants during movie viewing (Lankinen et al., 2014; Chang et al., 2015). More recently these scanning methods have been employed to examine the implicit tendency to segment our experiences (Silva et al., 2019). However, the underlying temporal dynamics of scenes and events remain essentially unaddressed questions.

MEG and EEG studies that explore the hierarchical relationship between the hippocampus and vmPFC are especially scarce. The most directly relevant study is that of Barry et al. (2019a), who tasked participants with immediately imagining full scenes in response to single scene-evoking cue words, such as “forest”, whilst undergoing an MEG scan (also see Barry et al., 2019b). Changes in anterior hippocampal and vmPFC theta were evident in response to imagining scenes, with the vmPFC engaging significantly earlier than the hippocampus. Moreover, the vmPFC drove hippocampal activity (Figure 15) during scene imagination when effective connectivity was assessed using dynamic causal modelling (DCM; Friston et al., 2003; Friston, 2009; Kiebel et al., 2009). Kaplan et al. (2017) also found neural activity in these two regions to be synchronous during the imagination of scenes. Several MEG studies have examined autobiographical memory recall (Fuentemilla et al., 2014, 2018; Jafarpour et al., 2014; Hebscher et al., 2019, 2020),

also noting theta power changes in the medial temporal lobe (MTL) and medial prefrontal cortex (Fuentemilla et al., 2014). Oscillatory activity within the theta range, in particular, is often pinpointed as the means for disparate brain regions to communicate during episodic memory processes involving hippocampal and prefrontal regions, in both MEG and EEG studies (Colgin, 2013; Fuentemilla et al., 2014; Garrido et al., 2015; Backus et al., 2016; Berens and Horner, 2017; Kaplan et al., 2017; Sans-Dublanc et al., 2017; Nicolás et al., 2021), with evidence for theta underlying associative processing in the brain (Herweg and Kahana, 2018; Herweg et al., 2020). While the findings outlined here broadly support the model I have described in Section 1.4, a key test concerns how scene and event construction are initiated.

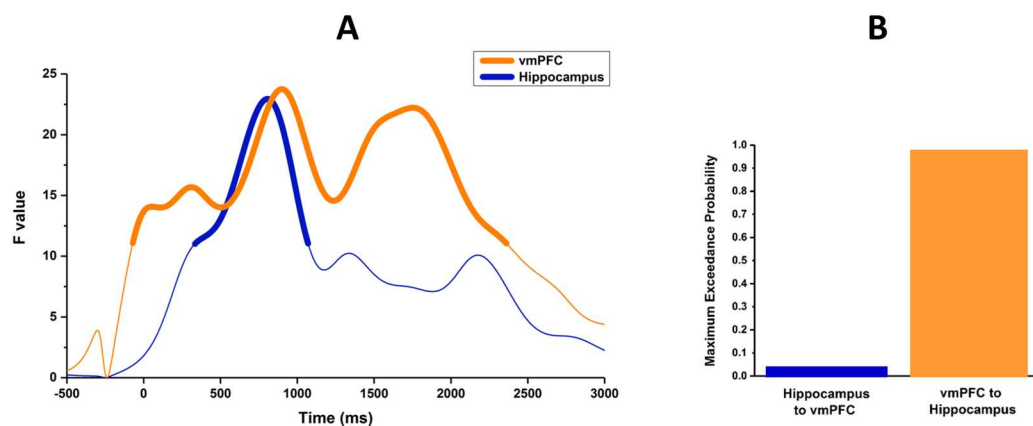


Figure 15. Hippocampal-vmPFC interactions during the construction of novel scene imagery. (A) An early peak in vmPFC (orange) theta power preceded that of the hippocampus (blue) following the onset of the 3000 ms scene imagination period. Time periods where power changes significantly differed from baseline are indicated by a thickened line ($p < 0.05$ FWE). **(B)** Results of the DCM analysis showing evidence for each of two competing models. Stronger evidence was found for the vmPFC driving hippocampal activity (orange bar). Adapted with permission from Barry et al. (2019a).

Only one study has examined this issue, and specifically in relation to autobiographical memory retrieval – McCormick et al. (2020). They asked healthy participants to recall autobiographical memories, and found that the hippocampus and

vmPFC showed the greatest power changes across the whole brain. Interestingly, the response of the hippocampus lagged significantly behind that of the vmPFC during the very earliest phase of retrieval. Moreover, using DCM, the authors showed that the vmPFC drove hippocampal activity during recall initiation and the subsequent unfolding of the events in memory.

The results of this handful of MEG studies lend preliminary support for the idea that the vmPFC initiates scene and event construction, and drives the hippocampus. In this thesis I sought to fill in the gaps in our knowledge that lie between the Barry et al. (2019a) study – where single scenes were mentally constructed – and McCormick et al. (2020) study – where complex, naturalistic personal events were mentally re-constructed. Therefore, after outlining the general methods I employed in conducting a series of MEG studies (Chapter 2), the thesis is structured as follows:

Chapter 3: I begin by examining scene-preferential effects in the hippocampus and vmPFC, by adapting for MEG the paradigm originally used in fMRI by Dalton et al. (2018) to investigate how scene imagery is built. I compare this to the construction of non-scene arrays of objects that do not evoke scene representations.

Chapter 4: I then sought to investigate whether objects previously shown to elicit a sense of surrounding 3D space (space-defining) affect the neural dynamics of scene construction.

Chapter 5: To address how a single scene may then be prepared for inclusion into an event, I next examined BE using MEG.

Chapter 6: Finally, I directly examined how individual scenes might be linked together to form an evolving event. I did this by devising a novel paradigm to compare the construction of events built from scenes with that of events built from non-scene abstract patterns. Events were simple visual animations built from individual image frames, enabling me to look at the transition from a single scene to a linked event.

I then draw together the results from all of these experimental chapters for a discussion in Chapter 7, considering in particular what can be learned about the involvement of the hippocampus and vmPFC, and their relationship, in constructing scenes and events, and how this relates to the cognitive model proposed and the predictions I made.

1.6 Publications

The following publications arose from the work in this thesis:

- Monk AM, Dalton MA, Barnes GR, Maguire EA (2021) The role of hippocampal-ventromedial prefrontal cortex neural dynamics in building mental representations. *J Cogn Neuro* 33:89-103.
- Monk AM, Barnes GR, Maguire EA. (2020) The effect of object type on building scene imagery – an MEG study. *Front Hum Neurosci* 14:592175.

- Monk AM, Patai EZ, Barnes GR, Intraub H, Maguire EA. Very early neuronal responses are associated with top-down scene processing (submitted).
- Monk AM, Barry DN, Litvak V, Barnes GR, Maguire EA. Evoked response signatures during free viewing of animated movies distinguish unfolding events built from scenes and non-scenes (in preparation).

During the course of my PhD I was also involved in the following publications, not described in detail in this thesis:

- Clark IA, Monk AM, Maguire EA (2020) Characterising strategy use during the performance of hippocampal-dependent tasks. *Front Psychol* 11:2119.
- Clark IA, Monk AM, Hotchin V, Pizzamiglio G, Liefgreen A, Callghan MF, Maguire EA (2020) Does hippocampal volume explain performance differences on hippocampal-dependent tasks? *NeuroImage* 221:117211.
- Ciaramelli E, De Luca F, Monk AM, McCormick C, Maguire EA (2019) What "wins" in VMPFC: Scenes, situations, or schema? *Neurosci Biobehav Rev* 100:208-210.
- Clark IA, Hotchin V, Monk A, Pizzamiglio G, Liefgreen A, Maguire EA (2019) Identifying the cognitive processes underpinning hippocampal-dependent tasks. *J Exp Psychol Gen* 148:1861-1881.

2 General methods

2.1 Précis

In this chapter I describe the methods used to collect and analyse the data. First, I explain the set-up that I employed across all my experiments for the collection of behavioural and neuroimaging data. I then discuss the basic principles of measuring brain responses using MEG, the neuroimaging technique that I used for all of my experiments. Following this, I describe the pre-processing steps involved in preparing the raw MEG data for analysis, the statistical methods I used to subsequently analyse the data, and the software toolboxes utilised to implement them. The core analytic methods I employed ranged from an examination of evoked responses at the sensor level to effective connectivity analyses at the source level.

2.2 Participants

All participants were aged between 18 and 35 years old, were right-handed, had either normal or corrected-to-normal vision, had no history of neurological or psychiatric disorders, and were fluent in English. Participants were a mixture of undergraduate and masters-level students, and non-students recruited from a range of sources: the UCL Institute of Cognitive Neuroscience's participant database, the UCL Division of Psychology and Language Sciences SONA participant database, the Call for Participants volunteers pool (<https://www.callforparticipants.com>), recruitment via masters student programme emailing lists at UCL, Birkbeck University of London, and Queen Mary University of London, as well as Facebook posts. When students were recruited, an effort

was made to include people studying on non-psychology courses, who were more likely to be naïve to the aims of the experiments. Except for experiments reported in Chapters 4 and 5, participants only took part in one experiment. For these two exceptions, the experiments were typically run on the same day; the shorter Chapter 4 experiment took place in the morning, and the Chapter 5 experiment in the afternoon, following a 2-hour break.

Before being assigned an appointment, participants were screened on the telephone and completed a written screening questionnaire to ensure they were suitable for a scan in MEG. They also read the participant information sheet and had the opportunity to ask questions. Participants had no history of surgery involving the insertion of metal implants or devices, no dental wires or metal dental fillings. On the day of their scan, they were instructed to wear non-magnetic clothing, wear no eye make-up or hair products and to remove all metal items (such as piercings). All participants gave written informed consent to participate in accordance with the University College London Research Ethics Committee (project ID: 1825/005). All data were acquired before the COVID-19 pandemic.

2.3 Overview of the experimental set-up

All experiments were conducted at the Wellcome Centre for Human Neuroimaging, UCL Queen Square Institute of Neurology, University College London, UK. Each experiment involved a behavioural and a neuroimaging component. Depending on the experiment (as detailed in the experimental chapters that follow), the behavioural element included pre-scan training, and post-scan debriefing or testing. Behavioural sessions were conducted in a quiet room at the Centre designed for the purpose of experimental testing,

where participants sat in front of a desktop PC. The neuroimaging component was conducted using an MEG scanner. The MEG system is described later in Section 2.5.1.

In each experiment, upon arrival, a participant gave written informed consent after re-reading the participant information sheet, and they had another opportunity to ask questions. The pre-scan training session was then completed on a desktop PC. Prior to entering the scanning environment, a participant was thoroughly screened, and asked to remove any metal items or make-up that could introduce artefacts to the MEG data. Inside the MEG scanner, a participant was in a seated position, with their head resting on the back of the MEG helmet. The set-up was consistent across all participants and experiments. Only the back-rest of the MEG chair was occasionally adjusted between participants, to ensure a comfortable posture. Cushions and foam pads were placed at a participant's sides and under their legs for further comfort and to prevent excessive movement.

Head position fiducial coils were attached to the three standard fiducial points on a participant's face (nasion, left and right preauricular) to monitor head position continuously throughout acquisition. During scanning, the door to the magnetically shielded room, within which the scanner is located, was closed (see Section 2.5.2.1). A participant was monitored throughout the scanning session via a camera, and a microphone enabled communication between a participant and experimenter from outside the scanner. A participant was also provided with an emergency alarm bell, taped to the left arm rest to be within easy reach, which they could use at any point during scanning to indicate they wished to exit the scanner.

An MEG-compatible (containing no metal components) 5-way button-box was held in the right hand to make responses. The recording of responses and delivery of visual stimuli was made using the Cogent2000 toolbox for Matlab (<http://www.vislab.ucl.ac.uk/cogent.php>). Visual stimuli and task instructions were presented via a projector (60 Hz frequency) situated outside the scanner room, projecting onto a screen via a mirror. The distance between the screen and a participant's eye was ~60 cm. A photodiode was used to accurately record the onset of stimuli during experiments that involved visual presentations (Chapters 5 and 6). I measured a jitter of approximately $36\text{ms} \pm 0.3\text{ms}$ between the trigger codes sent from the stimulus PC to the MEG acquisition computer and the projector screen viewed by a participant, which was found to be consistent across trials in both relevant experiments. I accounted for this jitter during the pre-processing of the data. When required, auditory stimuli were delivered via an MEG-compatible Etymotic Ear-Tone stereo sound system (Etymotic Research Inc, Illinois; Chapters 3 and 4). Sound travelled via plastic tubes to inner-ear foam inserts. Sounds were played using the Cambridge Research Systems Audiofile (Cambridge Research Systems Ltd) via commands coded in a Matlab script.

An Eyelink 1000 Plus (SR Research) eye tracking system with a sampling rate of 2000 Hz was used during MEG scanning to monitor task compliance, measure oculomotor behaviour, and compare experimental conditions. The eye tracker camera was positioned on a table in front of a participant. The right eye was used for both calibration and data acquisition, to record eye tracking across the full screen. On occasions when participants were missing eye tracking data, typically as a result of insufficiently accurate calibration, the number of participant samples available for eye tracking analysis is reported in the specific experimental chapter.

After scanning, a participant was behaviourally tested, if this was part of the experimental protocol, and debriefed. Each participant was paid £10 per hour to cover their travel and meal expenses.

2.4 The origin of the MEG signal

As the experimental questions I was interested in addressing required high temporal resolution, MEG was the method of choice to record brain responses in all my experiments. Consequently, I next briefly explain the physiological and biophysical basis of the signal measured in MEG. I also mention how MEG compares to EEG and fMRI, and provide an overview of how inferences are made about brain function using MEG technology. For further detail on the basic theoretical concepts underpinning the MEG signal, and the resultant considerations for analyses, I refer the reader to Hansen et al. (2010).

2.4.1 *MEG signal generation*

The brain communicates information via neural electrical activity, which also generates magnetic fields. Both electrical current flows, and the magnetic fields they give rise to, reach the scalp surface where they can be measured non-invasively outside of the brain, by EEG and MEG respectively. Post-synaptic (dendritic) activity in pyramidal cells of the cortex are thought to contribute the most to the measurable MEG (and EEG) signal, since these are oriented in parallel to one another and aligned perpendicularly to the cortical surface. The current flow along dendrites is dipolar and behaves much like electricity travelling along a wire, following the well-established right-hand rule of electromagnetism. Approximately 50,000 pyramidal neurons are required to be

synchronously active to generate neural activity measurable by MEG, depending on how these neurons are distributed within cortex, their relative alignment, and their orientation with respect to the pial surface (Figure 16). The sum of many dipolar pyramidal dendrites forms one “dipole”. Pyramidal cells are found mainly in layers II/III and V, with those in layer V being particularly large and therefore thought to contribute more to the MEG signal (Murakami and Okada, 2006).

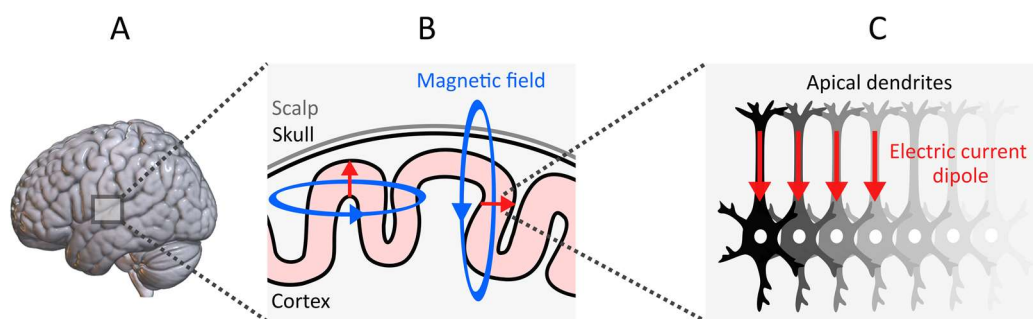


Figure 16. The generation of electric and magnetic fields in the brain. (A) The cortical surface is highly folded, creating gyri and sulci. Visualisation of a rendered brain was performed using MRICroGL (<https://www.nitrc.org/projects/mricrogl>). (B) Magnetic fields (blue arrows) of current dipoles (red arrows) are shown for a radial electrical current (gyrus; left-hand arrow), and a tangential electrical current (sulcus; right-hand arrow), in a perfectly spherical conductor (an approximation of the human head). Magnetic fields resulting from tangential dipoles are more detectable at the scalp. The magnetic field wraps around the current dipole, so this should be interpreted as a 3-dimensional ring. (C) A representation of a pyramidal cell assembly in the cerebral cortex needed to generate a detectable magnetic field, like the one illustrated in B. Approximately 50,000 synchronous dendritic currents are required for a detectable signal at MEG sensors.

2.4.2 Comparison between MEG and other imaging techniques

A change in the magnetic fields produced by neuronal current flow serves as a direct measure of “online” neuronal activity, allowing the tracking of responses at the millisecond level. This is notably in direct contrast to fMRI, where the latency of the haemodynamic response – the neural signature of fMRI – is slow (in the order of ~4 seconds). The fMRI signal also typically reflects changes in oxygen consumption, and so

is not a direct measure of neural activity. MEG ostensibly offers a more direct insight into rapid cognitive processes in the timeframe in which they occur (i.e. sub-second), where the temporal resolution is only limited by the sampling rate set during acquisition. All of the MEG data in this thesis were acquired at a sampling frequency of 1200 Hz.

When comparing MEG with EEG, it may initially appear there is little difference between them to warrant the use of one over the other. Although the origin of the signal measured by both technologies is relatively similar, magnetic fields are less distorted by the skull and scalp than electric fields. By the time the EEG signal reaches the scalp, it is severely smeared (Wolters et al., 2006). Therefore, although both technologies allow for non-invasive direct recordings of neural activity at a high temporal resolution, MEG provides better spatial resolution ($\sim 5\text{-}8$ mm; Brookes et al., 2010) than EEG, allowing researchers to answer not only questions pertaining to the timing of an effect, but also its likely brain region of origin.

This advantage in terms of spatial resolution has some caveats. Both MEG and EEG are sensitive to the folded nature of the brain (Figure 16A). Whether the signal is generated at the gyri (top of folds) or the sulci (inside folds) affects the ability of each modality to detect that signal. Given a spherical conductor (i.e. the approximate shape of the head), and the fact that a current flow generates a magnetic field clockwise around it, then current flow perpendicular to the scalp is more detectable by EEG, whereas current flow parallel to the scalp is more detectable by MEG (Figure 16B). Therefore, MEG is particularly good at detecting signals arising from sulci (mostly tangential dipoles), since pyramidal cells are aligned perpendicularly to the cortical surface, but is less effective at detecting signals arising from gyri (mostly radial dipoles). EEG is sensitive to both radial and tangential dipoles. However, the head is not a perfect sphere so, in reality, the radial

orientation is not well defined, meaning radially-oriented (gyral) sources are not invisible to MEG (Cuffin, 1990). Only a small area of cortex at gyral crests has truly poor resolvability (estimated to only be around 5%), therefore the radial dipoles that go undetected account for only a small proportion of potentially active cell assemblies (Hillebrand and Barnes, 2002).

2.5 Instrumentation

This section briefly describes the core components of the MEG system, including how the small magnetic signals generated by the brain are recorded, and the techniques that were employed to reduce the noise that would otherwise contaminate the measurements.

2.5.1 *The MEG scanner*

The MEG scanner used was a 275-sensor CTF Omega MEG system, housed in a magnetically shielded room (MSR; Figure 17A). The sensors, capable of detecting the small magnetic signals generated by the brain, are called Superconducting Quantum Interference Devices (SQUIDS), and are arranged in a helmet-like structure (also called a dewar; Figure 17B) as close to the scalp as possible. Liquid helium is used to keep SQUIDS cool, maintaining a temperature of approximately -270°C , as they use a superconducting loop made of niobium. The CTF MEG system contains axial (also called radial) gradiometers as flux transformers, magnetically coupled to each sensor (Figure 17C).

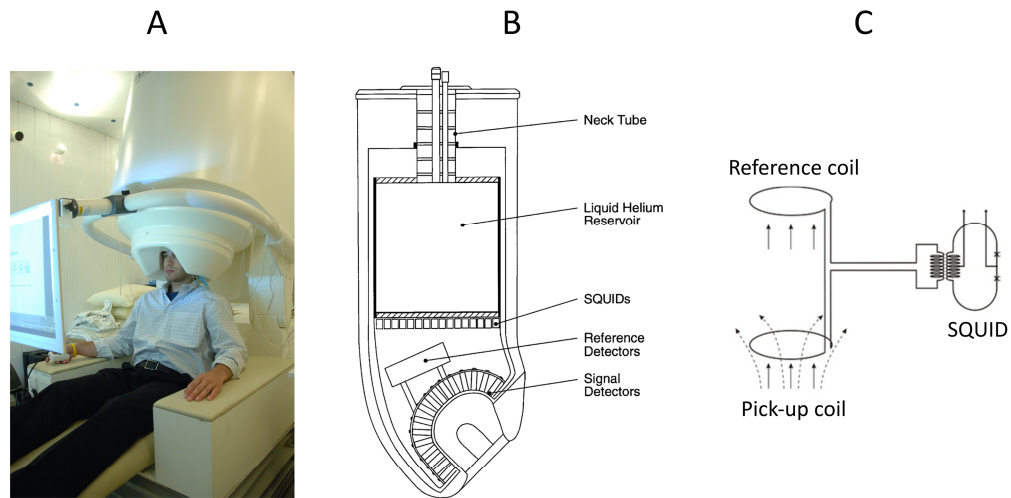


Figure 17. The MEG system. (A) The MEG scanner situated inside a magnetically shielded room, with a participant ready to be scanned in a seated position, looking at a screen displaying task stimuli. Courtesy of the National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services (http://infocenter.nimh.nih.gov/il/public_il/image_details.cfm?id=80) and copyright free. (B) Illustration of the MEG helmet-like structure, called the dewar. A large tank of liquid helium is used to lower the temperature of the SQUIDS to -270°C , used to detect magnetic fields produced by neural activity. Reference sensors help to reduce ambient noise (adapted from König et al., 2016, and copyright free). (C) An axial gradiometer flux transformer illustrated from the side. The two coils of the gradiometer greatly reduce noise from sources outside the brain, with a pick-up coil nearest to a participant's head. This works by applying magnetic flux to cause oscillations in a SQUID (one sensor is shown on the right; adapted with permission from Andrä and Nowak, 2006).

Gradiometers measure the gradient of magnetic flux between two coils (where one is positioned above the other), and are most sensitive to the magnetic fields perpendicular to the scalp surface. Due to their configuration, they considerably reduce environmental noise. The CTF system also contains an additional set of reference sensors that can be used to synthesise third-order gradiometers (see 2.5.2.1). These reference sensors measure interference and not neural activity. Data were digitised continuously at a sampling rate of 1200 Hz. I next explain how noise is minimised to obtain the most accurate MEG data possible.

2.5.2 *Noise reduction during data acquisition*

SQUIDS are capable of detecting miniscule differences in magnetic fields generated by the brain, which typically lie in the range of 10-1000 femtoTesla (fT)¹. Due to their sensitivity, however, SQUIDS also pick up magnetic fields in the surrounding environment. The earth's magnetic field alone is around 50 microTesla (μ T) and other signals in a typical urban environment can be as high as 1 μ T, far larger than the brain signals of interest (Figure 18). In addition, magnetic fields generated by the heart, eyes and muscular activity can contaminate the MEG recording. Consequently, different methods have been developed to reduce the influence of these external environmental, and internal physiological artefacts.

2.5.2.1 *External magnetic fields*

The main methods used to reduce noise from the surrounding environment involve axial gradiometers, passive shielding, and the option of applying synthetic gradiometry.

Axial gradiometers

The data from the CTF MEG system used were recorded with axial gradiometer SQUIDS (Figure 17C). Two aligned sensor coils (magnetometers), axial with respect to a participant's head, are wound in opposite directions so that the effect of distant magnetic fields is subtracted out, whilst retaining nearby magnetic fields measured at the scalp. The effect is a substantial reduction in ambient noise and better detection of magnetic fields at low frequencies emanating from the brain.

¹ Tesla is a unit of magnetic induction. 1 fT = 10^{-15} T; 1 μ T = 10^{-6} T

Passive shielding

In order to reduce external magnetic fields, the MEG scanner is housed in a MSR. This provides what is referred to as passive shielding against magnetic noise from the external environment, reducing it to around 10-20 nanoTesla (nT). Devices that may generate a magnetic field (such as electrical equipment containing magnetic components, including the projector used to display stimuli) are kept outside the MSR.

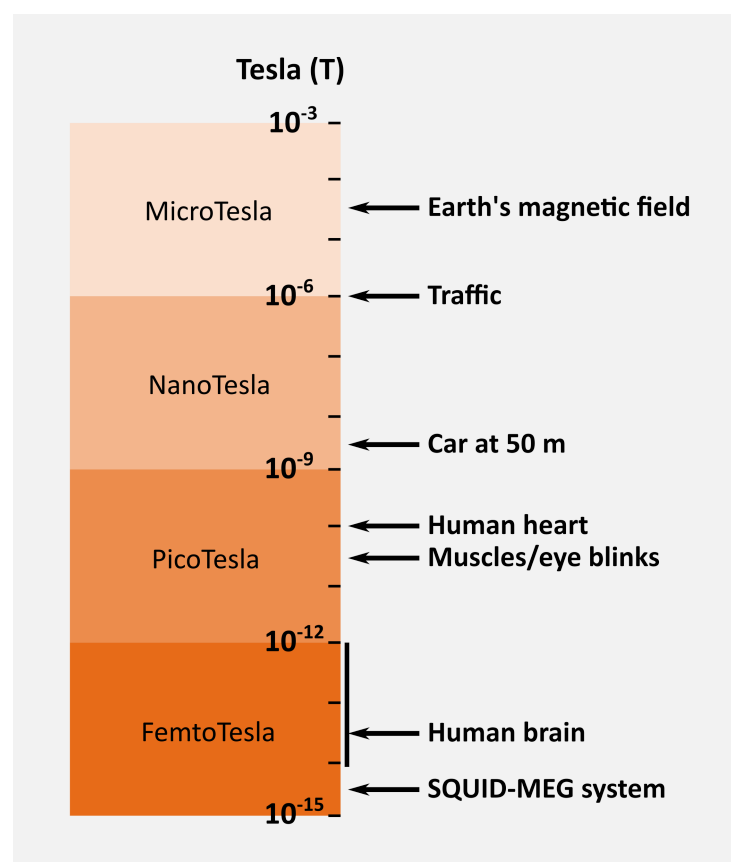


Figure 18. Strengths of physiological and environmental magnetic fields. The relative sizes of magnetic fields measured with MEG, comparing the human brain with other biomagnetic and environmental sources of noise that can contaminate MEG measurements. Attenuation of noise not originating from the brain is critical during data acquisition. The scale of magnetic fields is in Tesla (T).

Synthetic gradiometers

The CTF MEG system also provides the opportunity to create third-order synthetic gradiometers to reduce environmental interference, by using the signals from the reference sensors inside the MEG scanner to measure and subtract ambient magnetic noise. Although using this configuration is useful to improve the signal-to-noise ratio (SNR), it is not always necessary. All participants in this thesis were neurotypical, with no metal implants, and great care was taken during both recruitment and data acquisition to ensure each dataset was undisrupted by metal artefacts. This, too, is important in order to optimise the spatial reconstruction of the data by minimising the SNR. Furthermore, spatial specificity depends also on the accuracy of the forward modelling solution (i.e. models of the changes in magnetic fields due to a specific source of electrical activity). This is described in Section 2.7.1.

One experiment in this thesis did not use synthetic gradiometry (Chapter 3). This was due to a technical failure during data acquisition that meant synthetic gradiometry was turned off, and this remained undetected until after the completion of analyses. All analyses in this experiment were concerned with changes in spectral power, therefore it was important to check the accuracy of the source reconstruction in the absence of this higher-order gradiometry. I selected one source reconstruction result (i.e. where synthetic gradiometry was not applied) and re-ran the analysis with synthetic gradiometry retrospectively applied. Figure 19 clearly demonstrates the results were almost identical, including the detection of activity in deeper sources (such as the thalamus, in this example), suggesting the findings were robust.

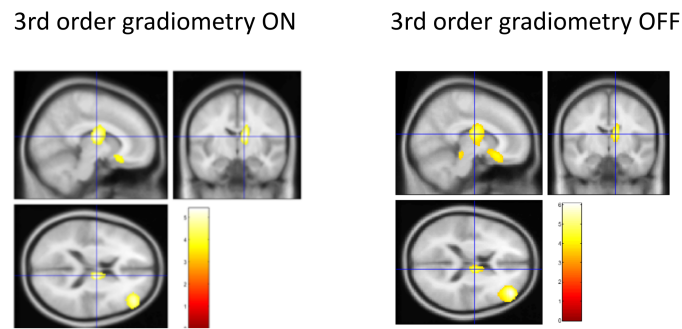


Figure 19. Effect of third-order gradiometry on results of source reconstruction. Comparing the results of source reconstruction using linearly constrained minimum variance (LCMV) beamforming demonstrates the effect of applying synthetic third-order gradiometry (left) is minimal in this case.

2.5.2.2 *Physiological magnetic fields*

Artefacts can also arise from physiological sources.

Muscular artefacts

Muscle tension can be reduced during data acquisition, and its effects on neural signals can be easily removed before analysis if necessary. In particular, artefacts resulting from muscular tension and bodily movement (e.g. in the head, neck, upper body) can be minimised by ensuring a participant is comfortable during scanning and that they take regular breaks. As a result, in addition to the set-up described in Section 2.3 that served to make a participant feel comfortable throughout scanning, the duration of individual scanning blocks in each experiment was between 6 and 7 minutes. Longer than this would be uncomfortable and lead to tension, movement and fatigue. A participant was also instructed to maintain a fixed head position, continuing to rest their head on the back of inside of the MEG helmet structure (which contains the sensors; see Figure 17B), and this was checked at the start of each scanning block.

Head movement

Small movements of a participant's head were recorded during data acquisition using localisation coils, which provided continuous measurement of the head position (see the set-up description in Section 2.3). Participants who moved more than 5mm at any one time were given further instructions about keeping still, and consistent movement of more than 6mm resulted in exclusion of data from subsequent analysis.

Ocular artefacts

Physiological artefacts can also arise from saccadic eye movements and blinking. The eye can be modelled as a magnetic dipole, therefore it can distort electrical fields coming from nearby brain regions, or might be erroneously attributed to brain activity. This is particularly problematic for frontal midline regions, such as the vmPFC. Of course, one cannot prevent a participant from moving their eyes completely, but a problem arises if the artefact is non-random; for example, if a participant always blinks at the onset of a stimulus for one condition, but not for the other condition to which it will be directly compared. In this case, and if low frequencies are of primary interest (as is the case in several of my experiments), eye movements could significantly contribute to any difference observed in the MEG signal. For this reason, tight control of eye movements is critical. However, what is the best strategy to achieve this?

The majority of MEG studies simply remove trials, or sections of them, as an easy fix for this problem. However, this loss of data may not be feasible if only a small number of trials are available, or for paradigms where oculomotor behaviour is systematic in particular conditions. Indeed, the gross removal of eye movements altogether may serve to “over clean” the data such that the effect of interest is lost. It is important to remember that eye movements and fixations are an important, natural part of human cognition and

behaviour – this is particularly relevant to the processing of naturalistic stimuli, such as scenes, which are used throughout this thesis.

Instead of discarding data in this way, paradigms that guide eye movements can control the distribution of fixations between conditions so that very similar fixation patterns are generated. The occurrence of eye movements, and their contamination of between-condition differences, can therefore be minimised in the first place. This is the approach I adopted to control eye movements and fixations. I also recorded eye data in real-time across the whole screen, using a high-quality eye tracker with a high sampling rate (2000 Hz; also see Section 2.3). This fine temporal resolution permits tracking of even small eye movements, which might otherwise be missed by using other methods, such as electrooculography. By combining temporally-resolved eye tracking and MEG, both could be recorded in real-time to determine whether eye movements and fixations could reasonably have had an influence on the MEG results.

2.6 MEG data pre-processing

To prepare raw MEG data for analysis, it must first be pre-processed. Here, I briefly describe the different pre-processing steps followed for each experiment, performed in Statistical Parametric Mapping version 12 (SPM12; Wellcome Centre for Human Neuroimaging, London, UK; <https://www.fil.ion.ucl.ac.uk/spm>). These steps include filtering, segmenting data into epochs, and the removal of possible artefacts. Baseline-correction is also mentioned here, the use of which varied between experiments due to differences in study design, which are discussed further in the specific experimental chapters.

2.6.1 Filtering and epoching

MEG data were filtered in several ways. First, a high-pass filter was applied in order to eliminate slow drifts in signals from the MEG sensors, followed by a stop-band filter of 48-52 Hz to remove the power line interference, and finally a stop-band filter of 98-102 Hz to remove its first harmonic.

Data were then epoched according to which experimental conditions were of interest. A baseline period was additionally epoched to be subtracted from each condition in certain circumstances. Specific details on how data were epoched and baseline-corrected are given in the each experimental chapter. General considerations that pertain to baseline-correction are, however, described in Section 2.6.2.

2.6.1.1 A note on high-pass filtering choices

Data were high-pass filtered at either 0.1 Hz or 1 Hz. In studies designed to examine event-related fields (ERFs), particular caution is needed when applying a high-pass filter. In cases where I used an ERF analysis approach (Section 2.7.2), a standard cut-off of 0.1 Hz was used. Greater than this (i.e. towards 1 Hz) could be problematic, as this could alter the shape of the resulting averaged ERF waveform, especially if the ERF is expected to occur early on after stimulus onset, and the duration of the epoch analysed is particularly short. These circumstances were certainly true for ERF experiments in Chapters 5 and 6, where a rapid effect was predicted (and found) and epochs lasted between only 250 and 700 ms respectively. A more rigorous cut-off frequency of 1 Hz was used in Chapters 3 and 4, because the epochs were longer (3000 ms in both cases), and where oscillatory power was the focus.

2.6.2 *Baseline-correction*

Baseline-correction is commonly applied to data before analysis as a means to remove bias introduced by drifts in the MEG signal. This is because there is effectively no “zero” level to the SQUIDS, and there is likely to be some low-frequency noise from both the sensors and the surrounding environment. This is normally achieved by subtracting the average of a specified pre-stimulus time period. This period should, of course, not contain any event-related activity, nor should it overlap with it, yet it should be close in time to the onset of the event. For different reasons, this may not always be feasible, and under some circumstances it is in fact undesirable.

First, baseline-correction by subtraction may be unnecessary altogether, since the important step of high-pass filtering (discussed above in 2.6.1.1) removes slow drifts and low-frequency components (Gross et al., 2013). There is some debate as to whether baseline-correction is an adequate substitute for high-pass filtering (see Maess et al., 2016 for a discussion on this topic). The efficacy of baseline-correction depends on the baseline period selected – it cannot be assumed there are no systematic differences between baselines for different conditions, and a duration of several hundred milliseconds would be required for a better, and more precise, estimate of its true mean. Therefore, baseline-correction ought to be an additional step, if useful, following the standard approach of high-pass filtering.

Second, rather than assuming baseline-correction is necessary (because it is a commonly used technique), a better approach would be to choose the baseline window according to the question being addressed by the study. For example, if a pre-stimulus baseline window of adequate length would be disruptive to the design, baseline-correction

may be unwise. Instead, low-frequency noise can be effectively minimised using a combination of high-pass filtering and practical steps employed during acquisition (see Section 2.5.2). Another option is for a form of baseline-correction to be implemented at a later stage of analysis – that is, rather than part of pre-processing, an appropriate baseline window could be subtracted from each condition at the single-subject level of analysis (also termed the first level). This was the approach used for experiments in Chapters 3 and 6. Furthermore, in the absence of a “clean” non-overlapped baseline close (i.e. immediately preceding) to the event of interest, it is advisable to choose a period with a greater degree of separation from task-related activity. This was the case for experiments reported in Chapters 3, 4 and 6, where there was no common, activity-free, baseline interval available immediately prior to each epoch. The main reason for this was that these studies required the continuous presentation of sequential epochs. The precise application of a baseline is described in the individual experimental chapters.

2.6.3 Artefact removal

Prevention of artefacts is always better than the cure. So far, approaches used to minimise the occurrence of different artefacts at the outset have been addressed, and in particular, the arguments against complete elimination of eye movements. Post-acquisition, I also visually inspected the data and rejected artefactual epochs using the FieldTrip visual artefact rejection tool, implemented in SPM12. This manual approach allowed me to selectively discard particularly noisy trials from subsequent analyses. However, steps taken during acquisition ensured that the data collected were of a high quality, significantly reducing the necessity to discard data. The extent of trial rejection is detailed in the specific experimental chapters.

2.7 MEG data analyses

Data were analysed with SPM12 and the FieldTrip toolbox (Donders Centre for Cognitive Neuroimaging, Nijmegen, the Netherlands; Oostenveld et al., 2011) within MATLAB R2018a (MathWorks, MA). I introduce the analytic methods employed in this thesis, and some of the main theoretical considerations for their use. In broad terms, methods examined changes in oscillatory power across different frequency bands (source reconstruction), transient evoked responses in the time-domain (sensor level ERFs), and effective connectivity between source-localised regions (DCM).

2.7.1 *Source reconstruction*

One of the main features studied in MEG data are induced neural oscillations. Changes in oscillatory power can be measured in response to different experimental conditions, localised to specific brain regions, as well as being time-resolved. The frequencies of oscillatory neuronal activity relevant to cognitive processes range between 0 and 200 Hz: commonly referred to as delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma (> 30 Hz). How is it possible to estimate the three-dimensional configuration of sources of this neural activity (i.e. its location, orientation, and time-course) when MEG sensors measure this activity from outside the brain?

2.7.1.1 *Source localisation and the inverse problem*

This transformation of data into source space is a major challenge for MEG. This is the so-called “inverse problem”. Even if the magnetic fields were measured at an infinite number of locations around the head, this would still be insufficient to uniquely estimate the distribution of current sources within the brain – it is a mathematically ill-posed

problem, with an infinite number of possible inverse solutions to any given forward model (Hämäläinen et al., 1993). Forward modelling refers to how changes in magnetic fields result from sources of electrical activity, and how this activity is propagated from the physiological source, through body tissues, before reaching the sensors where fields are observed. Unlike the inverse problem, the forward modelling problem is well-posed – electromagnetism is well-understood and there is only one solution for each current dipole.

Due to the difficulty of solving the inverse problem, it has been assumed that source localisation of focal brain activity in MEG (and EEG) is inaccurate. However, the development of beamforming has significantly improved this capacity, making it possible to accurately estimate the neural sources of activity measured with MEG (Van Veen et al., 1997; Hillebrand and Barnes, 2005; Brookes et al., 2008). Beamforming is an “adaptive” spatial filtering approach that estimates activity at predefined brain regions. This involves scanning the entire brain (i.e. the source space), sequentially scanning each location so that they are treated independently, so that a statistical volumetric image is generated that shows the brain regions engaged in a particular task. “Adaptive” here means that spatial filters rely on both the forward solution and the magnetic field measurements, and this typically leads to higher resolution than non-adaptive variants.

2.7.1.2 *Linearly constrained minimum variance (LCMV) beamforming*

In this thesis, the LCMV beamformer was used for source reconstruction, performed using the DAiSS toolbox (<https://github.com/SPM/DAiSS>) implemented in SPM12. This type of beamformer uses filter weights that linearly map the MEG sensors to source space, which are built based on data from the experimental conditions. Power is estimated

at each source location, whilst simultaneously minimising interference of contributions from other sources (thereby minimising the problem of leakage) and external noise captured in the data covariance matrix. This is why techniques like this are referred to as spatial filters (Van Veen et al., 1997). Importantly, this results in enhanced detection sensitivity of the target source activity. A frequency range is specified (e.g. 4-8 Hz if one is interested in theta activity), and two conditions are compared, to determine the distribution of brain activity that best explains the observed change in magnetic fields between conditions in this frequency range.

2.7.1.3 First-level analysis

In experiments where source reconstruction was performed, co-registration to the Montreal Neurological Institute (MNI) space was based on nasion, left and right preauricular fiducials (head positions), using a 5 mm volumetric grid. The forward model was computed using a single-shell head model (Nolte, 2003).

Here, a standard template brain was used for co-registration, but an alternative approach would have been to use individual structural MRIs. However, there is little evidence to suggest individual MRI images would provide any advantage in source localisation. In fact, these two approaches have been shown to produce very similar estimations of MEG activity (Holliday et al., 2003). Since brain anatomy varies greatly between participants, when functional images are combined for any group analysis, this will result in blurring (Thompson et al., 1996; Mazziotta et al., 2001). Therefore, any extra error introduced by using a standard template will have little impact on the source reconstruction. What is more important, instead, is to minimise the co-registration error, achieved by minimising participant head movement at acquisition.

At the first level, power in a particular frequency band was estimated to create one image per condition, per participant. The covariance matrix was estimated using a common spatial filter for all the conditions that would be later contrasted when performing group level statistics (the second level). This optimised the estimate and increased the reliability of the filters through which all trials would be subsequently projected. For some experiments, the time window of interest was particularly brief (<1000 ms; Chapters 5 and 6), which can adversely affect the beamformer estimation of source power. The LCMV beamformer is based on sensor measurements that are weighted, and these weightings are data-driven (Barnes and Hillebrand, 2003), based on the data covariance matrix specified. For accurate estimation of covariance, and therefore an accurate representation of source power, the window specified should ideally be longer, however the experimental design may impede this. This was true for Chapters 5 and 6, where the effects under investigation necessitated short time-windows (250 ms and 700 ms respectively). To mitigate this problem, the spatial filter was initially computed across a longer interval, encompassing time before and after the main (narrow) interval of interest. This has the advantage of a more stable covariance matrix. Whole brain power images per condition and per participant were subsequently generated only within the (narrow) window of interest.

2.7.1.4 *Spatial smoothing*

Images were smoothed using a $12 \times 12 \times 12$ mm Gaussian kernel before being entered into the second-level statistical test. Although this affects spatial accuracy, smoothing is a necessary step since neuronal currents display local coherence, and it helps to mitigate anatomical variability between participants.

2.7.1.5 *Second-level analysis*

At the second level, smoothed power images per participant were entered into a paired t-test, and between-condition contrasts were performed to address the study hypotheses. In this thesis, the output of the beamformer was a change in power between experimental conditions, visualised by superimposing it onto an anatomical MRI template image.

2.7.1.6 *Region-of-interest (ROI) analyses*

The source reconstruction was then constrained to a set of *a priori* defined ROIs, to determine the peak locations of power differences between conditions in each ROI. These analyses were performed using separate bilateral masks covering each region. Masks were created using the AAL atlas in the WFU PickAtlas software (<http://fmri.wfubmc.edu/software/pickatlas>). Each ROI was defined as follows across all experiments (also see Figure 20):

- **Anterior hippocampus:** defined from the first slice where the hippocampus can be observed in its most anterior extent until the final slice of the uncus (Poppenk et al., 2013; Zeidman and Maguire, 2016; Dalton et al., 2018).
- **Posterior hippocampus:** defined from the first slice following the uncus until the final slice of observation in its most posterior extent (Poppenk et al., 2013; Zeidman et al., 2015a; Dalton et al., 2018).
- **vmPFC:** comprising Brodmann areas 10, 14, 25, and parts of areas 32, 11, 12 and 13 (Catani et al., 2012; McCormick et al., 2018a).
- **PHC:** defined as the posterior portion of the PHC, positioned posteriorly from the first slice of the posterior hippocampus (Mullally and Maguire, 2011).

- **RSC:** comprising Brodmann areas 29–30 (Vann et al., 2009).

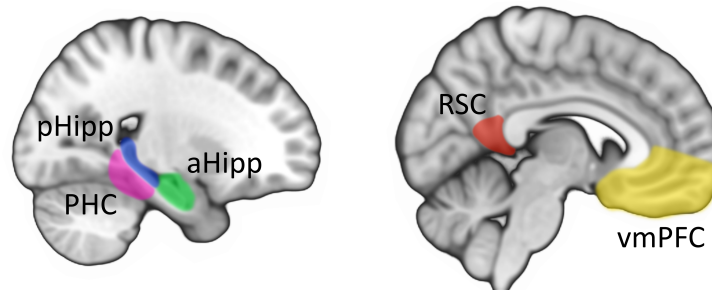


Figure 20. Anatomical regions of interest. Sagittal views showing the anterior hippocampus (aHipp; green), posterior hippocampus (pHipp; blue), ventromedial prefrontal cortex (vmPFC; yellow), posterior parahippocampal cortex (PHC; pink), retrosplenial cortex (RSC; red).

2.7.1.7 Statistical thresholding

When running beamformers to examine oscillatory effects (Chapters 3-4), the following statistical thresholds were used.

Based on strong *a priori* hypotheses from the extant literature about both the involvement of regions in the scene network (which, depending on the experimental paradigm, included the hippocampus, vmPFC, PHC, and RSC), and the relevance of theta and other low frequency oscillations, I applied a threshold of $p < 0.001$ (uncorrected for multiple comparisons) with a spatial extent threshold of >100 voxels for whole-brain analyses. This is held to provide a balance between protecting against false positives whilst enabling the detection of subtler signals.

Considering each experiment was motivated by strong *a priori* hypotheses, in a second step, the search volume could then be restricted to the predicted regions by performing ROI analyses. This was performed in Chapters 3 and 4, for experiments

where source oscillatory power was the primary focus. Here, a more stringent family-wise error (FWE) corrected threshold of $p < 0.05$ was applied, with a spatial extent threshold of >10 voxels for each ROI.

2.7.1.8 Estimation of deep neural sources

Despite the very high temporal resolution enabling dynamic brain activity to be recorded, and the minimal signal attenuation or distortion by skull and scalp tissue, MEG is assumed to be insensitive to deep sources due to the negative effect depth of electrical activity has on the amplitude of magnetic fields. While it is inevitable that spatial resolution does decrease with increasing depth (Hillebrand and Barnes, 2002), MEG has been successfully used to measure activity in very deep sources, including the thalamus and brainstem (e.g. Parkkonen et al., 2009; Papadelis et al., 2012; Attal and Schwartz, 2013; Coffey et al., 2016). Source reconstruction of the hippocampus, in particular, has been a topic of debate (Shigeto et al., 2002; Riggs et al., 2009; Mills et al., 2012; Ruzich et al., 2019). However, evidence has accumulated to convincingly establish that MEG can indeed localise activity in the hippocampus (Meyer et al., 2017; Pu et al., 2018; Ruzich et al., 2019), including during scene construction (Barry et al., 2019a), memory encoding (Crespo-García et al., 2016), and autobiographical memory retrieval (McCormick et al., 2020). Separate intracranial EEG (iEEG), MEG and fMRI studies using the same virtual reality (VR) paradigm have also revealed similar hippocampal theta activity between modalities (Doeller et al., 2008; Kaplan et al., 2012; Bush et al., 2017). Particularly compelling are studies using concurrent MEG and iEEG, where the ground truth is available, which have demonstrated MEG can successfully detect hippocampal activity using beamforming (Crespo-García et al., 2016) and at the sensor level (Dalal et al., 2013). Evidently, the utility of MEG is not restricted to imaging shallow cortical structures, and

beamforming has proven to be a powerful technique to reconstruct sources regardless of depth. In fact, the ability to detect deep sources depends on many factors, including the interference of extraneous noise and the forward model applied.

In addition to the modelling constraints outlined in this chapter, further constraints were imposed when making inferences about the source distribution of the data. Previous fMRI and neuropsychological data provided important evidence for the recruitment of specific brain regions in the cognitive processes being studied (as noted in Chapter 1). This is particularly powerful when comparable paradigms are used in fMRI and MEG, with corroborative findings. These sources of evidence form the basis of strong *a priori* predictions about ROIs, and lend greater credence to focal localisation of an effect from MEG data.

2.7.1.9 Interpretation of spectral power changes

Oscillatory activity can change in response to a task relative to a baseline period (during which time it is assumed that there is no task-related activity), or relative to a second task to better control certain properties such that a particular cognitive effect can be better characterised. The resultant change in activity may be expressed as increases or decreases in frequency power (i.e. the amplitude). Many spatial navigation studies have reported power increases (e.g. Ekstrom et al., 2005; Bohbot et al., 2017; Bush et al., 2017; Kaplan et al., 2017; Aghajan et al., 2017), seemingly reflecting increases in synchronised activity in a particular frequency band. However, there is accumulating evidence of power decreases when engaging in tasks related to memory, which suggests that power attenuation is also a source of rich information, and an important index of neuronal activity during episodic memory processes (for reviews, see Hanslmayr et al., 2012 and

Herweg et al., 2020). Although the mechanistic role of power decreases is not entirely clear, there is sufficient evidence to suggest these decreases have an active role in memory encoding (Hanslmayr et al., 2012), the immediate construction of scenes (Barry et al., 2019a, 2019b), and autobiographical memory recall (McCormick et al., 2020) using MEG.

How does the electrophysiological signal of MEG relate to the haemodynamic signal of fMRI – do they reflect the same underlying neurophysiological processes? Both signals originate from post-synaptic (dendritic) currents, however, it is not clear how close the correspondence is between these neuroimaging techniques, especially considering the large differences in resolution (both spatial and temporal). Nevertheless, power decreases in frequencies below 30 Hz during memory encoding have been inversely associated with the fMRI BOLD signal (Fellner et al., 2016), and related to enhanced subsequent memory recall (see Herweg et al., 2020). Similarly, iEEG studies consistently find decreases in theta during episodic memory tasks (e.g. Fellner et al., 2019; Solomon et al., 2019).

To summarise, decreases in localised oscillatory power observed during scene-related processing would, therefore, likely be indicative of increased neural engagement in those brain regions. This effect relates to low frequencies (<30 Hz), which were the focus of this thesis. It remains unclear, however, why some studies have also reported increases. Consequently, at the outset of each study I had an open mind about whether changes in oscillatory power would be expressed as increases or decreases, and report all findings in subsequent chapters. This is also a topic of discussion I return to in Chapter 7.

2.7.2 *Sensor-space ERFs*

The neural dynamics underlying higher-level cognitive processes can be characterised not only by oscillatory activity, but also by time-locked evoked responses, which I consider next. For a detailed account of evoked responses, see Luck (2014).

In order to observe oscillatory theta activity, at least 5-7 cycles would be needed, requiring an experimental time-window of at least 1000 ms. As outlined in the previous section, these brain responses are measurable when averaging over numerous trials of the same task for a particular frequency band. This contains information about the amplitude of the signal, but not its phase (i.e. the precise timing of the maxima and minima of the oscillations), making it only time-locked. In contrast, neural responses that are both time-locked and phase-locked to the onset of a stimulus – also known as evoked responses – are transient effects, normally expressed as deflections of the signed (positive or negative) signal. These can be thought of as the power of the average across similar epochs (as opposed to the average power, as observed in induced, oscillatory responses) and are measurable across a wide frequency band. This time-domain analysis can be performed with or without source localisation (see Section 2.7.2.2).

Ideally, for a measurable ERF, a large number of epochs per participant are needed. This is quite straightforward for studies using simple stimuli that be repeated multiple times, such as auditory tones. However, this is considerably more difficult when studying short-lived effects that decay over consecutive trials; a good example of this would be the BE phenomenon (described in Chapter 1 and examined in Chapter 5). Some study designs also do not allow for large trial numbers due to constraints on the stimuli, and in order for scanning sessions to remain manageable in length for the participant.

Chapter 6 illustrates these latter challenges. Consequently, a smaller number of trials contributed to the ERF in several of my experiments, but despite this being disadvantageous to detecting an effect, clear and significant ERFs were nevertheless identified in both studies. Furthermore, to compensate for the smaller number of trials, data were collected from ≥ 15 participants, all of whom were highly attentive, engaged in the tasks, kept very still, and for whom eye movements were tightly controlled. This ensured a high SNR, optimal for ERFs.

ERF measurements are aptly suited to investigating cognitive effects that are dependent on very rapid processing – for example, implicit or covert processing not detectable by slow, behavioural measures, such as reaction times. This is particularly true for the two aforementioned experiments, both designed to examine transient changes in power within time intervals too brief (700 ms or less) to be expressed as oscillatory responses.

In this thesis, I sought to determine the timing and channel locations of statistically significant differences in evoked activity between pairs of conditions. This approach was used to (a) characterise the time scale of the likely rapid process of BE (an implicit form of scene construction; Chapter 5), and to (b) examine the fine temporal dynamics of events built from scenes relative to those built from non-scenes (Chapter 6).

2.7.2.1 *ERF analysis*

Following basic filtering step (see 2.6.1), before ERFs could be calculated, data were low-pass filtered using a two-pass Butterworth filter and a frequency cut-off that was specified according to the study's requirements; in each case, a wide frequency band was selected.

I acknowledge this filter causes a small degree of temporal blurring in either direction, although the effect is minimal (~ 10 ms). Data were then visually inspected and baseline-corrected, as described in earlier sections. To ensure the observed difference between conditions was not already present in the pre-stimulus signal, which would suggest the presence of a confound rather than a stimulus-related effect, a pre-stimulus period of -100 to 0 relative to stimulus onset was included in the plotted ERFs. The amplitudes of the group-level effects during stimuli were later shown to not be similar in magnitude to differences in this pre-stimulus interval.

The robust average was then calculated to obtain the ERF per participant per condition. This averaging method down-weights outliers when computing the average and helps to suppress high-frequency artefacts to minimise trial rejection, particularly helpful when only a small number of trials are available (Wager et al., 2005), as was the case for the ERF experiments in Chapters 5 and 6.

ERFs were analysed using the FieldTrip toolbox (Oostenveld et al., 2011) implemented in Matlab R2018a. To statistically compare the evoked activity between any two conditions, I used a non-parametric cluster-based permutation test, correcting for multiple comparisons across all MEG channels and time samples across the specified time window (Maris and Oostenveld, 2007). This test controls for the Type I error rate by identifying clusters of significant differences over time and sensors, rather than performing separate tests for each sample of time and space. I report only effects that survived this correction (FWE, $p < 0.05$). Cluster-level statistics are the sum of t-values within each cluster, and this was calculated by taking the maximum cluster-level statistic (positive and negative separately), over thousands of random permutations of the observed data. The number of permutations varied between 5000 and 10,000, depending

on the length of the time-window, since the more permutations the more computationally intensive. The obtained p-value represents the probability under the null hypothesis (no difference between a pair of conditions) of observing a maximum greater or smaller than the observed cluster-level statistics.

In summary, this approach allowed me to assess whether there were significant temporal clusters of differential activity between any two conditions. The statistical test can be visualised as a topographic distribution of the difference between conditions, providing a broad indication of the location of the effect (see Figure 21 for the distribution of MEG sensors as they correspond to broad brain regions).

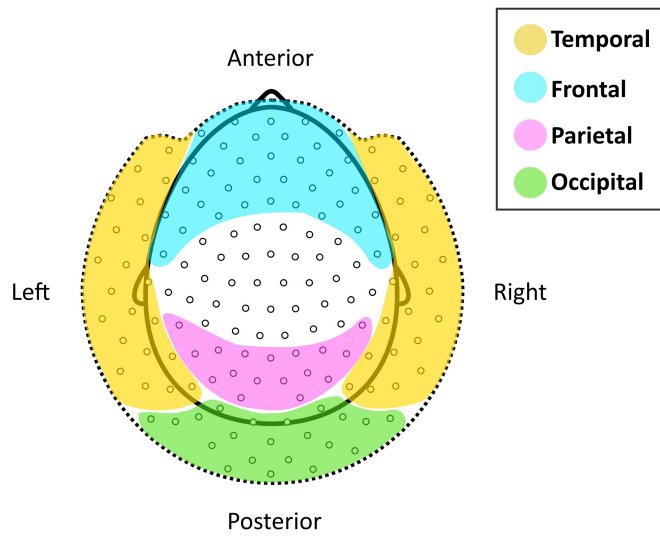


Figure 21. Distribution of MEG sensors over the head. The locations of sensors are displayed corresponding to temporal (yellow), frontal (blue), parietal (pink), and occipital (green) brain regions, visualised using the CTF151 helmet layout.

2.7.2.2 Source localisation of evoked responses

Having established a significant effect at the sensor level, a post-hoc analysis can be performed to reconstruct the sources of evoked activity, constrained using the temporal

information derived from the ERF. For example, if a significant difference is observed between conditions in a cluster between 100 and 200 ms, this is the time-window subsequently tested at the source level using beamforming. This can be performed in addition to the topographic visualisation of the statistical test already performed at the sensor level, and provides a more faithful representation of the distribution of the underlying neural activity. Rather than representing a second statistical test on the same data, this is simply performed to visually display the effect already determined at the sensor level. The resulting contrast images reveal where there were power changes in one condition relative to another (i.e. condition A minus B, and condition B minus A), and here were displayed at a liberal threshold ($p < 0.05$, uncorrected), chosen for illustrative purposes. There is some ambiguity regarding the appropriate thresholding to apply to the source reconstruction of ERFs once a significant difference is found at the sensor level. While it can be argued that a conservative statistical threshold may not be necessary, I also determined which of these active regions survived more stringent correction ($p < 0.005$, uncorrected), to obtain a sense of the spatial extent of activation comparable to more typical source reconstruction statistical thresholding levels (as applied in other experiments where the primary focus was on source reconstruction of oscillatory activity; Chapters 3 and 4).

Beamforming is commonly used to perform this source reconstruction. This is despite the fact that beamforming lies on the assumption that each source has its own unique time-course, not linearly correlated with another source (Hillebrand and Barnes, 2005). Partially correlated sources can be detected, but perfectly linearly correlated sources will likely result in little or no power being detected. The shorter time windows of ERFs (typically <100 ms) can also lead to low degrees of freedom and spurious

correlations. The challenge imposed by correlated sources means beamforming approaches are particularly tuned to reconstructing brain responses that occur over longer time intervals. However, transient evoked responses have been successfully localised with beamforming, despite the probable high similarity of the time-series of any two sources in evoked, time-locked activity (Cicmil et al., 2014; García-Pacios et al., 2015; Popov et al., 2018; Halder et al., 2019; Pscherer et al., 2020). This is likely because evoked activity is rarely perfectly correlated over every epoch – there will be some temporal jitter, permitting the computation of the covariance matrix in a time-domain beamformer like the LCMV method. Despite the success of using beamforming approaches to reconstruct sources of evoked responses, many researchers make the decision to not report source reconstruction.

It is worth remembering that when data are beamformed, the polarity of the ERF amplitude is lost, so that the resulting group images depict only a change in power. This makes any comparison between the polarity of the ERF difference, and the source reconstructed decreases or increases in power, problematic. That is to say, if the ERF shows a negative deflection for condition A, and a positive deflection for condition B, this does not necessarily correspond to a decrease in source power for A, and an increase for B. This also means that observing a negative or positive ERF does not mean a decrease or increase in brain activity.

2.7.3 Effective connectivity

Having established differences in neural engagement between experimental conditions, I then explored the effective connectivity between active sources. I used DCM (Friston et al., 2003; Kiebel et al., 2009) to infer how directional connectivity was modulated by tasks.

Using this approach, each *a priori* hypothesis of the causal influence of one region over another is embodied in a neural mass model based on particular data features (such as ERFs). Electrophysiological DCMs use neural mass models that summarise the activity within a specified cluster of neurons. I employed a convolution-based local field potential biophysical model, used for EEG/MEG data (Moran et al., 2009, 2013; Barry et al., 2019a). Intrinsic connectivity between different cell populations within a region are estimated. Extrinsic afferent inputs are categorised as forward, backward or lateral depending on which subpopulations these afferents project to (Felleman & Van Essen, 1991). Forward connections (“bottom up”) project to spiny stellate neurons in middle cortical layers, backward connections (“top down”) project to both excitatory pyramidal neurons and inhibitory interneurons, and lateral connections project to all cell populations. Each competing model is based on anatomical knowledge of the connections between these regions.

Once models were fitted to the MEG data, random-effects (RFX) Bayesian Model Selection (BMS) was then used to compare the evidence for these different biologically plausible models that varied according to which connections were modulated by the experimental condition (Stephan et al., 2009). The RFX procedure does not assume the optimal model is the same across all participants, making it more accurate and robust against outliers than fixed-effects approaches, since variability in brain activity across participants is to be expected when studying a complex cognitive process (Stephan et al., 2010). BMS therefore identifies the “winning” model out of these possible explanations of the data (Stephan et al., 2009).

Three simple models were specified: two unidirectional models (region A influences B; region B influences A) and a bidirectional model of mutual entrainment

(region A and B influences each other via reciprocal connections). Figure 22 provides an example of the application of these three models to the vmPFC and hippocampus. This allowed me to directly compare alternative unidirectional “master-slave” relationships, but also to directly compare each of these models with bidirectional mutual entrainment, where neither region exerted a stronger influence over the other. When applied to my data, I determined the winning model to be the one with the greatest exceedance probability (the maximum would be 100%), which provides a measure of how likely one model is, compared to all other models across the group of participants as a whole. Evidence for each model was a balance between the goodness-of-fit of the model with how parsimonious an explanation it provided, using the minimum number of parameters possible (also known as free energy).

In the analyses presented in Chapter 3 and 4, I used RFX BMS and plotted the exceedance probabilities for each model. In Chapter 3, where scene construction was examined step-by-step, separate models were specified for each stage of construction, allowing me to track changes in effective connectivity across time. In Chapter 4, where the type of object constructed at each of these stages was examined, the primary question concerned which connections were modulated by SD objects relative to SA objects. Therefore, a contrast between object conditions was performed. This was set as [1 -1], where the average of the two conditions is treated as the baseline, a set-up particularly suitable for paradigms where there is no low-level baseline condition.

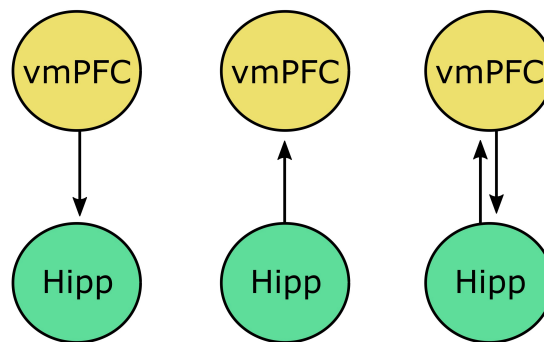


Figure 22. Example of dynamic causal models involving two regions. Three models are specified (left to right) in this example: the ventromedial prefrontal cortex (vmPFC) influences the hippocampus (Hipp), the hippocampus influences the vmPFC, and the two regions are mutually entrained.

A second analysis can be performed when seeking to compare only the two top models (those with the highest exceedance probability). To assess the consistency of the model fit of these two models, they can be compared by calculating the log Bayes factor (the ratio of the model evidence of model 1 to the model evidence of model 2). Log Bayes factor can be interpreted more easily by converting to a p-value, where a probability of 95% is the equivalent of a Bayes factor of 3. In Chapter 4, I calculated the log Bayes factor for the two unidirectional models, which in this case demonstrated the highest exceedance probability. Details on the particular DCM approach utilised are given in the specific experimental chapters.

Broadly speaking, two different kinds of DCM were applied to my data. Below, I briefly describe each in turn.

2.7.3.1 DCM for cross-spectral densities (CSD)

DCM for CSD (Kiebel et al., 2009; Moran et al., 2009) was used to assess how source localised pairs of brain regions dynamically interacted during the step-by-step mental construction of scenes and arrays studied in Chapters 3 and during scene construction involving different object types in Chapter 4. This was based on the spectral features of the data during each 3000 ms-long construction period. For this kind of DCM, to test competing biologically plausible hypotheses of the direction of information flow from one brain region to another, these relationships are modelled and fitted to the spectral components of the data. Furthermore, multiple models are directly compared in the same model space to establish which model best explains the observations. This revealed the different connectivity profiles of each experimental condition, contributing to the observed difference in spectral power. In each case, the three two-region models were specified.

2.7.3.2 DCM for event-related signals

DCM for event-related signals (Garrido et al., 2007) is a type of DCM applied to evoked responses measured with MEG, and was used in Chapter 5. For these data, effective connectivity was assessed between a number of brain regions identified from post-hoc source reconstruction. The time interval during which the BE effect was observed at the sensor level was the focus. A “NoBE” condition also existed, which was examined separately, thereby revealing different connectivity profiles for trials where BE occurred and where it did not. Due to the time interval being short (45.5 ms), I modelled the data across a broader interval (60 ms) within which the significant BE effect was detected, enabling better modelling of the early evoked response observed. For both conditions,

three simple models were again specified. The data fit between the event-related signals defined in the modelling was maximised, and BMS was again used to reveal the winning model.

2.8 MEG summary

In summary, by leveraging the fine temporal resolution of MEG, combined with its reasonable spatial resolution, this enabled me study the temporal dynamics underpinning construction cognitive processes, both related to scenes and non-scenes, that included imagining objects as they are integrated into a scene representation (Chapters 3 and 4), rapid implicit scene construction that serves to extend beyond the boundaries of the current scene (Chapter 5), and individual scenes as they unfold to give rise to an event (Chapter 6). Figure 23 provides an overview of the analysis pipeline.

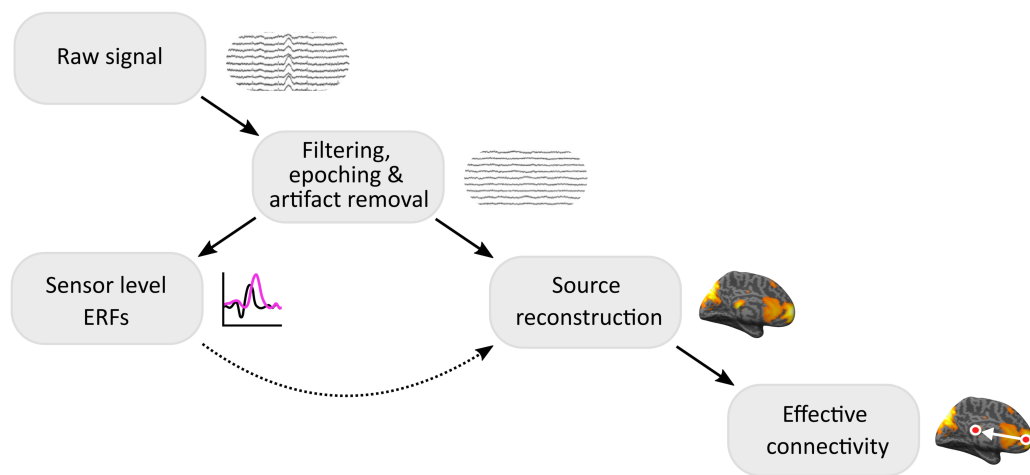


Figure 23. MEG data pre-processing and analysis pipeline. An overview of the steps taken to pre-process the raw data (filtering, epoching and artefact removal), before a source reconstruction or ERF analysis was performed. Following detection of differences at the source level, peaks of identified brain activity served as seed regions for analyses of effective connectivity. Source reconstruction of evoked responses is also subsequently performed.

3 Building scenes

3.1 Précis

The hippocampus and vmPFC are two brain regions that have been consistently implicated in supporting a range of cognitive functions including episodic/autobiographical memory recall and future thinking. As previously noted, it has been suggested that the ability to generate visual scene imagery may underlie these seemingly disparate processes. As a reminder, a scene is defined here as a naturalistic 3D spatially-coherent representation of the world typically populated by objects and viewed from an egocentric perspective. It is unclear whether scene processing in particular engages the hippocampus and vmPFC over and above other types of multi-component stimuli. In my first experiment, I leveraged the high temporal resolution of MEG to test participants as they gradually built scene imagery, and contrasted this with the closely matched construction of mental images of non-scene arrays. This paradigm permitted a closer examination of step-by-step mental construction than has been undertaken previously. This enabled me to better characterise the engagement of the hippocampus and vmPFC and their effective connectivity throughout construction, and determine whether their neural dynamics were particularly attuned to scene representations.

3.2 Introduction

The hippocampus plays a key role in episodic memory (Scoville and Milner, 1957), spatial navigation (O'Keefe and Dostrovsky, 1971), and a range of other cognitive domains (reviewed by Clark and Maguire, 2016; McCormick et al., 2018a). Perspectives differ on

precisely how the hippocampus supports these diverse cognitive functions. Nevertheless, there is general agreement that its contribution involves constructing representations comprised of numerous elements (Cohen and Eichenbaum, 1993; Lee et al., 2005; Hassabis and Maguire, 2007; Schacter and Addis, 2007; Yonelinas et al., 2019).

As discussed in Chapter 1, visual scenes are paradigmatic multi-element stimuli, perhaps explaining why scene stimuli have been used extensively to test hippocampal function. Patients with hippocampal damage show scene-related perceptual, imagination and mnemonic impairments (Lee et al., 2005; Hassabis et al., 2007b; Mullally et al., 2012b; Aly et al., 2013), and fMRI studies have pinpointed the anterior hippocampus in particular as being engaged during scene perception and imagination (Hassabis et al., 2007b; Graham et al., 2010; Zeidman et al., 2015a, 2015b; Hodgetts et al., 2016; Zeidman and Maguire, 2016). Lesions to the vmPFC also adversely affect scene imagination (Bertossi et al., 2016a). Moreover, a recent MEG study, where participants had 3000 ms to immediately imagine full scenes in response to single cue words (e.g. “jungle”), documented theta power decreases in the anterior hippocampus and vmPFC during scene imagination, with the latter driving activity in the hippocampus (Barry et al., 2019a).

Nevertheless, it remains unclear whether scenes preferentially engage the hippocampus and vmPFC. This is important to understand because it directly informs theories of how the hippocampus and vmPFC operate and, as a consequence, how they enable cognitive functions such as episodic memory. To reprise briefly, one account, the scene construction theory, posits that scenes are preferentially processed by the hippocampus (Hassabis and Maguire, 2007; Maguire and Mullally, 2013), with vmPFC driving hippocampal activity (McCormick et al., 2018a; Barry and Maguire, 2019a, 2019b; Ciaramelli et al., 2019). By contrast, the relational theory argues that it is the associating

of multiple elements that requires hippocampal input, irrespective of whether scenes are the outcome of such processing (Cohen and Eichenbaum, 1993; see also Yonelinas et al., 2019 for a related perspective).

To help adjudicate between these perspectives, Dalton et al. (2018) devised an fMRI paradigm which involved gradually building mental representations over three construction stages. This approach was based on previous work that reported three objects and a 3D space seem to be sufficient to generate the mental experience of a scene (Summerfield et al., 2010; Mullally and Maguire, 2013). Hence, in one condition Dalton et al. (2018) had participants build scene imagery from three successive auditorily-presented object descriptions imagined with a 3D space. This was directly compared to constructing images of non-scene arrays, composed of three objects and a 2D space. The scene and array stimuli were, therefore, highly matched in terms of content, and the associative and mental constructive processes they evoked. Focussing on the MTL and averaging across the full imagery construction period, Dalton et al. (2018) found that the anterior medial hippocampus was engaged preferentially during the mental construction of scenes compared to arrays during fMRI. This hippocampal involvement was not observed, however, when participants imagined 3D and 2D spaces on their own (i.e. with no objects; see Zeidman et al., 2012 for a similar result). Consequently, it would seem that the anterior hippocampus is particularly sensitive to representations that combine objects with a 3D space.

In the current study, I sought to extend previous work to offer novel insights into the construction of mental representations. In the Barry et al. (2019a) MEG study, participants imagined a full scene within 3000 ms, and this was compared to a low-level baseline (imagining single isolated objects). In the Dalton et al. (2018) fMRI study, scenes

were compared to non-scene arrays, but given the slow nature of the haemodynamic response, it was not feasible to examine the three construction stages separately. By contrast, here I combined these highly matched scene and array construction tasks with the high temporal resolution of MEG to study the neural dynamics associated with each of the three construction stages. I could therefore provide time-resolved insights into the step-by-step process of scene and non-scene array construction and the specificity of neural responses (if any) to scenes.

I focused on two regions of interest, the anterior hippocampus and the vmPFC, given the previous neuropsychological and neuroimaging evidence of their particular importance for scene construction. Moreover, my main interest was in theta, since there is a long history linking theta with hippocampal function, particularly from rodent studies (for a review, see Colgin, 2016; Karakaş, 2020). The focus on theta reflects this context, but also the previous MEG study associating theta with immediate full scene construction (Barry et al., 2019a). I predicted that theta power would be decreased relative to baseline, consistent with previous findings of attenuated power associated with the immediate construction of full scenes (Barry et al., 2019a, 2019b) and autobiographical memory recall (McCormick et al., 2020) using MEG. Moreover, as noted in Chapter 2, power decreases in frequencies below 30 Hz during memory encoding have been inversely associated with the fMRI BOLD signal (Fellner et al., 2016), and related to enhanced subsequent memory recall (for reviews see Hanslmayr et al., 2012 and Herweg et al., 2020). Similarly, iEEG studies consistently find decreases in theta during episodic memory (e.g. Fellner et al., 2019; Solomon et al., 2019; but see Ekstrom et al., 2005; Bush et al., 2017; Kaplan et al., 2017). Larger decreases observed during scene construction

would, therefore, likely be indicative of increased neural engagement of the hippocampus and vmPFC.

I also had hypotheses about how oscillatory power might change across the three construction stages. This prediction took account of the previous findings of anterior hippocampal and vmPFC activity during immediate scene construction (where participants had 3000 ms to immediately imagine a fully realised scene; Barry et al., 2019a), and also in the very earliest stage of autobiographical memory retrieval, which also involves scene construction (McCormick et al., 2020). Concordant with these previous findings, I predicted that anterior hippocampus and vmPFC power changes would be strongest during the first and possibly second stage of scene, but not array, construction.

Different views exist about the nature of interactions between the hippocampus and cortical areas such as the vmPFC in supporting memory and, I would add, building scene representations. For instance, some accounts place the hippocampus at the heart of such processing and believe it recruits neocortical regions in the service of this endeavour (Teyler and DiScenna, 1986; Teyler and Rudy, 2007). If this is the case, then the hippocampus should drive activity in vmPFC during the earliest stage of scene construction, and perhaps also during the subsequent stages.

This perspective stands in contrast to the model of scene construction that I articulated previously in Section 1.4 (McCormick et al., 2018a; Ciaramelli et al., 2019). Within this architecture, vmPFC initiates the activation of schematic and other knowledge in neocortex that is relevant for a specific scene, while inhibiting elements that are not relevant. This information is conveyed to the hippocampus, which starts to construct the scene. The vmPFC then engages in iterations via feedback loops with neocortex and

hippocampus to continually update the scene. I predicted that during the first stage of scene construction the vmPFC would drive the hippocampus. For the subsequent scene construction stages, I hypothesised there would be mutual entrainment between the two areas, reflecting the feedback loops between them.

My hypotheses aligned with the scene construction theory. However, given the similarity of the scene and array conditions, my paradigm was also capable of revealing if scenes and arrays were treated similarly by the hippocampus, which I believe would be the prediction of relational theorists, given their position that any form of associative processing should recruit this region.

3.3 Materials and methods

3.3.1 Participants

Twenty healthy, right-handed participants (13 female; mean age = 25.50 years; Std Dev = 4.70) took part in the experiment.

3.3.2 Stimuli

During MEG scanning, participants performed two closely matched tasks adapted from Dalton et al. (2018). Both tasks involved participants mentally constructing images with their eyes open while looking at a blank, grey screen (Figure 24).

For the scene task, participants were asked to first imagine a 3D grid covering approximately the bottom two-thirds of the blank screen. They then heard three object descriptions, presented one at a time, and they imagined each object on the 3D grid, each

in a separate location that was specified in advance by the trial's visual cue (Figure 24). The participants were explicitly instructed to link the three objects together and with the 3D space, as they formed their mental representations. Specifically, after the first object description, participants were told to imagine each subsequent object whilst continuing to maintain the image of the grid and previous object(s), without rearranging objects from their original cue-defined positions. By the final object description, the entire mental image created by participants was a simple scene composed of a 3D grid and three objects.

For the array task, participants were asked to first imagine a regular, flat 2D grid covering approximately the bottom two-thirds of the screen. They then heard three object descriptions, presented one at a time, and they imagined each object on the 2D grid, each in a separate location that was specified in advance by the trial's visual cue (Figure 24). The participants were explicitly instructed to link the three objects together and with the 2D space as they formed their mental representations. Specifically, after the first object description, participants were instructed to imagine each subsequent object whilst continuing to maintain the image of the grid and previous object(s), without rearranging objects from their original cue-defined positions. By the final object description, the entire mental image created by participants was a simple array composed of three objects on a flat 2D grid. In this array task it was emphasised to participants they should link the objects and the space together, but not in a way that would create a scene.

For both scene and array tasks, a visual cue determined the three specific locations where participants should imagine each object during a trial. The identity of the objects was not provided at this point, only where they ought to be placed when being imagined while looking at the blank screen. Object descriptions were auditorily presented to the participant during the subsequent mental construction stages (Figure 24). There were four

different cue configurations (Figure 25A) which were randomised across the experiment, and the frequency of each was matched across scene and array conditions. I emphasised to participants the importance of following the cue configurations as precisely as possible. The purpose of prescribing the object configuration at the start of a trial was to ensure matched oculomotor behaviour between the scene and array conditions, consistency across participants, and that objects were imagined as separate and non-overlapping, occupying the full extent of the grid. In MEG, more so than fMRI, it is crucial to control for eye movements when making between-task comparisons, as discussed in **Ocular artefacts** in Section 2.5.2.2, to avoid erroneously attributing movement of ocular dipoles to changes in brain activity. The final design of these four visual cues was informed by participant feedback during a behavioural pilot study I conducted prior to the main MEG experiment. By including a small number of distinct and easy to follow cue configurations, I believe that a balance was achieved between not over-complicating the task by introducing additional working memory load, and maintaining attention by having a small degree of unpredictability trial-to-trial in the object configuration to be followed.

Participants were asked to construct novel mental images of objects based on the descriptions alone, and not to draw on memories associated with the objects described, or to recall specific objects of their acquaintance. In addition, they were told to imagine only the objects described and not add other elements to a scene or array. All imagery was to remain static and they were required to maintain a fixed viewpoint, as though looking at the image in front of them, rather than imagine moving parts of objects or imagine themselves moving through space.

The object descriptions were the same as those used in the original fMRI study (Dalton et al., 2018), and each description was heard only once. The objects were rated

by a separate group of participants on a range of properties by Dalton et al. (2018), which enabled me to ensure that the scene and array conditions were closely matched in terms of the number and order of presentation of space-defining and space-ambiguous objects² ($t_{(71)} = 0.497, p = 0.620$) (Mullally and Maguire, 2011, 2013), ratings of object permanence ($t_{(71)} = 0.178, p = 0.859$) (Mullally and Maguire, 2011), as well as the number of syllables in the objects descriptions ($t_{(71)} = 0.327, p = 0.745$) and utterance duration ($t_{(71)} = 0.169, p = 0.866$). Furthermore, all objects were rated as highly imageable, obtaining a score of at least 4 on a scale from 1 (not imageable) to 5 (extremely imageable). Objects in each triplet were not contextually related to each other – as part of the piloting in Dalton et al.'s (2018) fMRI study, if a triplet was deemed consistently across participants to contain objects that were related, this triplet was not included in the main experiment. Furthermore, in my study, object triplets were counterbalanced across conditions such that for some participants a triplet was in the scene condition and for others it was in the array condition. In this way, I controlled for semantic elements across conditions. Importantly, and to reiterate, during both tasks the participants viewed a blank screen, so there was no difference in visual input between conditions.

A third task was also included in my study. Participants heard a backward sequence of three numbers, after which they were instructed to mentally continue counting backwards. The sole function of this task was to provide participants with relief from the effortful imagination required of the scene and array trials. This counting task was not subject to analysis.

² Of note, in Chapter 4 I report the results of a separate experiment, in a new group of participants, that specifically examined the space-defining/ambiguous object property (as well as permanence), using these same stimuli (see section 4.3.2).

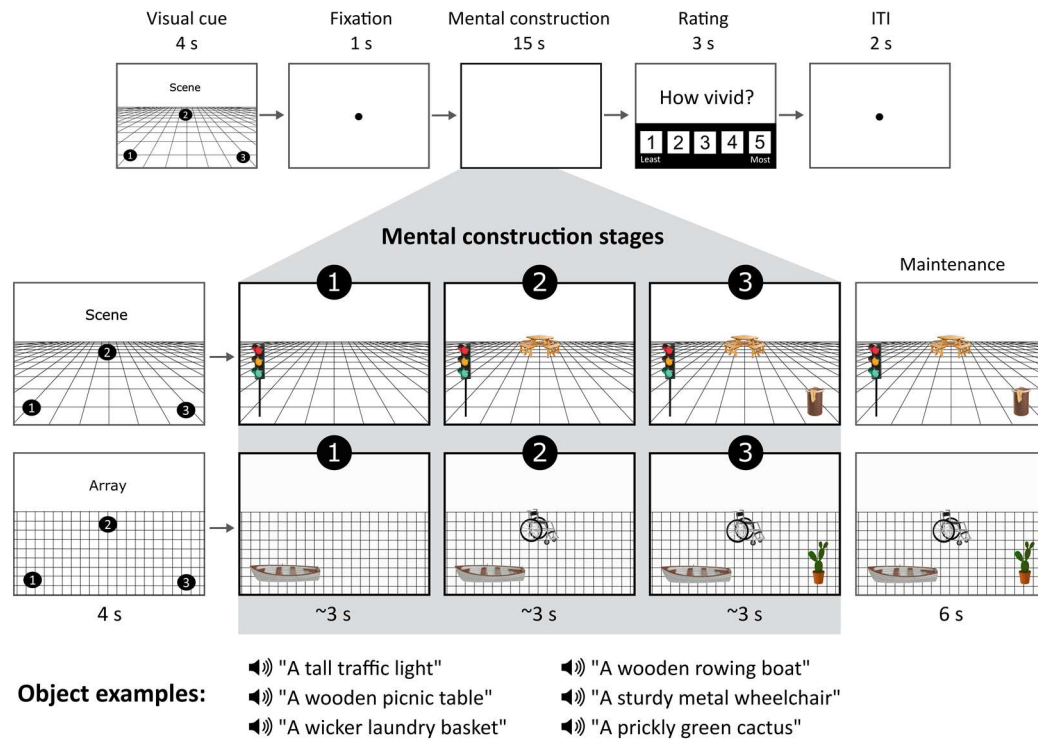


Figure 24. Structure of a trial. The top panel shows the timeline of, in this case, a scene trial. In the middle panels, an example cue configuration is shown on the left, and what participants imagined as they mentally constructed a scene or an array is shown on the right. Of note, during the 3 s construction stages and 6 s maintenance period, participants were looking at a blank screen, and images here merely serve as illustrations for the reader. Examples of the auditorily presented object descriptions during each stage are provided in the bottom panel.

To summarise, scene and array tasks were identical except for one manipulation – imagining objects on a 3D grid induced a sense of depth and perspective, resulting in a mental representation of a scene, in contrast to imagining objects on a 2D grid which resulted in a mental representation of a non-scene array. This close matching of the scene with the array tasks meant any differences in neural activity could not simply be attributed to more general mental construction processes. Moreover, the use of MEG meant that I could now examine the step-by-step construction of these representations with millisecond precision.

3.3.3 Pre-scan training

Before entering the scanner, participants were trained to become familiar with the task requirements. They completed four practice trials per task on a desktop computer in a darkened room, to liken the experience to that in the scanner environment. After each trial they rated the vividness of the mental imagery on a scale of 1 (not vivid at all) to 5 (extremely vivid). If they gave a rating of 3 or lower on any trial, the instructions and practice items were repeated. A score of 4 or more qualified them to proceed to the MEG experiment. After one of the practice trials per task, participants drew what they were able to imagine (Figure 25B). The main purpose of the drawings was to ascertain whether a participant drew objects corresponding to the auditory descriptions, in the correct cue-specified locations, and on a 3D or 2D grid, indicating whether objects were correctly imagined in a 3D or 2D space. I checked that all three criteria were met after each drawing.

3.3.4 Task procedure

On each trial (Figure 24), a visual cue was presented for 4000 ms that contained both the identity of the task (scene or array) and the configuration of locations at which the objects should be imagined on the grid following each auditory object description. Following the cue, participants fixated on the centre of the screen for 1000 ms before the start of the imagination period (~15 seconds), during which they engaged in mental imagery whilst hearing three auditory object descriptions one after another. Participants were instructed to move their eyes to where they were imagining each object on the screen. Each object description lasted approximately 2000 ms, followed by a silent 1000 ms gap before the presentation of the next object description, so that each object imagination period was 3000 ms in duration. These three construction stages were my main interest, and the focus

of the data analyses, since my hypotheses concerned the construction process. After the final object description, there was an additional 6000 ms silent period where participants maintained the image of the fully realised scene or array they had created, which was included in order to avoid an abrupt transition between trials (Figure 24). This maintenance period was examined as a secondary analysis. Participants then rated how vividly they were able to imagine the scene or array on that trial on a scale of 1 (not vivid at all) to 5 (extremely vivid), self-paced with a maximum duration of 3000 ms. An inter-trial interval (ITI) of 2000 ms preceded the next trial.

A total of 24 trials per task were presented in a pseudorandomised order across four separate blocks – two blocks of twenty and two blocks of nineteen trials. This ensured that no more than two trials belonging to one condition were presented consecutively. To ensure participants attended to the tasks throughout scanning blocks, I included an additional six catch trials (two per task, including the low-level counting backwards task) pseudorandomly presented across all four scanning blocks such that every block contained at least one catch trial. During a catch trial, participants were required to press a button when they heard a repeated object description or number within a triplet.

3.3.5 In-scanner eye tracking

An Eyelink 1000 Plus (SR Research) eye tracking system with a sampling rate of 2000 Hz was used during MEG scanning to monitor task compliance. For some participants the calibration was insufficiently accurate, leaving 15 data sets in the eye tracking analyses.

3.3.6 Post-scan surprise memory test

Following the experiment, participants completed a surprise item memory test for the scene and array tasks, outside of the scanner on a desktop computer. Participants were presented with all individual auditory object descriptions heard during the scanning experiment (72 scene and 72 array objects – 24 trials of each condition comprised of three objects each) and an additional 72 lure object descriptions not previously heard. Auditory descriptions were presented one at a time, and the order of object descriptions was randomised. After hearing each description, participants pressed a button to indicate “yes” if they remembered hearing that object description during the scanning experiment or “no” if they did not. They then rated how confident they felt about their answer on a scale from 1 (low confidence) to 5 (high confidence). Both responses were self-paced, with a maximum of 5000 ms for each.

I chose not to include an associative memory test. This decision was informed by the previous experience of Dalton et al. (2018) using the same task during fMRI. They reported that in a post-scan surprise associative memory test there was no difference between the scene and array conditions, but participants performed at chance. As part of a separate behavioural pilot study, I asked participants to perform a source memory test, to see whether participants would perform at a similar (chance) level to the associative memory test, or whether it would be a useful addition to the surprise memory testing post-scan. In this task, participants indicated whether each object description was originally encountered as part of a scene or an array. Again, participants performed at chance, with no difference between scenes and arrays, and consequently this memory test was not included in the main study. These previous and pilot findings are not surprising. Tests of associative and source memory are challenging in the context of this paradigm,

because participants were not explicitly told in advance that memory would be subsequently tested. The scene and array tasks were instead designed to be “in the moment” imagination tasks.

3.3.7 Post-scan debriefing

Upon completion of the memory test, participants were asked about their experience of performing the scene and array tasks. Ratings completed during this debriefing are reported in the Results section (also see Table 1).

The visit of each participant took approximately 2.5 hours in total, including training, set up in the MEG scanner and eye tracker calibration, the experimental task (~40 minutes), and the post-scan session.

3.3.8 Behavioural data analysis

Ratings for scene and array per-trial vividness during scanning were compared using paired-samples t-tests, as were the ratings from the post-scan debriefing session. Eye tracking comparisons between scene and array tasks were performed using paired-samples t-tests to examine eye movements and fixations across the full 9000 ms construction period. Where relevant, two-way repeated measures ANOVAs were also used to assess the effects of task (scene, array) and construction stage (first, second, third) in the eye tracking data. For the post-scan surprise memory test, to determine whether memory for object descriptions was above chance for scenes and arrays, one-sample t-tests were performed. The memory test data were also analysed using two-way repeated measures ANOVAs which assessed the effects of task (scene, array) and construction

stage (first, second, third). Statistical analyses were performed using SPSS25 (<https://www.ibm.com/products/spss-statistics>), using a significance threshold of $p < 0.05$.

3.3.9 MEG data acquisition

MEG data were acquired using the whole-head 275 channel CTF Omega MEG system with a sampling rate of 1200 Hz, as describe in Chapter 2.

3.3.10 MEG data pre-processing

MEG recordings were filtered with a 1 Hz high-pass filter, 48-52 Hz stop-band filter, and 98-102 Hz stop-band filter, to remove slow drifts in signals from the MEG sensors and power line interference. Epochs corresponding to the three scene construction stages and three array construction stages were each defined as 0 to 3000 ms relative to the onset of each object description, while epochs corresponding to the scene and array maintenance periods were defined as 0 to 6000 ms relative to the offset of the third construction stage. Scene and array baseline periods were defined as 1000 to 2000 ms relative to the onset of the ITI preceding a scene or array trial. This final 1000 ms of the ITI was chosen as it provided the most separation from task-related activity – the on-screen visual cue was not an appropriate baseline since this contained important preparatory information pertinent to the upcoming scene or array to be constructed. Epochs were concatenated across trials for each condition and across scanning blocks.

3.3.11 MEG data analysis

All MEG analyses were conducted using SPM12 (www.fil.ion.ucl.ac.uk/spm).

3.3.11.1 MEG source reconstruction

Source reconstruction was performed using the DAiSS toolbox (<https://github.com/SPM/DAiSS>) implemented in SPM12 (www.fil.ion.ucl.ac.uk/spm) and results were visualised in MRICroGL (<https://www.mccauslandcenter.sc.edu/mricrogl>) using the MNI 152 T1 image. The LCMV beamformer algorithm (Van Veen et al., 1997) was used to generate maps of power differences between conditions at each construction stage. For each participant, a set of filter weights was built based on data from all six construction stages (three for scenes and three for arrays) within the theta band (4-8 Hz) and a 0 to 3000 ms time window relative to object description onset. Coregistration to MNI space was performed using a 5 mm volumetric grid and was based on nasion, left and right preauricular fiducials. The forward model was computed using a single-shell head model (Nolte, 2003).

At the first level, whole brain theta power was estimated within each construction stage for each condition, creating one weight-normalised image per construction stage per participant. The images were smoothed using a 12 mm Gaussian kernel. To determine the peak locations of power differences between the scene and array tasks in the anterior hippocampus and vmPFC, at the second level I performed a t-contrast at each of the three construction stages (for example, scene construction stage 1 relative to array construction stage 1). An uncorrected whole brain threshold of $p < 0.001$ with a cluster extent threshold of >100 voxels was applied to each contrast, in line with numerous other MEG studies (e.g. Lieberman and Cunningham, 2009; Hanslmayr et al., 2011; Kveraga et al., 2011; Kaplan et al., 2012; Cushing et al., 2018; Barry et al., 2019a). This is held to provide a balance between protecting against false positives whilst enabling the detection of subtler signals. This was followed by ROI analyses where I focused on my two regions

of interest, the anterior hippocampus and the vmPFC using separate bilateral masks covering each of these two regions, again conducted in SPM12. I included a separate bilateral posterior hippocampal mask for completeness. A FWE corrected threshold of $p < 0.05$ was applied, with a spatial extent threshold of 10 voxels for each ROI. Peaks in the hippocampus and vmPFC provided the seed regions for subsequent effective connectivity analyses. The beamformer analysis was also repeated in the alpha (9-12 Hz) and gamma (31-100 Hz) bands to examine whether effects of task at each stage were confined to the theta range. For completeness, and while not of primary interest, a similar beamformer analysis was performed for the maintenance period (0 to 6000 ms) that followed the third construction stage of scenes and arrays.

To investigate whether engagement of the anterior hippocampus and vmPFC changed across the three construction stages during each condition, I used the same source extracted theta power obtained with the beamformer for an additional analysis. Using the previously defined masks (anterior hippocampus, vmPFC and posterior hippocampus), I extracted ROI-specific power values from each of the smoothed first level images. Power was subsequently averaged across each ROI, to produce an average for each participant at each of the six construction stages. The same source reconstruction protocol was performed for a pre-trial fixation baseline time-window, which was then subtracted from each task stage to ascertain whether theta power originating from each ROI during task engagement corresponded to an increase or attenuation of power. Data were then entered into two-way repeated measures ANOVAs implemented in SPSS25, allowing me to evaluate changes in power between scene and array tasks and across construction stages (task \times stage). Where the assumption of sphericity was violated using the Mauchly test, Greenhouse-Geisser corrected degrees of freedom were reported.

Paired-sample t-tests were performed for each ROI in cases of a significant main effect or interaction, using a conservative significance level of $p < 0.01$.

3.3.11.2 *Effective connectivity*

To assess how the hippocampus and vmPFC interacted during each of the three construction stages for both scenes and arrays, I used DCM for cross spectral densities (Moran et al., 2009). This is described in some detail in Section 2.7.3.1, but to reprise briefly, this approach permits the comparison of biologically plausible models representing my *a priori* hypotheses of the directed influence one region has over another, as well as mutual entrainment between regions (Friston, 2009; Kahan and Foltynie, 2013). A convolution-based local field potential biophysical model used for EEG/MEG data (Moran et al., 2009, 2013; Barry et al., 2019a) was applied. Intrinsic connectivity between different cell populations within a region are estimated as part of every model. Extrinsic afferent inputs are categorised as forward, backward or lateral depending on which subpopulations these afferents project to (Felleman and Van Essen, 1991). Bayesian model selection (BMS) was then used to compare the evidence for different competing models that varied according to which connections were modulated by the experimental condition (Stephan et al., 2009).

For each construction stage, I tested three models corresponding to three competing hypotheses regarding hippocampal-vmPFC connectivity: (1) hippocampus drives vmPFC, (2) vmPFC drives hippocampus, and (3) hippocampus and vmPFC are mutually entrained. When applied to my data, the exceedance probability (EP) provided a measure of how likely each model was, compared to all other models across the group

of participants as a whole. The model with the highest exceedance probability represented the winning model.

3.4 Results

The goal of this study was to examine whether the anterior hippocampus and vmPFC supported step-by-step scene construction whilst using the array condition to control for neural responses associated with objects, and general associative and constructive processing. I first examined the behavioural data to assess whether the participants adhered to the task instructions and to ascertain if there were any differences between the scene and array tasks that might have influenced the neural activity.

3.4.1 Behavioural results

3.4.1.1 *Did participants engage in the scene and array tasks and adhere to instructions?*

Several methods were used to assess whether participants paid attention during scene and array trials, engaged in mental imagery, and complied with task-specific instructions throughout the experiment.

Pre-scan training

After performing scene and array practice trials during the pre-scan training, participants drew what they imagined on each trial, to ensure they could closely follow the task instructions and cue configurations. Scene and array drawings demonstrated participants formed the appropriate mental representations, related the objects and the space together in a coherent way, and were able to mentally place the objects in the locations indicated by the cue configurations (see examples in Figure 25B).

Catch trials

Catch trials were included in each scanning block whereby participants were instructed to press a button whenever they heard a repeated object description (or number) within a trial. On average 98.33% (Std Dev = 0.31) of catch trials were correctly identified, suggesting participants remained attentive throughout the experiment.

In-scanner eye tracking

Heat maps of the spatial pattern of fixations for the cue configurations confirmed not only remarkably close adherence to the instructions trial-by-trial, but also consistency across participants (Figure 25A). The eye tracking data were also examined for any differences between the scene and array tasks over the full 9000ms construction period, of which none were evident – fixation count ($t_{(14)} = 0.239, p = 0.814$), fixation duration ($t_{(14)} = 0.559, p = 0.585$), saccade count ($t_{(14)} = 0.322, p = 0.752$), saccade amplitude ($t_{(14)} = 0.752, p = 0.465$). I also examined the two fixation variables for any effects of task (scene, array) and construction stage (first, second, third). For both fixation count and fixation duration, there was no significant effect of task (fixation count: $F_{(1, 14)} = 1.441, p = 0.250$; fixation duration: $F_{(1, 14)} = 2.964, p = 0.107$) or construction stage (fixation count: $F_{(2, 28)} = 0.372, p = 0.693$; fixation duration: $F_{(2, 28)} = 0.650, p = 0.530$), and there was no significant task \times construction stage interaction (fixation count: $F_{(2, 28)} = 2.394, p = 0.110$; fixation duration: $F_{(2, 28)} = 0.356, p = 0.704$). Oculomotor behaviour was therefore closely matched across all construction stages, and there was no significant difference between the scene and array tasks.

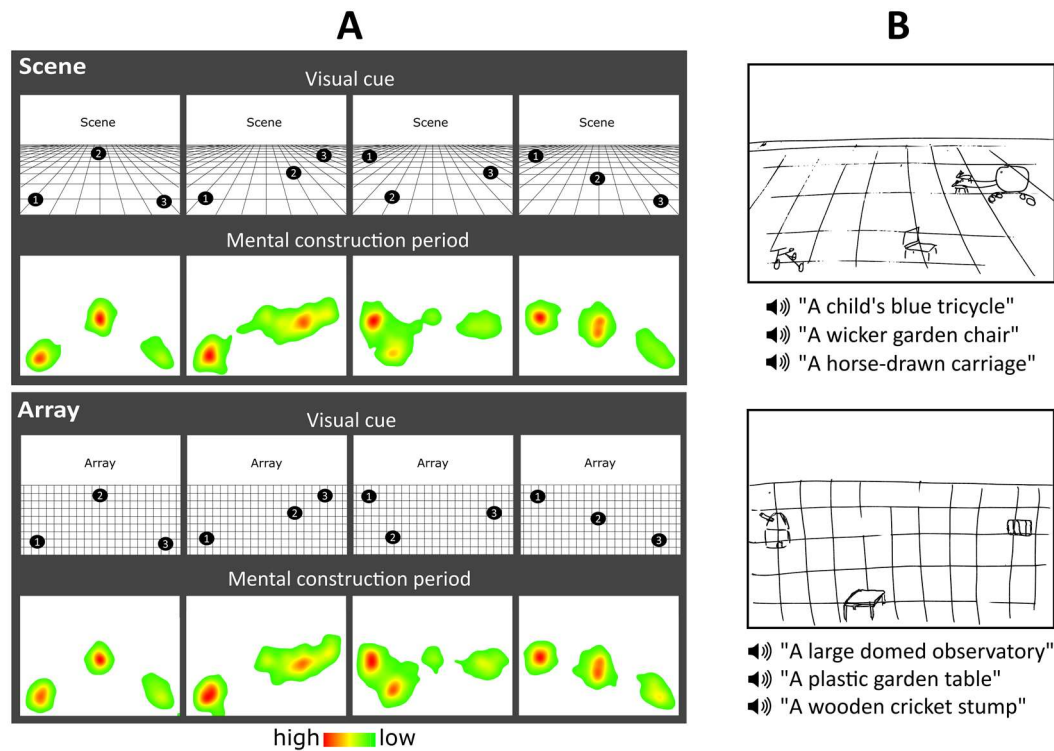


Figure 25. Eye fixations and example drawings. (A) Heat maps for each cue configuration for the scene and array tasks show the fixation count during the 9000 ms mental construction period. Each heat map is an aggregate of fixations on the blank screen across all trials for that cue configuration across all participants with eye tracking data ($n=15$). Red indicates higher fixation density and green lower fixation density. (B) Pre-scan training drawings for the scene (top) and array (bottom) tasks from an example participant.

3.4.1.2 Did vividness of mental imagery differ between scene and array tasks?

Participants rated the vividness of mental imagery immediately after each scene and array trial in the scanner on a scale of 1 (not vivid at all) to 5 (extremely vivid). On average, both scene (mean = 4.15, Std Dev = 0.43) and array (mean = 4.08, Std Dev = 0.49) trials were rated high on vividness, and there was no significant difference between the tasks ($t_{(19)} = 0.873, p = 0.394$).

3.4.1.3 *Were there any other subjective differences in mental imagery between the scene and array tasks?*

After the experiment was completed, participants reported on their subjective experience of performing scene and array tasks in a debriefing session (questions and summary data are shown in Table 1). Mind-wandering was reported as very low for both scene and array tasks, and did not differ significantly between the two conditions ($t_{(19)} = 1.831, p = 0.083$). Mental imagery was performed with ease in both conditions ($t_{(19)} = 0.000, p = 1.000$). Participants also had no difficulty following the cue configuration instructions for either condition ($t_{(19)} = 1.000, p = 0.330$). As expected, scene trials were perceived as significantly more 3D than array trials, while arrays were rated as more 2D ($t_{(19)} = 17.168, p < 0.0001$). Participants also reported that they rarely needed to suppress scene-like mental imagery during array trials.

3.4.1.4 *Did memory performance differ between scene and array tasks?*

After scanning, participants engaged in a surprise item memory test. They performed above chance at recognising individual object stimuli from scene trials (mean = 84.31%, Std Dev = 9.63; $t_{(19)} = 39.135, p < 0.0001$), array trials (mean = 85.42%, Std Dev = 9.43; $t_{(19)} = 40.482, p < 0.0001$) and identifying novel items (mean = 87.54%, Std Dev = 7.48; $t_{(19)} = 52.289, p < 0.0001$). Participants, therefore, paid sufficient attention during both tasks to successfully encode a large number of stimuli, even though they were never instructed to memorise. A repeated-measures ANOVA revealed no significant effect of task ($F_{(1, 19)} = 0.456, p = 0.508$), construction stage ($F_{(1.487, 28.262)} = 1.480, p = 0.240$), and no task \times construction stage interaction ($F_{(1.554, 29.526)} = 0.739, p = 0.484$).

Table 1. Debriefing ratings.

Question	Rating (1-5 Likert scale)	Task	Mean \pm Std Dev
To what degree did you feel your mind was wandering, that you were thinking of things other than the task?	1 = no mind wandering... 5 = always mind wandering	Scene	1.20 \pm 0.41
		Array	1.35 \pm 0.49
How difficult was it to imagine the objects on the grid?	1 = easy... 5 = very difficult	Scene	1.75 \pm 0.64
		Array	1.75 \pm 0.72
How difficult was it to follow the cue's instruction of where to imagine the objects on the grid?	1 = easy... 5 = very difficult	Scene	1.2 \pm 0.41
		Array	1.15 \pm 0.37
Did you have any sense of 3D space during array trials? Please rate the extent to which you felt the image had a sense of depth.	1 = completely 2D... 5 = completely 3D	Scene	4.3 \pm 0.44
		Array	1.55 \pm 0.51
For the array trials, did you feel that you imagined the objects in a scene at all and tried to repress this?	1 = always repressed... 5 = never had to repress	Array	4.55 \pm 0.60

Std Dev = Standard Deviation.

3.4.1.5 Behavioural results summary

Having determined across a range of behavioural measures that the participants adhered to the task instructions, formed the appropriate mental images, with no differences in terms of key variables such as oculomotor behaviour, vividness and memory performance between the scene and array tasks, I next examined the neural data.

3.4.2 MEG results

I began by conducting a whole brain analysis, followed by ROI analyses focused on the anterior hippocampus and vmPFC, as my primary question was whether these two regions would be preferentially engaged during the step-by-step construction of scenes.

3.4.2.1 Source space power differences for each construction stage separately

I first compared the scene and array tasks during each of the three construction stages. Whole brain analyses revealed attenuated theta power for scenes relative to arrays at each stage in a number of areas, including the right vmPFC during the first construction stage (peak in BA11; $x, y, z = 20, 30, -20$; see Table 2 for all brain areas that were engaged). No differences were observed for the arrays relative to scenes contrast.

Given my specific anatomical hypotheses, analyses were then restricted to the anterior hippocampus and vmPFC ROIs using volumetric masks. For the first stage, I observed a significant attenuation of theta power for the scene relative to the array task in the left anterior hippocampus (peak = $-24, -20, -16$; Figure 26A), left vmPFC (peak in BA32; $-12, 40, -14$; Figure 26A) and right vmPFC (peak in BA11; $16, 26, -18$). The contrast of scene and array conditions during the second construction stage revealed an attenuation of theta activity for scenes in the vmPFC, extending bilaterally with an overall peak ($-4, 26, -4$) and subpeak ($-12, 32, -18$) in the left hemisphere (Figure 26A). No significant differences in anterior hippocampus or vmPFC were apparent between the scene and array tasks during the third construction stage, or during the ensuing maintenance period. For completeness, no task-related theta power changes were found for any of the construction periods, or the maintenance period, in the posterior hippocampus.

To investigate whether there were differences between scene and array tasks within the ROIs in frequencies other than theta, I performed the same ROI analyses for alpha (9-12 Hz) and gamma (30-100 Hz) bands. No significant changes were evident in the anterior (or posterior) hippocampus or vmPFC.

Table 2. MEG whole brain power changes during scenes relative to arrays at each construction stage.

Region	MNI peak coordinate (x, y, z)	Peak level <i>t</i> -value	Cluster size
Stage 1			
R thalamus	10, -10, 16	5.4	456
R inferior frontal gyrus	48, 40, 8	5.33	890
L ventral frontal cortex (BA47)	-32, 40, -14	4.67	814
L cerebellum	18, -36, -40	4.66	888
R vmPFC (BA11)	20, 30, -20	4.32	588
R putamen	32, 6, -4	4.03	245
Stage 2			
R thalamus	4, -26, 10	5.92	2552
R inferior frontal gyrus	60, 24, 6	4.94	965
L medial frontal cortex	-2, 22, 0	5.11	<i>Sub-peak</i>
L thalamus	-2, -6, 2	4.73	<i>Sub-peak</i>
Stage 3			
R middle occipital gyrus	36, -76, 20	4.29	293

R = right; L = left; vmPFC = ventromedial prefrontal cortex.

3.4.2.2 Theta power changes across construction stages

Having observed between-task differences in the two ROIs when each construction stage was considered separately, I next analysed the data in a different way to examine how

theta activity changed across the construction stages. For both scene and array conditions, theta power was attenuated across all stages relative to baseline (Figure 26B), which can be regarded as reflecting increased task-related neural activity. An ANOVA assessing the effects of task (scene, array) and construction stage (first, second, third) on anterior hippocampal activity revealed no significant effect of task ($F_{(1, 19)} = 0.694, p = 0.415$) or construction stage ($F_{(2, 38)} = 0.036, p = 0.965$), but a significant task \times construction stage interaction ($F_{(2, 38)} = 5.762, p = 0.007$; Figure 26B). This was driven by a significantly greater theta power decrease during the first construction stage of the scene task ($t_{(19)} = -2.713, p = 0.014$), indicating increased involvement of the anterior hippocampus during this scene stage, with engagement then decreasing over the subsequent two construction periods.

For the vmPFC, the ANOVA revealed a significant main effect of task ($F_{(1, 19)} = 17.252, p = 0.0005$; Figure 26B), but not of construction stage ($F_{(1.488, 28.272)} = 0.894, p = 0.392$), and there was no significant interaction ($F_{(2, 38)} = 2.586, p = 0.089$). The task effect was driven by larger theta power decreases during the scene relative to the array task during both the first ($t_{(19)} = -2.959, p = 0.008$) and second ($t_{(19)} = -3.723, p = 0.001$) construction stages, indicating the greatest vmPFC involvement during these periods of the scene task.

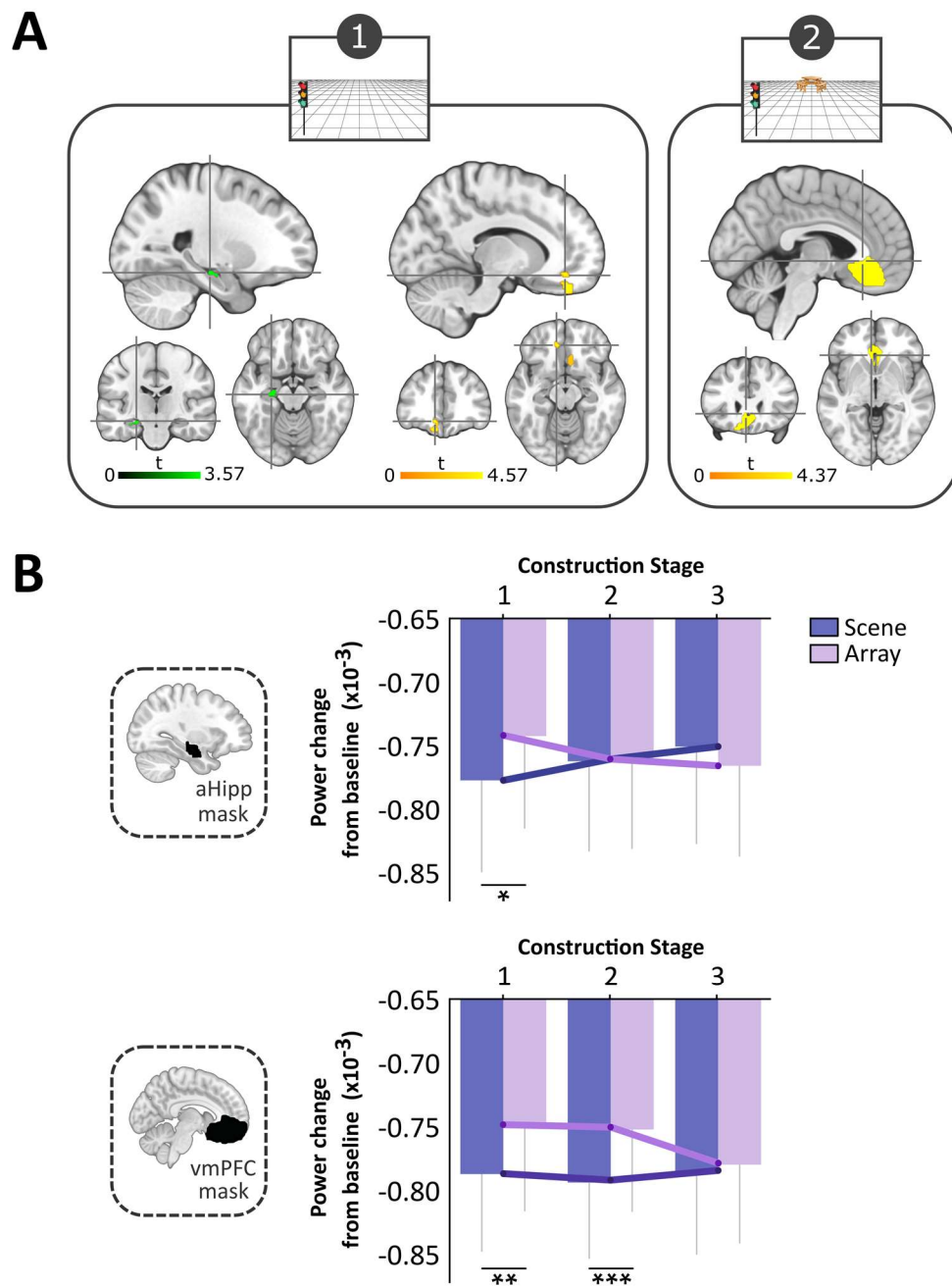


Figure 26. MEG ROI results. (A) Significant theta power changes within our two regions of interest, the anterior hippocampus (aHipp, shown in green) and ventromedial prefrontal cortex (vmPFC, shown in yellow) during scene relative to array construction (FWE $p < 0.05$, corrected for the ROI). Slices are centred on the overall peaks (crosshairs) in the left anterior hippocampus ($x, y, z = -24, -20, -16$) and left vmPFC ($-12, 40, -14$) during the first construction stage, and the vmPFC ($-4, 26, -4$) during the second period of construction. (B) The task \times construction stage interaction for the change in theta power relative to baseline in the anterior hippocampus, and the main effect of task in the vmPFC. For the hippocampus, power was significantly attenuated for the scene relative to the array condition during the first stage of construction. For the vmPFC, power was significantly attenuated for the scene compared to the array task for both the first and second construction periods. $*p = 0.014$, $**p = 0.008$, $***p = 0.001$. Error bars indicate ± 1 SEM.

3.4.2.3 *Effective connectivity during scene construction stages*

The source localisation analyses described above established the peak locations of theta power changes in the anterior hippocampus and vmPFC during scene compared to array construction. I next examined the direction of information flow between these regions during each task using DCM for cross spectral densities using the left hemisphere peaks in the anterior hippocampus and vmPFC identified during the first construction stage (Figure 26A). Three connectivity models were proposed based on established anatomical projections, representing the hypotheses for the hierarchical architecture of the hippocampus and vmPFC (Figure 27).

The first model predicted hippocampal activity drove that of the vmPFC. Hippocampal projections via the fornix terminate in the middle layers of the ventral portion of the vmPFC (Aggleton et al., 2015), so this connection was classified as “forward”. The second model specified the vmPFC as the influencing region. Connections from the vmPFC to the hippocampus are via the entorhinal cortex, where vmPFC inputs arrive at all cell populations (Rempl-Clower and Barbas, 2000), so were classified as “lateral” connections. The third model predicted that the hippocampus and vmPFC drove each other through reciprocal connections (mutual entrainment), so both forward and lateral connections were specified. I then fitted the models to each of the three 3000 ms scene and array construction stages. The approach adopted here to classifying DCM connections is also in line with previous work modelling hippocampal and vmPFC interactions in MEG (Barry et al., 2019a). As an extension to this previous DCM framework, I not only compared alternative unidirectional “master-slave” relationships, but also directly compared these models with bidirectional mutual entrainment, where neither region exerted a stronger influence over the other.

The models were compared using random effects BMS to determine the winning model at each construction period at the group level (Figure 27). During the initial construction period for scenes, the model most likely to explain the data across all participants was the vmPFC driving anterior hippocampal activity, with an EP of 62.45%, almost double that for the mutual entrainment model (35.78%). Furthermore, there was very little evidence for the opposing model of the anterior hippocampus driving vmPFC (1.77%). During the second scene construction stage, the mutual entrainment model was the most probable, with an EP of 55.08%, compared to the hippocampus-driven model (31.18%) and the vmPFC-driven model (13.74%). During the third scene construction stage, the winning model was again the mutual entrainment model (73.32%) compared to the hippocampus-driven model (22.28%) and the vmPFC-driven model (4.40%).

DCM of array construction revealed a different pattern of effective connectivity between the two ROIs. At the first stage of array construction, the model that best fitted the data was the anterior hippocampus driving vmPFC activity, with an EP of 76.74%. There was little evidence for either the vmPFC driving the hippocampus (12.06%) or mutual entrainment (11.20%). During the second array construction stage, there was no clear winning model (hippocampus-driven model 38.54%; vmPFC-driven model 30.38%; mutual entrainment model 31.08%). During the third array construction stage, the mutual entrainment model was the most probable (56.10%), followed closely by the vmPFC-driven model (40.75%), and least likely was the hippocampus-driven model (3.15%).

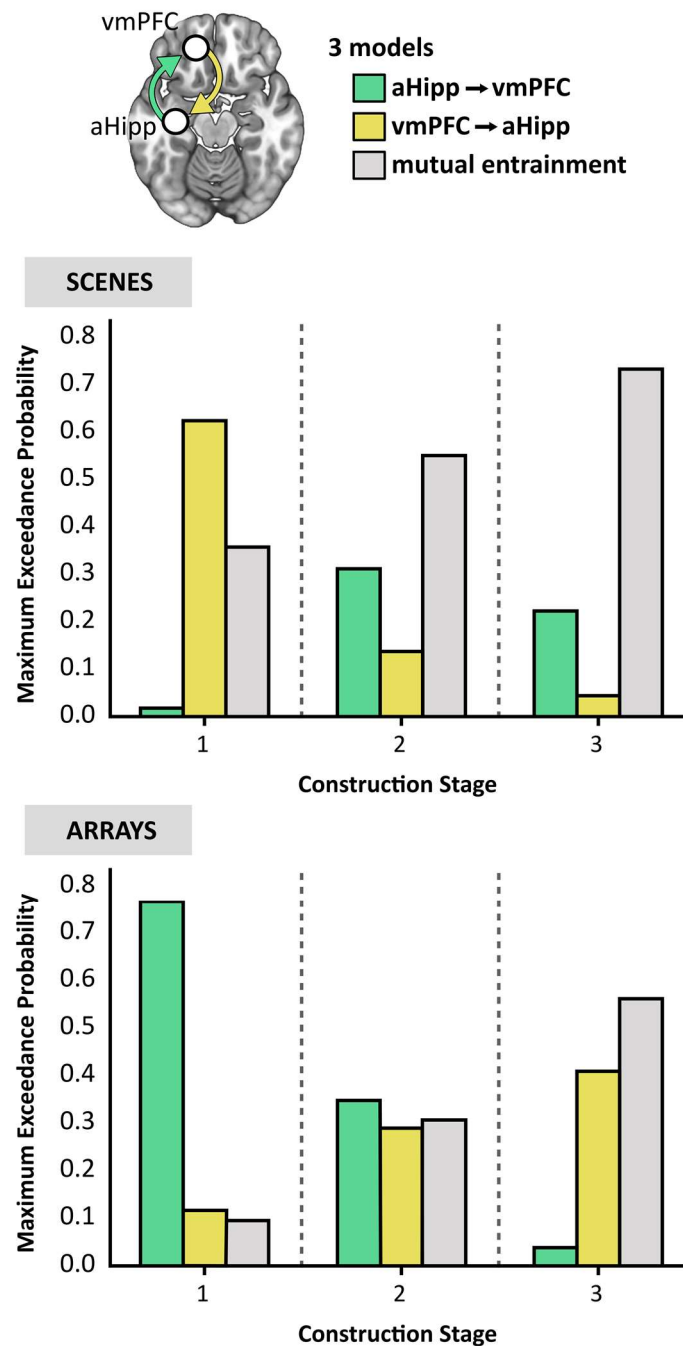


Figure 27. Effective connectivity between the anterior hippocampus (aHipp) and ventromedial prefrontal cortex (vmPFC) during step-by-step scene and array construction. The **top panel** shows the three models of effective connectivity. The **middle panel** displays the results of Bayesian model selection whereby the model with the strongest evidence during scene construction stage 1 was the vmPFC driving anterior hippocampus, while for the latter two stages, there was stronger evidence for the anterior hippocampus and vmPFC driving each other through reciprocal connections (mutual entrainment). The **bottom panel** shows the results for array construction, where the winning model during construction stage 1 was the anterior hippocampus driving the vmPFC, for the second stage there was no clear evidence for any model, and for the third stage there was stronger evidence for mutual entrainment.

3.5 Discussion

In this study I focused on the anterior hippocampus and vmPFC and investigated how their neural dynamics were associated with the step-by-step mental construction of scene and non-scene array imagery. Across three successive construction stages, scene-related attenuation of theta power was observed in the anterior hippocampus during the initial stage, and in the vmPFC during the first two stages, relative to array construction. I also found that the anterior hippocampal theta activity during the first phase of scene construction was driven by the vmPFC, with mutual entrainment between the two brain regions thereafter. By contrast, array construction was associated with the anterior hippocampus driving vmPFC activity during the first construction stage.

The power changes observed were, as predicted, specific to theta, aligning with the long association between theta and hippocampal function (reviewed in Colgin, 2016; Karakaş, 2020), and a previous MEG study of immediate full scene construction (Barry et al., 2019a). Finding significantly greater *attenuation* in theta power for the first (anterior hippocampus, vmPFC) and second (vmPFC) stages of scene, relative to array, construction also accords with the extant literature (Fellner et al., 2019; Solomon et al., 2019), where such theta power decreases also appear to reflect increased task-related neural activity (Fellner et al., 2016).

The question of primary interest was whether increased engagement of the anterior hippocampus and vmPFC would be specifically associated with the construction of scenes. As predicted, both brain regions were preferentially recruited during the mental construction of scenes compared to arrays, particularly during the early stages. My results converge with the previous fMRI findings of Dalton et al. (2018) and the MEG findings

of Barry et al. (2019a), and extend them in several ways. For example, in the latter MEG study, participants imagined a full scene within three seconds, and this was compared to a low-level baseline (imagining single objects) (Barry et al., 2019a). Dalton et al. (2018) also compared scenes and arrays, but could not examine the three construction stages separately given the slow nature of the haemodynamic response. By contrast, in the current study I examined the gradual step-by-step construction of scenes and showed that even at the earliest stage, with one object and a simple 3D space, this was sufficient to engage the anterior hippocampus and vmPFC, and more so than the first stage of array construction. Hence I could drill down further than before into how a scene develops, with a much tighter comparison in the form of the arrays, and in doing so show that a scene is “set” early on.

The scene and array tasks were very well matched in terms of the requirement for constructive and associative processing. However, it could be argued that the 3D space in the scenes somehow required more relational processing than the 2D space in the arrays, although it is unclear how the experience of depth (inherent to scenes) would translate into additional associative processing. Of note, when the imagined 3D and 2D spaces alone (without objects) were examined during fMRI by Dalton et al. (2018), neither was associated with increased hippocampal activity (see also Zeidman et al., 2012). Moreover, proponents of the relational theory are clear that mental operations involving 3D space – navigation being a prime example – engage the hippocampus because they are quite simply an example of associative processing, and that non-spatial stimuli involving associations should therefore also engage the hippocampus. Consequently, and given the close similarity of the stimuli in the scene and array tasks, I believe that, *a priori*,

relational theorists would have predicted similar hippocampal involvement for the scenes and arrays, which I did not find.

Nevertheless, if the scene stimuli involved more associative processing, how might I detect this in my data? I may have expected to observe a difference between scenes and arrays in terms of oculomotor behaviour and fixations, yet none were evident. If more relational processing was required for scenes then I might have predicted differential performance on the surprise post-scan item memory test, but none was found. Of note, when Dalton et al. (2018) examined associative memory for the same stimuli in a surprise post-fMRI scan memory test, there was also no difference between scenes and arrays (in fact, both were at chance, as this is a very challenging task). As part of my own piloting for the current study, a separate group of participants performed an equally difficult surprise source memory task, where they were asked to decide whether each object description was originally encountered as part of a scene or an array. Again, there was no difference between tasks, with both performed at chance level. Of course, I acknowledge that the absence of evidence can be difficult to interpret. However, given the simplicity of the stimuli (three objects and a space) and their similarity on measures of item memory and oculomotor behaviour, I believe that a parsimonious interpretation of this paradigm is that the two conditions were comparable in terms of associative processing.

Why might scene construction be particularly engaging for the anterior hippocampus and vmPFC? Scene imagery seems to be important for, and is frequently deployed in the service of, hippocampal-dependent tasks including episodic memory, imagining the future and spatial navigation (Clark et al., 2019, 2020). For example, using a large sample ($n = 217$) of participants and multiple cognitive tests with a wide spread

of individual differences in performance, Clark et al. (2019) reported that the ability to construct scene imagery explained the relationships between episodic memory, imagining the future and spatial navigation task performance. The prominence of scene imagery was further emphasised in another study involving the same sample, where the explicit strategies used to perform episodic memory recall, future thinking and spatial navigation tasks was assessed (Clark et al., 2020). In each case, the use of scene visual imagery strategies was significantly higher than for all other types of strategies (also see Andrews-Hanna et al., 2010). The influence of scene imagery across cognition is perhaps not surprising, given that it mirrors how most people experience the world.

My findings also provide insights into the interactions between the anterior hippocampus and vmPFC, addressing the prediction made in Chapter 1 that the vmPFC would initiate scene construction. McCormick et al. (2018a) and Ciaramelli et al. (2019) proposed that the vmPFC initiates the activation of schematic and other knowledge in neocortex that is relevant for a scene, whilst inhibiting elements that are not relevant. This information is conveyed to the hippocampus which constructs the scene. vmPFC then engages in iterations via feedback loops with neocortex and hippocampus to mediate the schema-based retrieval and monitoring necessary to flesh-out the scene. That the vmPFC drove activity in the anterior hippocampus during the first scene construction phase supports the perspective that the vmPFC instigates this process. After this initial scene-setting, hippocampal and vmPFC activity was mutually entrained, perhaps reflecting iterative feedback loops between them as each new object was integrated to form a fully realised scene.

The activation of schemas by vmPFC may provide an initial rudimentary template into which a scene is constructed in collaboration with the hippocampus (van Kesteren

et al., 2010a; Ghosh et al., 2014). Although my scene stimuli were composed of object triplets that were not contextually associated (van Kesteren et al., 2012), there is evidence for schema-guided construction of novel representations, driven by the vmPFC (Garrido et al., 2015). Furthermore, the observed tonic response of the vmPFC throughout the first two construction stages in my study may reflect the prediction and preparatory processing necessary for the subsequent integration of scene elements (McCormick et al., 2018a; Ciaramelli et al., 2019).

A very different result was evident when connectivity between the two ROIs was examined during the array condition. For the first stage of this construction, the anterior hippocampus drove vmPFC activity. Why I observed this direction of connectivity for the arrays is not clear. I speculate that during array construction, with the vmPFC much less engaged compared to the scenes during the first two construction stages, perhaps the hippocampus was “reaching out” for guidance from the vmPFC. Interestingly, at the third stage, the vmPFC became equally engaged for arrays and scenes, and the DCM analysis showed mutual entrainment to be the most probable explanatory model. This may reflect the vmPFC stepping in to provide some input. Of course, this is conjecture, and future studies are required to explore this finding further. It does suggest, however, that scenes elicit a different kind of neural processing to arrays, during which the vmPFC is the co-ordinator. Arrays, by contrast, are not mental representations within which a person could reasonably operate, and consequently do not generate the same cascade of neural processing initiated by the vmPFC.

In conclusion, the hippocampus and vmPFC are particularly implicated in supporting cognitive functions such as episodic memory, imagining future events and spatial navigation. While it is certainly likely these two brain areas participate in associative

and constructive processing more generally (Dalton et al., 2018), here I have shown that their neural activity seems to be especially tuned to the earliest stages of building scene representations. In the future, with the use of high spatial resolution neuroimaging, it will be important to elucidate the contributions of other brain regions, including within the posterior parietal cortex (Sakata and Kusunoki, 1992; Epstein and Higgins, 2007; Dalton and Maguire, 2017), to this visuospatial process and to characterise putative anatomical pathways that facilitate hippocampal-vmPFC interactions in the human brain.

Having established a preferential response in the hippocampus and vmPFC to imagining scenes, another consideration is the influence of the objects that compose these mental representations, and this is the subject of the next chapter.

4 The influence of objects on building scenes

4.1 Précis

In the previous chapter, I investigated the neural dynamics associated with the step-by-step mental construction of scenes, and whether this processing was associated with scenes in particular, or shared by multi-element representations more generally. I found that anterior hippocampal and vmPFC theta power was preferentially modulated by scenes, and this was driven predominantly by vmPFC activity at the earliest stages of scene construction. In this second study, I turned my attention to a particularly prominent component of scenes – objects. One property of objects that might be pertinent to scene construction is whether they are space-defining or space-ambiguous. Using the same scene construction paradigm and stimuli as those deployed in Chapter 3, in the current experiment I focussed on the neural dynamics associated with space-defining and space-ambiguous objects while they were being incorporated into scene representations.

4.2 Introduction

As described in Chapter 1 and briefly recapitulated here for convenience, neuroimaging and neuropsychological studies have identified a number of brain areas that seem to be particularly engaged during the viewing and imagination of scenes including the vmPFC (Zeidman et al., 2015a; Bertossi et al., 2016; Barry et al., 2019a), the anterior hippocampus

(Hassabis et al., 2007a, 2007b; Summerfield et al., 2010; Zeidman et al., 2015a, 2015b; Dalton et al., 2018; reviewed in Zeidman and Maguire, 2016), the PHC (Epstein and Kanwisher, 1998; reviewed in Epstein, 2008 and Epstein and Baker, 2019), and the RSC (Park and Chun, 2009; reviewed in Epstein, 2008; Vann et al., 2009; Epstein and Baker, 2019). While spatial aspects of scenes have been amply investigated and linked to the hippocampus (Byrne et al., 2007; Morgan et al., 2011; Epstein et al., 2017; Epstein and Baker, 2019), the higher-order properties of objects within scenes have received comparatively less attention (Auger et al., 2012; Troiani et al., 2014; Julian et al., 2017; Epstein and Baker, 2019), and yet they could influence how scene representations are constructed by the brain. How might they contribute to the building of scene representations?

In Chapter 1, I introduced one particular object attribute that seems to play a role in scene construction. Mullally and Maguire (2011; see Kravitz et al., 2011 for related work) observed that certain objects, when viewed or imagined in isolation, evoked a strong sense of 3D local space surrounding them (SD objects), while others did not (SA objects), and this was associated with engagement of the PHC during fMRI. In other words, SD objects seem to anchor or impose themselves on their surrounding space and in doing so they define the space around them, despite no background or spatial context being present. By contrast, SA objects do not anchor themselves in their surrounding space in the same way. Participants in the Mullally and Maguire study (2011) were asked to describe the difference between an SD and an SA object, and a typical response included: “SD items conjure up a sense of space whereas SA items float - they go anywhere.” Many participants spontaneously used the word “float” when discussing SA objects, emphasizing the detachment of these objects from an explicit sense of

surrounding space. Mullally and Maguire (2011) also found that SD objects tended to be less portable, maintaining a permanent location, although a considerable number of SA objects had reasonably high permanence also, suggesting that permanence, while certainly related to SD, is not the sole basis of the SD effect. Indeed, object permanence has been linked to the RSC rather than the PHC (Auger et al., 2012, 2015; Auger and Maguire, 2013; Troiani et al., 2014). Similarly, Mullally and Maguire (2011) found that object size and contextual associations did not account for the SD effect.

In a subsequent behavioural study, participants showed a strong preference for SD objects when given a choice of objects with which to mentally construct scenes, even when comparatively larger and more permanent SA objects were available (Mullally and Maguire, 2013). Moreover, when deconstructing scenes, participants retained significantly more SD objects than SA objects. It therefore seems that SD objects might enjoy a privileged role in scene construction.

Mullally and Maguire (2011) examined SD and SA objects in isolation. However, given their apparent influence during scene construction (Mullally and Maguire, 2013), in the current study I compared neural responses to SD and SA objects while they were being used to build imagined scene representations. I adapted the paradigm from Dalton et al. (2018) and my previous experiment reported in Chapter 3, where participants gradually built a scene image from three successive auditorily-presented object descriptions and an imagined 3D space.

In order to capture the neural dynamics associated with the points during scene construction when either SD or SA objects were being imagined, I once again leveraged the high temporal resolution of MEG. In previous MEG studies, changes in anterior

hippocampal and vmPFC theta were noted when participants imagined scenes in response to scene-evoking cue words (Barry et al., 2019a, 2019b). This was also true in the previous chapter when scene imagery was gradually built. However, the effects, if any, of SD and SA objects on neural responses during scene construction remains unknown.

Here, I performed source-level power analyses to examine whether a difference between SD and SA objects could be detected in the vmPFC, anterior hippocampus, PHC, and RSC, given their acknowledged involvement in scene processing. I also characterised the effective connectivity between any brain regions that emerged from this analysis. My focus was on theta (4-8 Hz), due to accumulating evidence for a functional role of theta in episodic memory processes (e.g. see Colgin, 2013; Fuentemilla et al., 2014; Garrido et al., 2015; Backus et al., 2016; Berens and Horner, 2017; Kaplan et al., 2017; McCormick et al., 2020) and scene construction (Barry and Maguire, 2019a; Barry et al., 2019b), including my own MEG study (Chapter 3). I predicted a decrease in theta power during scene construction when SD objects were being imagined, consistent with previous findings of attenuated power associated with the immediate construction of scenes (Barry et al., 2019a, 2019b) and autobiographical memory recall (McCormick et al., 2020) using MEG. As noted in Chapter 3, in light of the association between low frequency power attenuations and fMRI BOLD increases (Fellner et al., 2016), I considered decreases in theta localised to a particular brain region during the processing of SD objects to represent that area's enhanced engagement. While my main interest was in theta, I nevertheless also examined other frequencies.

The obvious prediction, given the previous Mullally and Maguire (2011) fMRI study, was that PHC would be engaged by SD objects. However, because all stimuli were scenes – unlike previous investigation of the SD/SA object property – and the key

manipulation of SD and SA objects within scenes was so subtle, I retained an open mind about which brain areas might distinguish between the two object types.

4.3 Materials and methods

4.3.1 *Participants*

Twenty three healthy, right-handed participants (13 females; mean age = 25.35 years; Std Dev = 3.69) took part in the experiment.

4.3.2 *Stimuli and task procedure*

The task involved participants gradually constructing simple scenes in their imagination from a combination of auditorily-presented SD and SA object descriptions (see examples in Figure 28A) and a 3D space. The auditory object descriptions were identical to those used in my previous study, where they composed either scenes or non-scene arrays (see Chapter 3), a paradigm originally implemented during fMRI by Dalton et al. (2018). However, in the current experiment, participants only mentally constructed scenes.

Dalton et al. (2018)'s study involved a separate group of participants who rated each object as SD or SA. For use in the current study, I ensured these stimuli types were closely matched on utterance length ($Z = 1.643, p < 0.1$) and number of syllables ($Z = 1.788, p < 0.074$). Unsurprisingly, SD objects were rated as more permanent than SA objects ($Z = 5.431, p < 0.001$). All objects were rated as highly imageable, obtaining a score of at least 4 on a scale from 1 (not imageable) to 5 (extremely imageable). The participants in the current MEG experiment were unaware of the SD-SA distinction. Just as was the case for my experiment in Chapter 3, objects in each triplet were not

contextually related to each other. A detailed description of the scene imagination task is given in Chapter 3, however, as there are some small differences in this second study, I briefly reprise key details.

Each trial started with a visual cue (4000 ms), which displayed the configuration of locations at which objects should be imagined in the upcoming trial (Figure 28B). Four different cue configurations (Figure 29) were randomised across the five scanning blocks. During pre-scan training, I emphasised the importance of following the cue configurations as precisely as possible. This ensured consistency of oculomotor behaviour across participants, and that objects were imagined as separate and non-overlapping. Participants then fixated on the screen centre (1000 ms). During the scene construction task (~9000 ms; Figure 28B), keeping their eyes open whilst looking at a blank screen, participants first imagined a 3D grid covering approximately the bottom two-thirds of the blank screen. Upon hearing each of three auditory descriptions, one at a time, they imagined the objects in the separate, cue-specified positions on the 3D grid. They were instructed to move their eyes to where they were imagining each object on the screen, but also to maintain imagery of previous objects and the grid in their fixed positions. Each construction stage consisted of a ~2000 ms object description and a silent 1000 ms gap before the presentation of the next object. An additional 1000 ms at the end of scene construction avoided an abrupt end to the task. By the end of a trial, participants had created a mental image of a simple scene composed of a 3D grid and three objects. Vividness of the entire scene was then rated on a scale of 1 (not vivid at all) to 5 (extremely vivid). An ITI (2000 ms) preceded the next trial.

Participants imagined a total of 66 scenes (composed of 99 SD and 99 SA objects). Each object description was heard only once. The order of presentation of SD and SA

objects within triplets was balanced across scenes with an equal number of SD and SA objects in the first, second and third construction stages (33 in each). In total there were six orders, half of which began with SD objects, and the other half with SA objects³. A similar number of scenes followed each order (between 10 and 12), presented pseudorandomly across five scanning blocks. This ensured that no more than two trials comprising the same object order were presented consecutively. As an example, Figure 5B illustrates the order SD-SD-SA. Each order contained a mixture of both SD and SA objects. The orders SD-SD-SD and SA-SA-SA were therefore not used. This careful control over SD/SA order ensured any SD/SA differences were not driven by over-representation of either SD/SA objects at particular points in the scene construction process.

Another task was also included in this study. As in the Chapter 3, participants heard a backward sequence of three numbers, after which they were instructed to mentally continue counting backwards. The sole function of this task was to provide participants with relief from the effortful imagination required of the scene and array trials. This counting task was not subject to analysis. Seven additional catch trials (5 scenes, 2 counting) were also included to ensure that participants maintained attention – participants pressed a button upon hearing a repeated object description or number

³ SD-SA-SA
SD-SD-SA
SD-SA-SD
SA-SD-SD
SA-SD-SA
SA-SA-SD

within a triplet. These additional trials were pseudorandomly presented across blocks such that every block contained at least one catch trial.

The visit of each participant took approximately 2.5 hours in total, including training, set up in the MEG scanner and eye tracker calibration, the experimental task (~50 minutes), and the post-scan memory test.

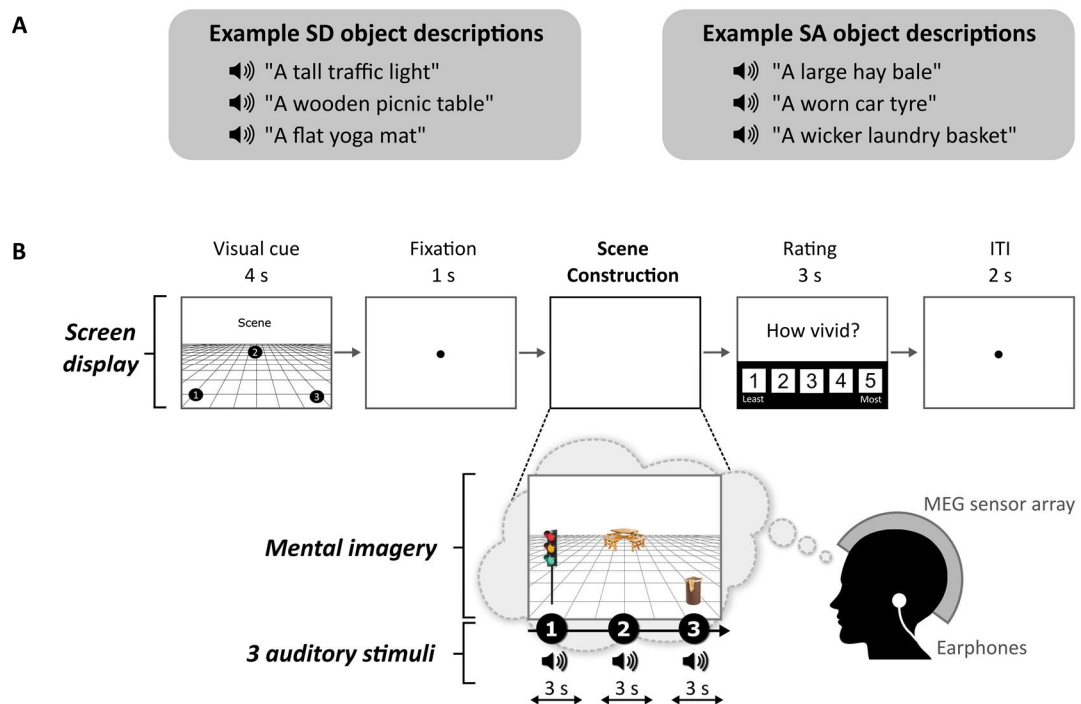


Figure 28. Example stimuli and trial structure. (A) Examples of space-defining (SD) and space-ambiguous (SA) object descriptions, three of which are used in the example trial in part B. (B) The structure and timings of an example trial. An example cue description is shown on the left. Note that during the scene construction period the participants imagined the simple scenes while looking at a blank screen, and never saw visual objects. Images here merely serve as illustrations for the reader.

4.3.3 In-scanner eye tracking

Eye movements and fixations were recorded during the MEG scan using an Eyelink 1000 Plus (SR Research) eye tracking system with a sampling rate of 2000 Hz to monitor task

compliance. For some participants the calibration was insufficiently accurate, leaving 19 data sets for the eye tracking analyses.

4.3.4 Post-scan surprise memory test

If I found a neural difference between SD and SA objects, it could be argued this was due to a difference in encoding success. For example, perhaps SD objects were more readily encoded into memory than SA objects, even though there was no explicit memory encoding requirement in this task and participants were only instructed to imagine stimuli as vividly as possible. By conducting a surprise memory test for SD and SA objects immediately post-scan, I could examine this issue. Participants were presented with a randomised order of all previously heard auditory object descriptions (198 objects in total) and an additional 33 SD and 33 SA new object descriptions (lures). After hearing each item, they indicated whether or not they recalled hearing the object description during the scan, and their confidence in their decision (1 = low, to 5 = high).

4.3.5 Behavioural data analysis

In-scanner vividness was compared between SD-majority (where 2 out of the 3 objects in the scene were SD) and SA-majority scenes (where 2 out of the 3 objects in the scene were SA) using a paired-samples t-test. An assessment of individual vividness per SD and SA object was not possible, since participants scored the vividness of the overall scene at the end of each trial. To obtain individual scores it would have been necessary to interrupt the construction of a scene each time an object was imagined, which would have changed the nature of the task entirely. Eye tracking data were analysed using two-way repeated measures ANOVAs, looking at individual object stages.

Memory performance was assessed using the sensitivity index d-prime (d'), based on Signal Detection Theory (Macmillan and Creelman, 1990). This takes into account the relative proportion of “hits” (objects identified as familiar when they were familiar) and “false alarms” (object identified as familiar when they were novel), using the formula:

$$d' = z(H) - z(FA)$$

where z denotes a z-score transformation, H is the hit rate (hits/[hits + misses]), and FA is the false alarm rate (false alarms/[false alarms + correct novel]). Perfect scores for H and FA were adjusted using the formulae (Macmillan and Creelman, 1991):

$$H = 1 - \frac{1}{2N} \quad FA = \frac{1}{2N}$$

where, N is the number of familiar items for the H correction, or the number of novel items for the FA correction. I also quantified the bias of participants to control for any tendency participants might have to give one type of response more than another. This response bias ' c ' was calculated using the formula (Macmillan and Creelman, 1990):

$$c = -0.5(H + FA)$$

Differences in d' and c as a function of object type (SD, SA) and construction stage (first, second, third) were each analysed using a two-way repeated measures ANOVA.

Statistical analyses were performed in SPSS25 (<https://www.ibm.com/products/spss-statistics>) using a significance threshold of $p < 0.05$. In cases where Mauchly's test found sphericity violated, Greenhouse-Geisser adjusted degrees of freedom were applied.

4.3.6 MEG data acquisition

MEG data were recorded using the whole-head 275 channel CTF Omega MEG system with a sampling rate of 1200 Hz, as described in Chapter 2.

4.3.7 MEG data pre-processing

MEG recordings were filtered with a 1 Hz high-pass filter, 48-52 Hz stop-band filter, and 98-102 Hz stop-band filter, to remove slow drifts in signals from the MEG sensors and power line interference. Epochs corresponding to each scene construction period were defined as 0 to 3000 ms relative to the onset of the SD or SA object description. A baseline period was defined as 0 to 1000 ms relative to the onset of the fixation preceding the imagination period. This period provided the most appropriate baseline against which task-related activity could be compared, since the ITI (another fixation period immediately following the rating) data were comparatively more noisy. Epochs were concatenated across trials for each condition and across scanning blocks.

4.3.8 MEG data analysis

All MEG analyses were conducted using SPM12 (www.fil.ion.ucl.ac.uk/spm).

4.3.8.1 Source reconstruction: SD and SA objects

Source reconstruction was performed using the LCMV beamformer (Van Veen et al., 1997), performed using the DAiSS toolbox (<https://github.com/SPM/DAiSS>), and visualised in MRICroGL (<https://www.mccauslandcenter.sc.edu/mricrogl>) using the MNI 152 T1 image. I used this approach to estimate power differences between SD and SA objects in source space within selected frequency bands: theta (4-8 Hz), alpha (9-12 Hz),

beta (13-30 Hz), and gamma (30-100 Hz). By repeating the beamformer analysis in bands > 8 Hz I could determine whether effects were confined to the theta range.

For each participant, a set of filter weights was built based on data from the SD and SA conditions within each frequency band, and a 0 to 3000 ms time window relative to object description onset, encapsulating a construction period. Coregistration to MNI space was performed using a 5 mm volumetric grid and was based on nasion, left and right preauricular fiducials. The forward model was computed using a single-shell head model (Nolte, 2003).

At the first level, power in a particular frequency band was estimated to create one image per object type (SD or SA) per participant. Images were spatially smoothed using a 12 mm Gaussian kernel and entered into a second-level random effects paired *t*-test to determine power differences between SD objects and SA objects across the whole brain. An uncorrected whole brain threshold of $p < 0.001$ with a cluster extent threshold of > 100 voxels was applied to each contrast, in line with numerous other MEG studies (e.g. Lieberman and Cunningham, 2009; Hanslmayr et al., 2011; Kveraga et al., 2011; Kaplan et al., 2012; Cushing et al., 2018; Barry et al., 2019a). This is held to provide a balance between protecting against false positives whilst enabling the detection of subtler signals. This was followed by ROI analyses, using a FWE corrected threshold of $p < 0.05$ and a spatial extent threshold of 10 voxels for each ROI, performed using separate bilateral masks covering the vmPFC, anterior hippocampus, PHC, and RSC. *A priori*, I identified these regions as having particularly important roles in the mental construction of scenes, SD objects, and a related object property of permanence. For completeness, the posterior hippocampus was also included.

The contrast of primary interest was the direct comparison between SD and SA objects. As a secondary issue, I was mindful that it was also of interest to know whether an effect represented an increase or a decrease in power from baseline. Including baseline correction in the original beamformer would be the usual way to examine this. However, as was the case in Chapter 3, this was challenging to implement in the case of the current design because an SD or SA object was not necessarily preceded by a clean baseline. On each trial, three object descriptions were heard one after the other. While the first object in a triplet was preceded by a clean 1000 ms fixation baseline, objects 2 and 3 had an object imagination stage preceding them. Objects at different construction stages would therefore be corrected against different types of baseline (i.e., object 1 vs preceding fixation, object 2 vs preceding object 1, and object 3 vs preceding object 2). Conducting this analysis could have introduced spurious effects. Consequently, having established that a difference was apparent between SD and SA objects when they were compared directly, I then sought to ascertain the direction of power change. By using a separate beamformer where the 1000 ms pre-stimulus fixation period was the only baseline for all objects (whether a SD/SA, or first/second/third object), and hence did not overlap with any stimulus period, I was able to establish the direction of power change for SD and SA objects in a straightforward way. To clarify, each experimental condition (SD, SA) comprised objects from across all three construction stages – SD or SA objects within each individual stage could not be examined separately due to having only a small number of trials per stage and object type. This is considered further in Section 4.4.2.3.

4.3.8.2 *Source reconstruction: permanent and non-permanent objects*

It was important to also determine whether any observed source localised differences in power for the SD vs SA contrasts were distinct from the effect of object permanence.

The same LCMV beamforming protocol described above was followed when objects were re-categorised as permanent and non-permanent, with the number of permanent and non-permanent objects equalised to 65 in each category. The categorisation of permanence was made according to ratings from a separate group of participants in the Dalton et al. (2018) study.

4.3.8.3 *Effective connectivity*

Brain areas identified in the whole brain SD versus SA beamformer provided the seed regions for the effective connectivity analysis, which was conducted using DCM for cross spectral densities (Moran et al., 2009). This followed the same procedure as that described in Chapter 3. This approach permitted me to compare different biologically plausible models of how one brain region influences another, as well as mutual entrainment between regions (Friston, 2009; Kahan and Foltynie, 2013; also see Section 2.7.3.1 in Chapter 2 for a more detailed explanation of this DCM approach). Random-effects BMS was performed to compare the evidence for each specified model that varied according to which connections were modulated by SD relative to SA objects (Stephan et al., 2009). I determined the winning model to be the one with the greatest exceedance probability. As only a single DCM was performed (SD versus SA), when the two top models were identified by BMS (i.e. those with the highest exceedance probabilities), these two models could be further compared by calculating the log Bayes factor for each participant (see also Section 2.7.3). This enabled an assessment of the consistency of the model fit across participants, and is quite simply the ratio of model evidences between the two top models.

4.4 Results

4.4.1 Behavioural results

Adherence to task-specific instructions and any behavioural SD compared to SA differences that could have influenced the electrophysiological data were examined first.

There was no significant difference in the vividness of mental imagery between SD-majority (mean = 3.91, Std Dev = 0.69) and SA-majority (mean = 3.89, Std Dev = 0.66) scene trials ($t_{(22)} = 0.464, p = 0.647$). Participants correctly identified on average 97.52% (Std Dev = 0.39) of catch trials, indicating that they attended throughout the experiment.

The effect of object type (SD, SA) and construction stage (first, second, third) on in-scanner fixation count (FixCount) and fixation duration (FixDur) showed that there were no significant main effects of object type (FixCount: $F_{(1,18)} = 1.908, p = 0.184$; FixDur: $F_{(1,18)} = 0.086, p = 0.772$) or construction stage (FixCount: $F_{(2,36)} = 0.292, p = 0.748$; FixDur: $F_{(2,36)} = 0.535, p = 0.590$), and no object type \times construction stage interaction (FixCount: $F_{(2,36)} = 0.710, p = 0.499$; FixDur: $F_{(2,36)} = 1.871, p = 0.169$). Heat maps of the spatial patterns of fixations during the task demonstrated a consistent adherence to cue configuration instructions across participants (Figure 29).

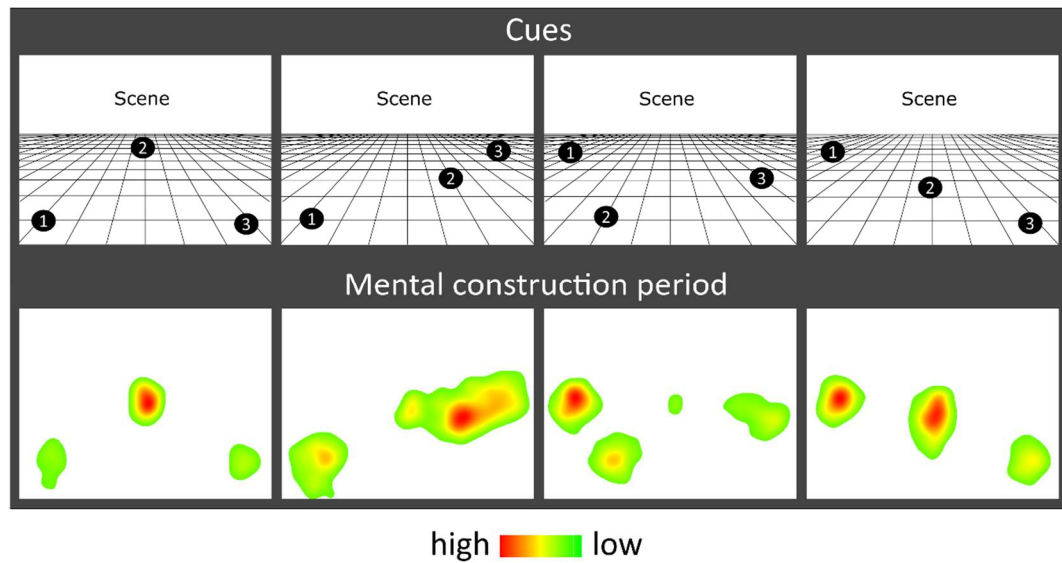


Figure 29. Eye fixations. Heat maps of the fixation count during the 9000 ms mental construction period following each cue configuration. Each heat map is an aggregate of fixations on the blank screen across all trials for that cue configuration across all participants with eye tracking data ($n=19$). Red indicates higher fixation density and green lower fixation density.

In terms of recognition memory (see Table 3), performance exceeded 80% correct for both SD and SA objects, and for d' and c there were no significant effects of object type (d' : $F_{(1,22)} = 0.469$, $p = 0.500$; c : $F_{(1,22)} = 0.012$, $p = 0.915$), construction stage (d' : $F_{(2,44)} = 2.383$, $p = 0.104$; c : $F_{(2,44)} = 0.120$, $p = 0.887$), nor were there any interactions (d' : $F_{(2,44)} = 1.431$, $p = 0.250$; c : $F_{(2,44)} = 0.035$, $p = 0.965$).

Table 3. Results of the surprise post-scan object recognition memory test.

	SD objects	SA objects
	Mean (Std Dev)	Mean (Std Dev)
% correct	81.063 (7.332)	80.074 (8.198)
Hit rate	0.809 (0.079)	0.801 (0.091)
False alarm rate	0.156 (0.108)	0.156 (0.103)
d'	2.045 (0.581)	2.001 (0.611)
c	0.091 (0.282)	0.088 (0.272)

SD = space-defining; SA = space ambiguous; Std Dev = Standard Deviation; % = percent; d' = d -prime; c = response bias discrimination.

4.4.2 MEG results

4.4.2.1 Source space power differences between SD and SA objects

A whole brain beamforming analysis revealed significant theta power attenuation for SD compared to SA objects in only two regions: the right vmPFC (peak x, y, z = 12, 60, -8; t -value = 3.66; cluster size = 1960) and right superior temporal gyrus (STG; peak x, y, z = 66, -6, -12; t -value = 3.76; cluster size = 1197) (Figure 30A). Engagement of the vmPFC was further confirmed in an ROI analysis, whereas changes in power in the other ROIs (anterior hippocampus, posterior hippocampus, PHC, and RSC) were not evident. The opposite contrast (SA versus SD) revealed no relative attenuations in power.

A subsequent contrast between each object type and the baseline revealed that these theta power changes were decreases, echoing numerous previous reports of power attenuation during the construction of scene imagery (e.g., Guderian et al., 2009; Barry et al., 2019a, 2019b) and episodic memory recall (e.g., Solomon et al., 2019; McCormick et

al., 2020). No power increases were found for either condition relative to baseline. In line with this preceding work, I interpret these observed decreases as enhanced involvement of the vmPFC and STG during the imagination of SD objects.

Analysis of alpha, beta and gamma frequency bands showed no significant power differences when SD and SA objects were compared.

4.4.2.2 *Source space power differences between permanent and non-permanent objects*

Having identified a significant effect of SD relative to SA objects on theta power in two brain regions, I next examined whether similar theta power changes were associated with permanent compared to non-permanent objects. As outlined in the introduction to this chapter, SD/SA and object permanence are related constructs and, unsurprisingly, a difference in perceived permanence was observed between SD and SA objects in the stimulus set used in the current study. However, permanence has been previously linked with the RSC, while the SD effect has been associated with the PHC (e.g. see Mullally and Maguire, 2011; Auger et al., 2012, 2015; Auger and Maguire, 2013; Troiani et al., 2014), suggesting these properties are represented differently in the brain.

No significant differences in theta power were observed between permanent and non-permanent objects across the whole brain, or in any of the ROIs. Indeed, when the whole brain threshold was lowered further (uncorrected $p < 0.005$) and with no extent thresholding, I only found small clusters with peaks in the visual association cortex, inferior parietal lobule, and dorsolateral prefrontal cortex. Notably, no power changes were apparent in the vmPFC or STG. This suggests that although permanence and

SD/SA are related features, they engaged different brain regions, as has been previously reported.

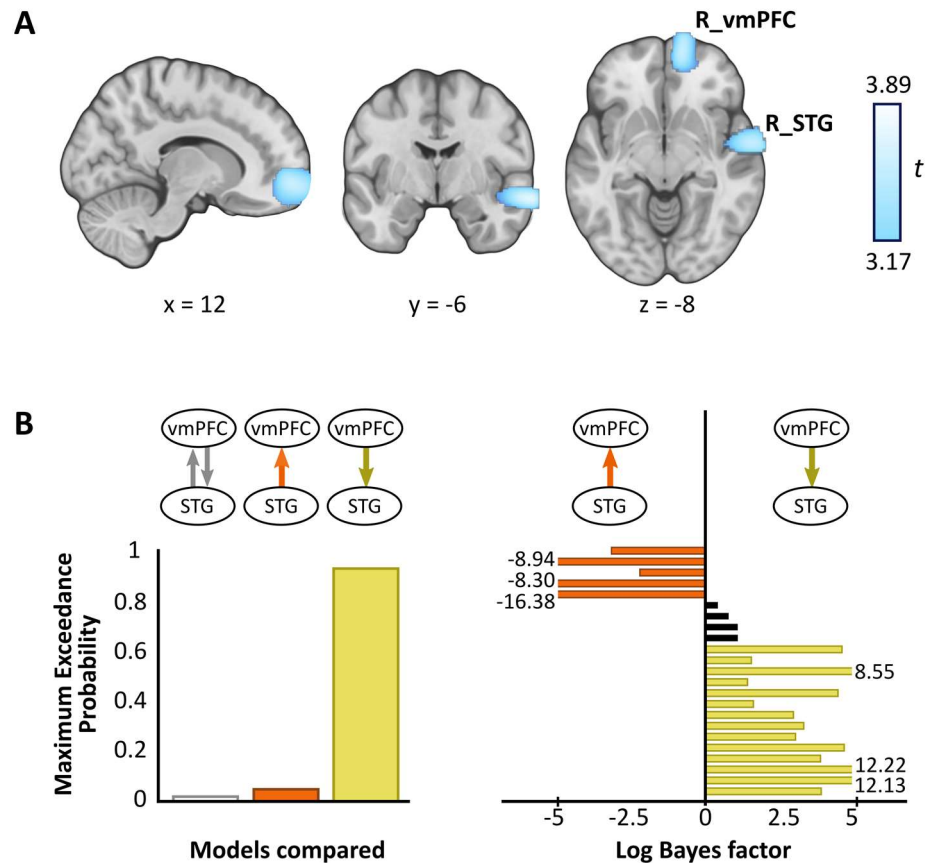


Figure 30. MEG results. (A) Source reconstruction of theta (4-8 Hz) power changes during SD relative to SA objects revealed attenuation in the ventromedial prefrontal cortex (vmPFC) and superior temporal gyrus (STG). Images are superimposed on the MNI 152 template and thresholded at uncorrected $p < 0.001$. (B) Effective connectivity between vmPFC and STG was examined using DCM. Three models were compared, with vmPFC driving STG (yellow) theta activity during SD compared to SA objects being the model that best explained the data (**left panel**). Log Bayes factors per participant (**right panel**) showed positive to strong evidence for this model in most participants (yellow bars). Participants for whom there was no conclusive evidence for either model are represented by black bars. Where log Bayes factors exceeded five, bars are truncated and the exact values are adjacently displayed.

4.4.2.3 *Source space power differences for each construction stage separately*

The aim of the current study was to address whether SD and SA object construction couched within scenes could be differentiated using MEG, firstly in terms of power in a specific frequency band, as outlined above, and secondly in terms of effective connectivity, which I consider in the next section. Before I address the latter, I was also interested to ascertain whether there would be sufficient statistical power to examine SD/SA differences in spectral power at each construction stage separately (first, second, third). This means there would be just 33 objects per category. This study was not designed for such comparisons.

Nevertheless, to explore this point further, I re-ran the SD versus SA comparison with 33 objects in each category, randomly selected per participant (instead of the 99 versus 99 objects in the full analysis). This mirrors the total number of objects that would be available per category per construction stage. As anticipated, I no longer found any theta power differences between the SD and SA conditions, either across the whole brain (even when lowering the threshold to $p < 0.05$ uncorrected) or in the ROI analyses. This strongly suggests there would be insufficient sensitivity to detect effects of SD versus SA at each construction step, making definitive inferences difficult, even in the event of finding differences.

An obvious question is why I did not simply increase the number of scenes in each condition enabling me to boost power for SD/SA comparisons at each construction stage, but this was not possible for several reasons. Participants were already in the MEG scanner for more than one hour, were required to remain very still in a seated position, alert and to create vivid mental imagery throughout. A large number of additional trials

would have made the scanning session unmanageably long for participant comfort, with an effect not only on motion artefacts, but also most likely on their ability to concentrate and therefore the vividness of mental imagery.

4.4.2.4 *Effective connectivity during SD relative to SA objects*

Having established a response to object type in the vmPFC and STG, I next examined the effective connectivity between these regions. I tested three simple hypotheses: (1) vmPFC and STG are mutually entrained, (2) STG drives vmPFC, or (3) vmPFC drives STG. I embodied each hypothesis as a DCM where models differed in terms of which connection could be modulated by SD relative to SA objects.

BMS identified the winning model to be vmPFC driving STG during SD more so than SA objects, with an exceedance probability of 91.62% (Figure 30B, left panel). This model was also the most consistent across participants (Figure 30B, right panel). Using Kass and Raftery's (1995) classification, there was “positive” evidence (log Bayes factor 1.1 to 3) in favour of this model for 14 of the 23 participants (Figure 30B, right panel, yellow bars), compared to only 5 participants for the STG driving vmPFC (Figure 30B, right panel, orange bars). For 4 participants, there was no conclusive evidence for either model (Figure 30B, right panel, black bars). It should be noted that “strong” evidence (log Bayes factor ≥ 3) for vmPFC driving STG was observed in only 11 participants overall, suggesting the SD vs SA effect on effective connectivity is relatively small. This is unsurprising, considering the likely subtlety of the effect and the well-controlled nature of the stimuli and paradigm.

4.5 Discussion

In this study I focused on an object property, SD or SA, that has been shown to influence how scene imagery is constructed (Mullally and Maguire, 2013). I found that while these object types were being imagined during scene construction, SD objects elicited significant theta changes relative to SA objects in two brain regions, the right vmPFC and right STG. Furthermore, the vmPFC drove STG theta activity during the imagination of SD objects relative to SA objects.

SD and SA objects were matched in terms of the vividness of mental imagery, oculomotor behaviour and incidental memory encoding. Importantly, all objects were incorporated into the same three-object scene structures within which the order of SD or SA object presentation and object locations were carefully controlled. I also examined object permanence, and found that this property did not engage the vmPFC or STG. My findings are therefore unlikely to be explained by these factors.

Interestingly, most of the brain regions typically associated with scenes did not respond to SD objects during scene construction. This is probably because scene processing was constant throughout the experiment, and so there was no variation required in the activity of these areas. It is particularly notable that the PHC, which was active during fMRI in response to SD objects when they were viewed or imagined in isolation and devoid of a scene context (Mullally and Maguire, 2011), did not exhibit power changes during scene construction. It may be that examining objects in isolation afforded a “purer” expression of SD, whereas, once these objects were included in scene building, higher-order areas then came online to direct their use in constructing scene representations, a possibility that I discuss next.

Considering first the involvement of the STG, while this region has been linked to speech processing (e.g. Hullett et al., 2016), the close matching of auditory stimuli and the absence of activity changes in other auditory areas suggests this factor does not account for its responsivity to SD objects. Perhaps more germane is the location of the STG within the anterior temporal lobe, a key neural substrate of semantic and conceptual knowledge that supports object recognition (Peelen and Caramazza, 2012; Chiou and Lambon Ralph, 2016). Patients with semantic dementia caused by atrophy to the anterior temporal lobe, lose conceptual but not perceptual knowledge about common objects (Campo et al., 2013; Guo et al., 2013).

This could mean that SD objects provide conceptual information that is registered by the STG. Why might this be relevant to scene construction? Prior expectations have a striking top-down modulatory influence on our perception of the world, enabling us to process complex surroundings in an efficient manner (Bar, 2003; Bar et al., 2006; Summerfield and Egner, 2009), and resolve ambiguity (Chiou and Lambon Ralph, 2016). Without this knowledge, we are unable to understand how and where an object should be used (Peelen and Caramazza, 2012). Therefore, objects are an important source of information about the category of scene being imagined (or viewed), facilitating a rapid, efficient interpretation of the scene “gist” without the need to process every component of a scene (Oliva and Torralba, 2006; Summerfield and Egner, 2009; Clarke and Tyler, 2015; Trapp and Bar, 2015). For example, if we see a park bench, this might indicate the scene is from a park. Although in the current study the scenes were deliberately composed of semantically unrelated objects, this may not have impeded the STG in nevertheless registering SD objects more so than SA objects because SD objects would normally offer useful conceptual information to help anchor a scene.

The operation of the STG might be facilitated by the vmPFC, which also showed a preferential response to SD objects in this study. Converging evidence across multiple studies has shown that the part of the vmPFC that was active in response to SD objects plays a role in the abstraction of key features across multiple episodes (Roy et al., 2012). These contribute to the formation of schemas, which are internal models of the world representing elements that likely exist in a prototypical scene, based on previous exposure to such scenes (Tse et al., 2007; van Kesteren et al., 2013; Gilboa and Marlatte, 2017). For instance, a park typically contains benches, trees and flowers. SD objects may be particularly useful in building scene schema, and hence the response to them by the vmPFC.

Patients with damage to the vmPFC exhibit deficits that suggest aberrant schema re-activation (Ciaramelli et al., 2006; Gilboa et al., 2006; Ghosh et al., 2014), and this has led to the proposal that vmPFC may activate relevant schema to orchestrate the mental construction of scenes performed elsewhere – for example, in the hippocampus (McCormick et al., 2018a; Ciaramelli et al., 2019). My previous experiment (Chapter 3) lends further support to this proposal regarding hippocampal-vmPFC dynamics. In the current study, my DCM findings extend this work by showing that the vmPFC also exerts influence over the STG, indicating it may be engaging in top-down modulation of conceptual object processing by the STG, specifically during the processing of SD objects when couched in scene representations. My results may therefore indicate that SD objects help to define a scene by priming relevant schemas in the vmPFC which then guide conceptual processing in areas such as the STG. This may speak to the simple model of scene (and event) construction outlined at the start of this thesis (Section 1.4; McCormick

et al., 2018a), by providing some insights into vmPFC-neocortical interactions that then feed the hippocampus with the elements that are relevant for building a scene.

There is another possible explanation for my findings. In the current study, the scenes were deliberately composed of semantically unrelated objects, and this could have introduced ambiguity about a scene's identity. vmPFC and STG engagement may therefore be evidence of additional neural processing that was required to resolve incongruences inherent to acontextual scenes (Chiou and Lambon Ralph, 2016; Brandman and Peelen, 2017; Epstein and Baker, 2019), perhaps by drawing upon existing schemas in the pursuit of an appropriate scene template. Indeed, connectivity between medial prefrontal and medial temporal cortices has been shown to increase when processing novel information that was less congruent with pre-existing schematic representations (van Kesteren et al., 2010a; Chiou and Lambon Ralph, 2016).

The acontextual nature of the scenes, and the effortful nature of the scene construction task, may also have precluded observation of a schema-related memory advantage for SD objects (see more on schema and memory in Tse et al., 2007; Gilboa and Marlatte, 2017; McCormick et al., 2018a). It should be noted, however, that the current study was not designed to investigate schema, and consequently these possible interpretations remain speculative. Future studies will be needed to further elucidate the SD/SA difference revealed here, perhaps by comparing semantically related and unrelated objects during scene construction, and by adapting the current paradigm to test patients with vmPFC or STG damage. In addition, testing with a longer retention interval, and examining memory for the constructed scenes and not just single objects, may reveal differential SD/SA results. Another notable feature of my findings was the right hemisphere location of the responses. This may be related to the visual nature of the

imagined scene, and could also be probed further in future studies by comparing patients with left and right-sided lesions. Finally, in a future study it may be possible to find some way to examine the order of SD and SA objects during scene construction and whether this has an influence on the brain areas engaged and their connectivity.

In conclusion, this study revealed the neural dynamics associated with a specific object property during scene construction, and I suggest that SD objects in particular may serve to activate schematic and conceptual knowledge in vmPFC and STG upon which scene representations are then built. In the next chapter, I consider how individual scenes might be prepared for inclusion into events by invoking the phenomenon of BE.

5 Preparing scenes for inclusion into events

5.1 Précis

The experiments described so far have focussed on the imagination of single scenes. In Chapter 3, I observed early responses in the anterior hippocampus and vmPFC during the construction of scene imagery, with the vmPFC driving this scene-specific hippocampal activity. In Chapter 4, I showed that SD objects, which seem to define their surrounding space, also preferentially engaged the vmPFC, which drove activity in the STG during scene construction. In both experiments, the vmPFC influenced activity in downstream brain areas, concordant with the account of the vmPFC initiating and orchestrating the construction of consciously generated scene imagery. How the brain supports the transition from one single scene to another is the focus of this next experiment. I investigated whether the hippocampus and vmPFC, along with other brain regions typically engaged by scene processing, were engaged by the boundary extension (BE) effect, a well-documented implicit form of scene construction that I described in Chapter 1. This concerns the tendency we have to automatically and rapidly extrapolate outward beyond the edges of the scene currently in view. I contend that BE may contribute to preparing individual scenes for inclusion into evolving events. BE has not been probed previously using MEG, a technique that has the potential to reveal the finer temporal unfolding of this process.

5.2 Introduction

An important feature of perception is that our visual input is necessarily constrained. It is not possible to infinitely sample our surroundings such that every aspect of the scene in front of us is “complete”. Instead, we frequently blink and our view of parts of the scene may be obstructed. These limitations do not, however, interfere with our sense of continuous perception of the scene as coherent and “whole”. Pertinent to my thesis, this completion process may explain how our ongoing mental representation of an event feels seamless, perhaps by integrating successive scenes. In the current experiment I sought to address how this might be achieved.

We know that the brain is able to make predictions about what is likely to exist beyond our view (Friston, 2010). This capacity is crucial for adaptive decision-making in humans, and is exemplified by BE, a cognitive phenomenon outlined and illustrated in Chapter 1 (see Figure 7). People tend to recall a larger extent of a scene than was actually viewed, thought to result from an anticipatory form of scene construction whereby one implicitly constructs a mental representation of the scene that extends beyond its visible borders (Intraub and Richardson, 1989). BE is a robust effect, persisting even when the observer is made aware of the phenomenon (Intraub and Bodamer, 1993), suggesting it is not completely under conscious control. This effect also specifically occurs in response to images of scenes, and not isolated objects (Gottesman and Intraub, 2002).

In Chapter 1 I described the two phases thought to comprise BE, which I briefly recapitulate here. When a close-up scene is initially perceived (phase 1), in healthy people, automatic extrapolation beyond its borders occurs which has the consequence of creating a mental representation of the scene wider than the picture depicted (Intraub and

Richardson, 1989; Intraub et al., 1992, 1996; Seamon et al., 2002; Chadwick et al., 2013; Czigler et al., 2013; Spanò et al., 2017). It is during the viewing of this initial study picture that the BE effect occurs, evidence for which is observed at the subsequent test phase in the form of a memory error (phase 2). Thus, when the same scene is presented again after a brief delay, it is mistakenly perceived as being a closer-up view compared to the original. BE can be observed when the two scenes presented are identical (Mullally et al., 2012b; Chadwick et al., 2013; De Luca et al., 2018), as was used in the current study, and when they are mismatched close-to-wide angle views (see Intraub and Richardson, 1989; Park et al., 2007; Hubbard et al., 2010). Since BE has been observed even with a delay as brief as 42 ms between the two scene presentations – equivalent to a saccade (Dickinson and Intraub, 2008; Intraub and Dickinson, 2008) – this also suggests that it is a very rapid process, consistent with the notion of it being implicit.

The BE memory error measured at the second phase provides a diagnostic measure of BE having occurred earlier, or not – the picture is rated as “the same” (the correct response in the case of the current experiment), “closer-up” or “farther away” (both errors). The initial extrapolation of scenes during the original scene picture can then be examined by comparing neural activity categorised according to a participant’s judgement at test. To date, only one fMRI study has examined this first phase, and in this manner. Using a brief presentation of 250 ms, Chadwick et al. (2013) found the hippocampus and PHC were both involved in the BE effect. Interestingly, they further observed that the hippocampus drove early visual cortices during BE, not the reverse, suggesting this structure in the MTL plays a key role in implicit scene construction and not only in its explicit form (Hassabis et al., 2007a; Epstein, 2008; Dalton et al., 2018; Barry et al., 2019a, 2019b; Clark et al., 2019; Epstein and Baker, 2019). To note, in another

fMRI study, Park et al. (2007) studied the neural correlates of the retrieval phase of BE in healthy controls, where the PHC and RSC both emerged as regions sensitive to the BE error. They did not report on the study phase, which is of key interest here.

Further evidence of the necessity of the hippocampus for scene extrapolation was provided by patients with selective bilateral hippocampal damage, who demonstrated attenuated BE relative to healthy controls (Mullally et al., 2012b). Interestingly, patients with damage to the vmPFC – another region associated with scene construction, as I've outlined previously (Zeidman et al., 2015b; Barry et al., 2019b, 2019a; McCormick et al., 2020) – showed a level of BE attenuation comparable to that of hippocampal-damaged patients (De Luca et al., 2018). This suggests that the hippocampus and vmPFC both play roles in the implicit construction of mental scene imagery.

Together, the extant, albeit limited, neuroscientific literature on BE suggests that areas within the MTL, including the hippocampus and PHC, as well as the vmPFC are involved in implicit scene construction. In the current experiment, I examined the involvement of these regions at the study phase (the BE effect). This is because it is at this point in time we can glean the most information about how the brain reaches beyond the current view, potentially laying the groundwork for linking one scene snapshot to the next, as a precursor to representing events. As in Chadwick et al.'s (2013) fMRI study, I did not examine neural activity during the retrieval phase (the BE error) as it is likely confounded by memory-related and preparatory processing as the participant anticipates making a response. The first extrapolation BE phase has not been previously investigated electrophysiologically. To the best of my knowledge, there has been only one EEG study on BE, providing a finer temporal resolution than was possible with fMRI, but which examined only the second phase – the BE error (Czigler et al., 2013). The early event-

related potential (ERP) observed within the first 285 ms of the test picture suggested that the extrapolated region of the scene committed to memory was made rapidly available when assessing the scene a second time.

Here, I used MEG to measure the neural activity during the first phase, where the BE effect occurs. This temporally-resolved technique also affords sufficient spatial resolution for localising any ERF effect observed. I presented the study picture for a brief duration (250 ms), since this paradigm has been previously shown to elicit BE (Intraub et al., 1996, 2006; Intraub and Dickinson, 2008; Chadwick et al., 2013), and a rapid presentation is well-suited to MEG since neural processing can be tracked with millisecond precision. In this way, a simple comparison could be performed between the evoked responses to those scenes when BE occurred, and the scenes when it did not.

It is notable that ERF studies normally have a large number of trials which are thought to optimise the chances of detecting an evoked response. In the BE literature, however, it is usual to test only a small number of stimuli (normally fewer than 40). This is because increasing the duration of the testing session with a larger number of scenes weakens the BE effect and introduces other memory errors that can interact with BE over time (see Intraub et al., 1992; Intraub, 2020). Therefore, in line with the previous neuroimaging investigation of the BE effect (Chadwick et al., 2013), here I used a total of 60 trials which provided the best balance between trial number and eliciting BE.

I predicted that an effect of BE would be detectable early on in the evoked response to the study picture, consistent with an automatic process. How early within the 250 ms remained, however, an open question, since there is a lack of extant electrophysiological evidence to guide a more specific hypothesis. I subsequently

performed source-level beamforming to localise any ERF effect detected at the sensor level. Given that all stimuli during the study phase were scenes, it was highly likely that the hippocampus and vmPFC would be engaged. I further reasoned that extrapolation beyond the view of scenes when BE occurred would be associated with increased scene construction and so increased neural activity in these two brain regions relative to scenes where BE did not occur. Such a finding would also align with the attenuated BE observed in patients with hippocampal or vmPFC damage (Mullally et al., 2012b; De Luca et al., 2018). I also predicted involvement of other brain regions such as the PHC and RSC, as well as early visual cortices, considering these regions have been previously implicated in the BE effect and scene processing more generally (e.g. Park et al., 2007; Chadwick et al., 2013).

In a final part of the analysis, I used DCM to characterise the effective connectivity between regions revealed by the source reconstruction. Previous investigations, including my experiment in Chapter 3, of the effective connectivity of the vmPFC during scene construction and autobiographical memory have shown this region to exert an influence over the hippocampus (Barry et al., 2019b, 2019a; McCormick et al., 2020). I therefore hypothesised this frontal region would drive MTL regions associated with scene processing (hippocampus, PHC) during BE.

5.3 Materials and methods

5.3.1 Participants

Thirty six healthy right-handed participants (21 females; mean age 24.88 years, Std Dev = 3.40) took part in the experiment. In order to ensure a balanced analysis, I required a

similar number of trials where BE occurred and those where it did not – NoBE trials – rather than the more typical preponderance of BE trials (Intraub et al., 1992; Intraub, 2010). Previous research has shown that the longer the exposure to a study picture, the more likely BE is to occur (Intraub et al., 1996; Intraub and Dickinson, 2008; Mullally et al., 2012b; Chadwick et al., 2013; Czigler et al., 2013). Consequently, a short exposure such as the 250 ms typically used in the RSVP task, produces more variability in responses, which is why I used this paradigm. Sixteen of the participants (8 female; mean age 24.31 years, Std Dev = 4.25) had response proportions within my cut-off range (38-58%). Consequently, the eye tracking and MEG analyses described below were performed on the data of these participants.

5.3.2 Stimuli and task procedure

The task was a modified version of the RSVP task used previously in fMRI (Chadwick et al., 2013). Participants initially performed two practice trials in the MEG scanner using stimuli not included in the main experiment. Experimental stimuli consisted of photographs of scenes presented on a screen 60 cm away from a participant in the scanner. A participant was told that on each trial they would view a scene twice, and that on the second presentation the scene may appear closer-up, exactly the same, or farther away, compared to the first presentation. They were instructed to maintain their gaze on the centre of the screen throughout the experiment.

Each trial began with a central fixation cross for 3000 ms, immediately after which a scene was presented in the centre of the screen for 250 ms (the study picture; Figure 31). This was followed by a dynamically changing visual noise mask for 200 ms, and then a static visual noise mask presented for a variable period of between 2000 and 4000 ms.

The duration of this second masking was pseudorandomised across trials. This second presentation of the scene was in the same central location as the first (Figure 31; the test picture) and, unbeknownst to the participant, the scene on every trial was exactly the same scene as the initial study picture.

After 1000 ms, five options appeared under the test picture. Participants were asked to make a selection to complete the statement “compared to the original, this picture seems____” from the following: “much closer-up”, “a little closer-up”, “the same” (always the correct answer), “a little farther away”, or “much farther away” (Figure 31; lower panel). Up to 5000 ms was permitted to select a response using a five-button MEG-compatible button-box in their right hand. A second set of options then appeared for participants to make a confidence judgement regarding their decision: “not sure”, “fairly sure”, or “very sure”. In addition, they could indicate whether they did not recall seeing the first picture at all. This allowed for the possibility that a participant may have missed the study picture, given its earlier rapid presentation. However, this option was never selected by any participant. Up to 4000 ms was provided to make this response, followed by a rest period of 2000 ms before the start of the next trial. Each participant completed 60 trials, with a different scene on each trial, presented in a randomised order.

When a participant chose the “much closer-up” or “a little closer-up” options, this indicated that BE occurred at the earlier point of the study picture – these trials were therefore categorised as BE. BE did not occur when the “same”, “a little farther away”, or “much farther away” options were selected – these trials were therefore categorised as NoBE (Figure 31; lower panel). This categorisation allowed me to compare the presence and absence of BE when analysing the MEG data at the point of the study picture.

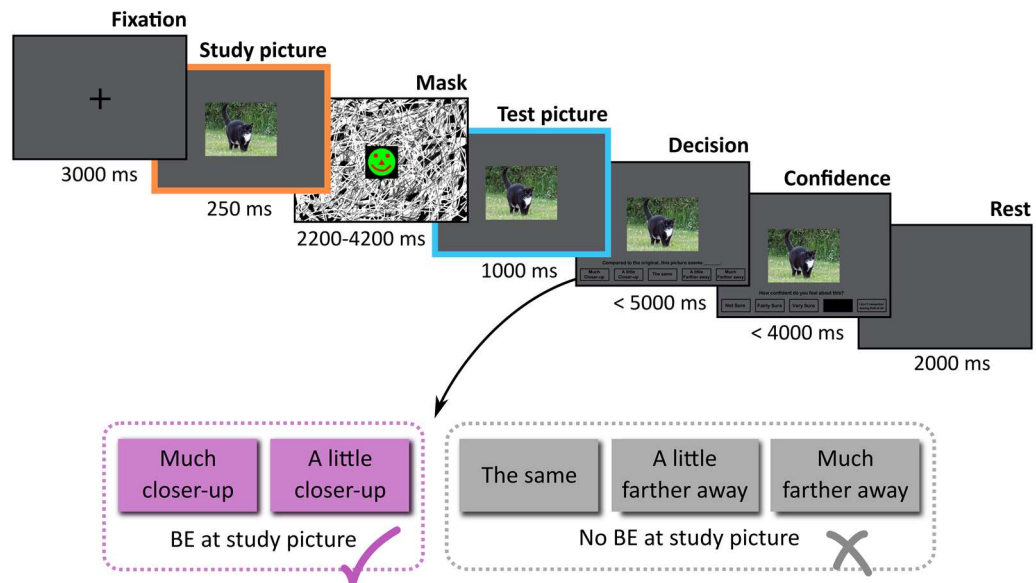


Figure 31. Trial timeline. The same scene picture is shown at study and test, separated by a masked delay period during which time the scene is no longer present. When we view a scene for the first time, we automatically extrapolate beyond the original view, called the “BE effect” (study picture; orange). When the scene is presented a second time, the extended mental representation is compared to the test picture, leading to the incorrect judgement that this second scene is “closer-up” than the first, even though they are identical, known as the “BE error” (test picture; blue). The five response options following the test picture are shown amplified below. Selection of either of the “closer” options was indicative of BE having occurred at the point of the study picture (pink), while “same” and “farther” options demonstrated BE was not experienced at the point of the study picture (grey).

The visit of each participant took approximately 50 minutes in total, including training, set up in the MEG scanner and eye tracker calibration, and the experimental task (~12 minutes).

5.3.3 In-scanner eye tracking

An Eyelink 1000 Plus (SR Research) eye tracking system with a sampling rate of 2000 Hz was used during MEG scanning to establish if participants maintained central fixation throughout the task. Eye tracking data were obtained for every participant.

5.3.4 Behavioural data analysis

Statistical analyses were performed using SPSS25. The number of responses in the BE and NoBE conditions were compared using a paired-samples t-test, using a significance threshold of $p < 0.05$. Eye tracking comparisons between BE and NoBE trials were performed using paired-samples t-tests to examine fixations across the duration of the 250 ms study picture, using a Bonferroni-corrected significance threshold. This was used to determine for how long participants fixated on the study image, as well as where (see Results, Section 5.4.1.2).

5.3.5 MEG data acquisition

MEG data were recorded using the whole-head 275 channel CTF Omega MEG system with a sampling rate of 1200 Hz, as described in Chapter 2.

5.3.6 MEG data pre-processing

Data were filtered with a 0.1 Hz high-pass filter, 48-52 Hz stop-band filter, and 98-102 Hz stop-band filter, to remove slow drifts in signals from the MEG sensors and power line interference. Epochs of interest were defined as -250 to 250ms relative to the onset of the study picture, corresponding to BE (closer ratings at test picture) and NoBE (same and farther ratings at test picture) trials. Since the time window subjected to analysis is narrow due to the nature of the effect under investigation (only 250 ms), it is difficult to resolve within a specific frequency band. Therefore, in order to best capture the BE effect, analyses involved a broad spectrum with a frequency cut-off of 90 Hz using a two-pass Butterworth filter with a filter order of 5. An interval of -100 to 0 before the onset

of the study picture was then used to perform baseline correction on BE and NoBE epochs prior to performing all analyses outlined below.

Following visual inspection of the data, an average of only 0.375 trials were rejected because they contained eye blinks and muscle artefacts. The robust average was calculated to obtain an ERF per participant and condition. This averaging method down-weights outliers when computing the average and helps to suppress high-frequency artefacts to minimise trial rejection; this is particularly helpful when only a small number of trials are available (as in this case; Wager et al., 2005).

5.3.7 MEG data analysis

5.3.7.1 Sensor level: ERFs

ERFs were analysed using the FieldTrip toolbox (Oostenveld et al., 2011), implemented in Matlab R2018a, on the robust averaged data per condition. To statistically compare the neural activity evoked in response to the study picture when BE did or did not occur, I used a non-parametric cluster-based permutation test, correcting for multiple comparisons across all MEG channels and time samples across the entire study picture window [0-250] (Maris and Oostenveld, 2007). This test controls for the Type I error rate by identifying clusters of significant differences over time and sensors, rather than performing separate tests for each sample of time and space. I report only effects that survived this correction (FWE, $p < 0.05$). Cluster-level statistics are the sum of t-values within each cluster, and this was calculated by taking the maximum cluster-level statistic (positive and negative separately), over 10,000 random permutations of the observed data. The obtained p-value represents the probability under the null hypothesis (no difference

between BE and NoBE) of observing a maximum greater or smaller than the observed cluster-level statistics.

5.3.7.2 *Source level: beamforming*

Source reconstruction was performed using the DAiSS toolbox (<https://github.com/SPM/DAiSS>) implemented in SPM12 (www.fil.ion.ucl.ac.uk/spm). The LCMV beamformer algorithm (Van Veen et al., 1997) was used to generate maps of power differences between BE and NoBE conditions following the calculation of the averaged evoked response. For each participant, the covariance matrix was estimated using a common spatial filter for all conditions, optimising the estimate and increasing the reliability of the filters through which all trials belonging to all conditions would be subsequently projected. For consistency, this was performed within the same broadband frequency spectrum as the ERF analysis. Due to the narrow time window of interest (< 250 ms), this spatial filter was computed across the entire epoch (-250 to 250 ms) including both pre- and post-response windows, as this has the advantage of a more stable covariance matrix. Whole brain power images per condition and per participant were subsequently generated within the shorter interval of interest identified by the ERF analysis (see Results, Section 5.4.2.1). Coregistration to MNI space was performed using a 5 mm volumetric grid and was based on nasion, left and right preauricular fiducials. The forward model was computed using a single-shell head model (Nolte, 2003). This resulted in one weight-normalised image per participant within the interval of interest for each condition, which was then smoothed using a 12 mm Gaussian kernel. At the second level I performed a t-contrast between BE and NoBE conditions.

5.3.7.3 *Effective connectivity*

Source localised differences in neural engagement between BE and NoBE provided the seed regions the effective connectivity analyses. I used DCM for event-related signals (Friston et al., 2003; Garrido et al., 2007; Kiebel et al., 2009) which was described in Section 2.7.3.2, but the key points are reprised here. This approach permits the comparison of biologically plausible models representing *a priori* hypotheses of the influence one region has over another, as well as mutual entrainment between regions (Friston, 2009; Kahan and Foltynie, 2013). Intrinsic connectivity between different cell populations within a region are estimated as part of every model. Extrinsic afferent inputs are categorised as forward, backward or lateral depending on which subpopulations these afferents project to (Felleman and Van Essen, 1991).

Random-effects BMS was then used to compare the evidence for competing models that varied according to which connections were modulated by the experimental condition (Stephan et al., 2009). For each condition I compared alternative unidirectional “master-slave” relationships between two regions, but also directly compared these models with bidirectional mutual entrainment, where neither region exerts a stronger influence over the other. When applied to my data, I determined the winning model to be the one with the greatest exceedance probability.

5.4 Results

5.4.1 Behavioural results

5.4.1.1 The BE effect

When a participant chose the “much closer-up” or “a little closer-up” options, this indicated that BE occurred at the earlier point of the study picture. BE did not occur (NoBE) when the “same”, “a little farther away”, or “much farther away” options were selected (Figure 31). The mean percentage of trials that elicited BE was 51.88% (Std Dev = 2.32), and NoBE was 48.12% (Std Dev = 2.32; comprised of 36.56% [Std Dev = 3.26] “same” responses, and 11.56% [Std Dev = 2.73] “farther” responses), and these were not significantly different ($t_{(15)} = 0.808, p = 0.432$; Figure 32A, left panel). There was therefore an even division between trials where the first presentation of scenes elicited BE and trials where it did not.

Confidence ratings for each condition also did not differ significantly (BE mean = 2.42, Std Dev = 0.34; NoBE mean = 2.28, Std Dev = 0.32; $t_{(15)} = 1.801, p = 0.092$). No participant selected the option “I don’t remember seeing that at all”, suggesting that all of the stimuli were processed. Note that there was no difference between BE and NoBE trials in terms of the length of jitter ($F_{(2, 30)} = 0.579, p = 0.567$).

I also examined whether there were any systematic differences across the stimuli in their propensity or not to elicit BE. I calculated the mean BE score for each scene image using the Intraub and Richardson (1989) BE scoring procedure, which scores each trial between -2 and 2, for “much closer-up” (-2), “a little closer-up” (-1), “the same” (0), “a little farther away” (1), and “much farther away” (2). As expected, there was a

significant BE effect across participants, with a negative mean BE score of -0.51 (SD = 0.21), reflecting a bias towards judging test pictures as “closer-up” and therefore indicative of the occurrence of BE at the study picture. This behavioural effect was highly significant ($t_{(15)} = -9.629, p < 0.0001$). Then, following an established rule for outliers (Chadwick et al., 2013; Intraub et al., 2015), I used a threshold of 3 Std Dev above or below the group mean to indicate whether a scene elicited a systematically strong or weak BE effect. This was not the case in my stimuli – the scene with the strongest BE effect was only 1.87 Std Dev from the mean, while the one with the weakest was only 2.42 Std Dev from the mean. I also calculated the cross-participant Std Dev per scene and found considerable variation (mean Std Dev = 0.82, Std Dev of the mean Std Dev = 0.27, range of the Std Dev = 0.33–1.50), thus further indicating that specific stimuli were not driving the effects I observed.

5.4.1.2 *Eye tracking data*

My focus was on the study picture, and the comparison between trials where BE occurred and those where it did not, as determined by the response given at the later decision point. Examining the eye tracking data, aggregate fixation heat maps confirmed that participants maintained their focus on the same central area of the study pictures on both BE and NoBE trials during their 250 ms presentation (Figure 32A, right panel), with no instance of a participant looking elsewhere, including at scene borders. Moreover, there were no significant differences between the BE and NoBE conditions in terms of fixation duration (BE mean = 200.05 ms, Std Dev = 56.74; NoBE mean = 194.99 ms, Std Dev = 51.16; $t_{(15)} = 0.645, p = 0.529$).

5.4.2 MEG results

5.4.2.1 ERFs

I examined a broad frequency range (1-90 Hz) given the short study picture presentation time window (0 to 250 ms). The cluster-based permutation test revealed one spatiotemporal cluster where BE and NoBE trials differed significantly in evoked activity between 12.5-58 ms (with the maximum difference at 40 ms) from the onset of the study picture (Figure 32B, left panel). The topographic distribution of this very early difference in neural responses involved fronto-temporal sensors (Figure 32B, right panel). A second later difference (~140 ms) was not statistically significant.

5.4.2.2 Source space power differences between BE and NoBE

Having demonstrated a significant effect of BE at the sensor level, I subsequently performed source reconstruction of the ERF using beamforming. I found that BE relative to NoBE elicited power changes in a set of brain regions that are typically engaged by the processing of scenes and events (Figure 32C; Svoboda et al., 2006; Park et al., 2007; Spreng et al., 2009; Chadwick et al., 2013): the middle temporal gyrus (peak $x, y, z = -58, -32, -2$), RSC ($-6, -42, 8$), anterior hippocampus ($32, -14, -12$), visual association cortex ($-42, -72, 18$), PHC ($-14, -36, -12$), thalamus ($13, -12, -2$), posterior hippocampus ($-22, -38, 4$), superior parietal lobule ($20, -72, 56$), and vmPFC ($-4, 58, -12$). No changes in power were evident for the opposite contrast of NoBE relative to BE. To note, the power changes within the anterior and posterior hippocampus, PHC, RSC, thalamus, and visual association cortex survived more stringent thresholding at $p < 0.005$ (uncorrected).

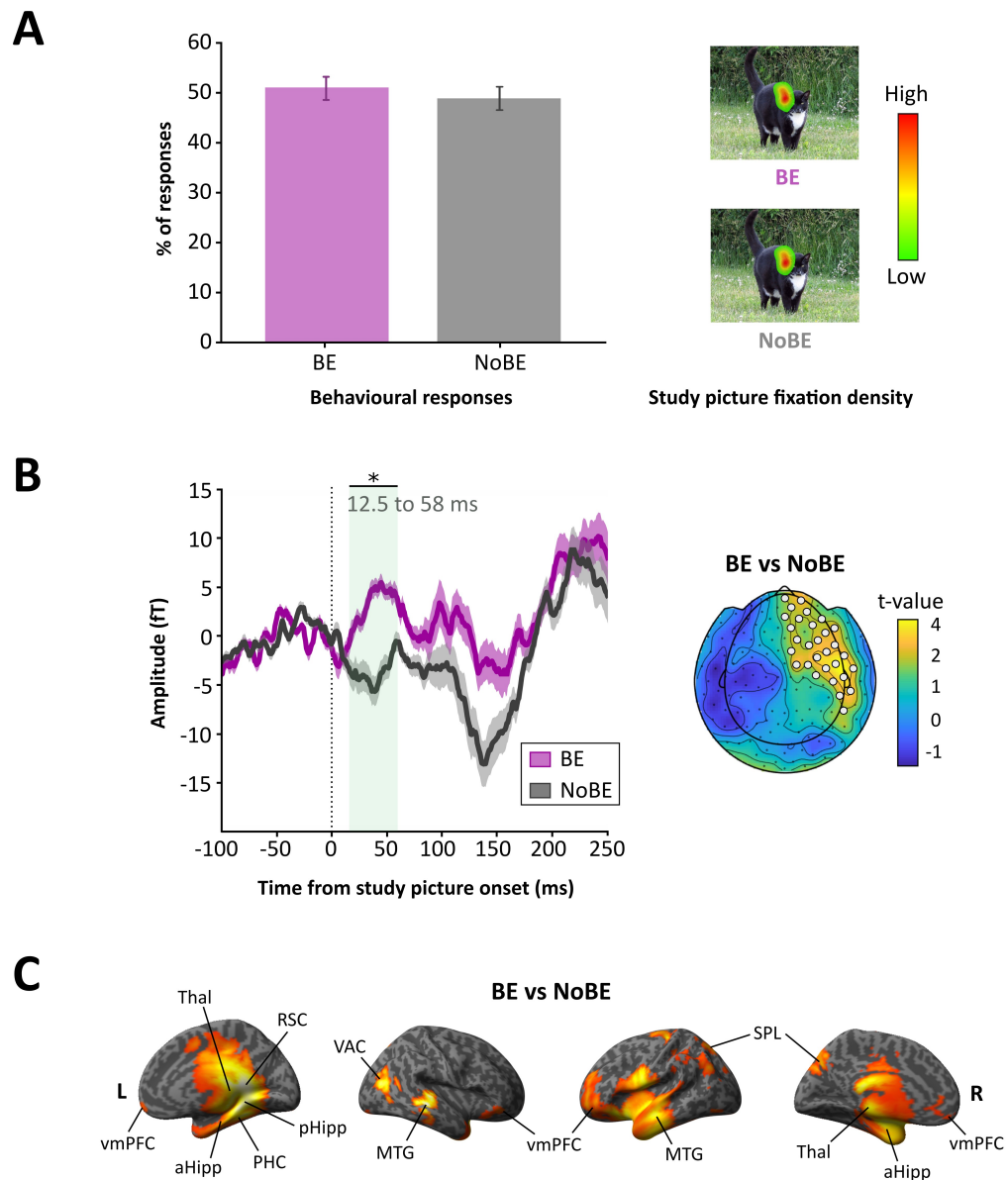


Figure 32. Behavioural and MEG results. (A) The percentage of trials where boundary extension (BE) did or did not (NoBE) occur (left panel); mean \pm SEM. Group eye fixation heat maps ($n = 16$) during the study picture (right panel) indicated no difference in where participants focused their gaze during BE and NoBE trials. Red indicates higher fixation density, and green lower fixation density. (B) The grand-averaged event-related field for the BE (pink) and NoBE (grey) conditions averaged over cluster of MEG sensors identified as significant by the permutation test (left panel). The line shading represents the SEM; the significant 12.5–58 ms time-window in which a statistically significant difference was observed is indicated by the light green box. White dots on the topographic representation of the data (right panel) mark the frontotemporal cluster of sensors at which this difference occurred, visualised using the CTF151 helmet layout. (C) Source reconstruction of evoked activity during in the 12.5–58 ms interval after the onset of the study picture, overlaid on a rendered inflated cortical surface, thresholded at $p < 0.05$ uncorrected for display purposes. L = left hemisphere, R = right hemisphere, vmPFC = ventromedial prefrontal cortex, aHipp = anterior hippocampus, pHipp = posterior hippocampus, PHC = parahippocampal cortex, MTG = middle temporal gyrus, Thal = thalamus, RSC = retrosplenial cortex, SPL = superior parietal lobule, VAC = visual association cortex.

5.4.2.3 *Effective connectivity*

Having established a difference between BE and NoBE in a number of brain regions, I next used DCM to assess the direction of information flow between the vmPFC and these other regions during the BE and NoBE conditions. I focussed the analyses on the left hemisphere, as the majority of peak differences were located there. For the two regions with primary peaks in the right hemisphere, I included secondary left hemisphere peaks (anterior hippocampus, -24, -14, -22; thalamus, -4 -25 11).

There are numerous ways in which these data could be analysed. My interest was in whether the rapid evoked response on BE trials was top-down driven. Consequently, I focused on examining the relationship of the vmPFC, which is typically regarded as a source of top-down modulation, with other regions revealed by the source reconstruction. This included brain areas often associated with scene processing (hippocampus, PHC, RSC), an important relay station (thalamus), and a potential source of bottom-up modulation (visual association cortex).

This wide set of brain regions presents a challenge for this DCM analysis due to the large number of possible models that could be evaluated if all brain regions were entered into the same model space. Since this would involve an unmanageable number of possible models, it is best to reduce this number by having strong hypotheses about which regions have causal influence. This is easier when examining well-established sensory and motor systems, however, the causal relationships of each region in this study are not precisely known. Whilst new approaches now exist that enable a search over the entire model space (see Friston et al., 2011, 2016; Moran et al., 2013; Razi et al., 2017), it is unlikely these could sufficiently reduce the complexity of the model space involving

five regions of interest, and they have not yet been widely applied to MEG data. As a result of these challenges, interpreting any results would be highly complex.

This situation calls for a simpler approach when specifying models. As a compromise, I examined pairwise connectivity of the five regions. For each brain region I tested three models of information flow: (1) a region drove vmPFC, (2) vmPFC drove a region, and (3) the two regions were mutually entrained. Due to the short time interval within which a significant effect of BE was localised (12.5–58 ms), I fit models to a window of 1 to 60 ms of the study picture, enabling better modelling of the evoked response observed. This was implemented for both BE and NoBE conditions. With this approach I could assess whether the vmPFC drove activity in other regions during BE, and whether this connectivity profile was specific to only trials where BE occurred.

BE trials

During BE trials, the vmPFC was more likely to be the driving influence: the EP for vmPFC driving anterior hippocampus was 62.32%, for posterior hippocampus it was 71.65%, PHC 88.33% (Figure 33), RSC 71.65%, thalamus 76.44%, and visual association cortex 45.24%. In contrast, the probability of the vmPFC being driven by each of these regions was low: anterior hippocampus-driven was 1.62%; posterior hippocampus-driven 11%; PHC-driven 1.03%; RSC-driven 22.6%; thalamus-driven 7.83%, and visual association cortex driven 1.62%.

NoBE trials

During NoBE trials, the vmPFC was also driving the RSC, thalamus and visual association cortex (61.83%, 91.55%, and 97.4% respectively, compared to the alternative unidirectional models: RSC-driven 17.41%; thalamus-driven 1.1%; visual association

cortex-driven 0.83%). For vmPFC connectivity with the anterior and posterior hippocampus, and the PHC, there was no clear winning model.

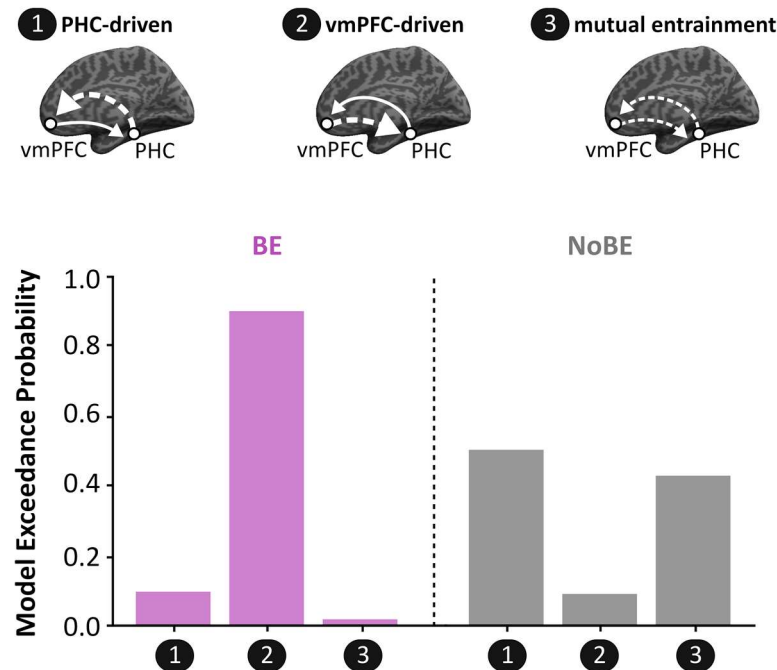


Figure 33. An example DCM. Three different models of connectivity between the ventromedial prefrontal cortex (vmPFC) and, in this example, the parahippocampal cortex (PHC) were tested: (1) the PHC driving the vmPFC, (2) the vmPFC driving the PHC, and (3) mutual entrainment. Bayesian model selection revealed the model with the strongest evidence during boundary extension (BE; pink) trials was the vmPFC-driven model. During NoBE trials (grey) there was poor evidence for any one model.

5.5 Discussion

This study examined BE, an implicit scene construction process whereby the view of a scene is remembered as containing more content than was physically present, reflecting the tendency to extrapolate beyond the scene's borders during its first viewing (Intraub and Richardson, 1989). To the best of my knowledge, this experiment represents the first time that MEG has been used to examine the neural dynamics of the BE effect in a time-resolved way. By tracking the BE effect millisecond-by-millisecond, this experiment

aimed to provide a window into how single scenes might be prepared for inclusion into unfolding mental events.

I found that when first viewed, scenes were processed differently depending on whether or not BE occurred, with BE evoking early ERFs within the first 58 ms of viewing the study picture. This BE effect was associated with power changes relative to NoBE in a set of brain regions that are typically engaged by scene and event processing, including the vmPFC, hippocampus, PHC, RSC, thalamus and visual association cortex. Given that all of the study pictures were scenes, I suggest that these power changes are related to the additional implicit scene construction that occurred during BE trials, over and above the general scene processing that was also present during NoBE trials.

Notably, analysis of the flow of information between the vmPFC and the other brain regions revealed a propensity for the vmPFC to exert top-down influence across the board. Interestingly, even on trials where BE was absent, the vmPFC drove the RSC, thalamus and visual association cortex during scene perception. However, effective connectivity between the vmPFC and hippocampus, and the vmPFC and PHC, was only causal (with vmPFC driving) during BE trials. These findings align with the few previous neuroscientific studies of the BE effect, which documented involvement of the hippocampus (Mullally et al., 2012b; Chadwick et al., 2013), the PHC (Chadwick et al., 2013), and the vmPFC (De Luca et al., 2018).

By supporting the ability to imagine what is beyond the view of a single scene implicitly and very rapidly, the vmPFC and MTL areas, and the associated additive scene construction, may be crucial for underpinning the subsequent linking of individual scenes into seamless events. For example, when we imagine what is beyond the view in the real

world, this may provide overlap in mental imagery from scene to scene that then facilitates our sense of continuity by smoothing over the transitions between scenes. Moreover, by predicting what might be beyond the view, this can then be quickly verified or refuted when the scene changes, helping to potentiate action. This is an efficient and effective way to view the world, and so while BE gives rise to a memory error when tested in paradigms such as RSVP, in the real world it is highly advantageous.

My MEG findings extend our understanding of BE in showing how quickly this effect occurs, and the nature of the information flow that supports it. I will consider these two points in turn.

First, it was notable that the differentiation between BE and NoBE trials was detectable between 12.5-58 ms after onset of the study picture. Although I acknowledge that the filtering used will have caused a small degree of burring in time, pushing backward the effect by approximately 10 ms, this still leave an early effect within the first 60 ms of picture onset. It is now well-established that scene processing in humans can be rapid, with as little as 13 ms exposure needed to discern scene content and “gist” (Potter, 1975; Thorpe et al., 1996; VanRullen and Thorpe, 2001; Rousselet et al., 2005; Joubert et al., 2007; Castelhana and Henderson, 2008; Greene and Oliva, 2009; Potter et al., 2014). This speedy response is observed even when participants are engaged in a distractor task (Li et al., 2002) or when a scene is displayed in their peripheral vision (Thorpe et al., 2001).

Most studies in this area have used explicit tasks such as scene categorisation, relying on overt reaction times to provide a behavioural measure of visual processing (Oliva, 2005; Oliva and Torralba, 2006; Fabre-Thorpe, 2011). Typically, neural responses measured using MEG have been observed less than 200 ms after scene onset, with

simultaneous responses in the MTL and visual cortex (e.g. Ramkumar et al., 2016). However, debate surrounds exactly what is processed so early, since many different features are likely to contribute to judging scene category, from lower-level elements such as textures to higher-level conceptual features. In a recent elegant EEG study, Greene and Hansen (2020) found that high-level visual features could distinguish scenes only 80 ms after low-level features were processed. Interestingly, only high-level features were correlated with behavioural measures of scene categorisation (Greene and Hansen, 2020). Together, extant studies suggest that the basic meaning or semantic gist of a scene can be extracted quickly, even within a single fixation, without exhaustive processing of scene objects, perhaps based on global scene features (Li et al., 2007; Greene and Oliva, 2009; Larson et al., 2014).

However, my BE findings can now push our understanding further. The BE effect I examined was implicit, unlike previous scene processing tasks which were typically explicit, such as scene categorisation. Moreover, BE involves prediction about what might be beyond the view, perhaps based on schema, then sourcing additional elements of the scene congruent with what is physically present in the scene, and finally constructing this extra portion of a scene. This undoubtedly constitutes top-down processing and yet I have shown that it occurs below 60 ms from scene onset. This appears to imply that high level cognition may occur much faster than traditional accounts of visual processing suggest (e.g. Di Lollo et al., 2000; Lamme, 2006; Del Cul et al., 2007), consistent with recent evidence for conceptual processing of scenes being more rapid than previously believed (Bar et al., 2006; Potter et al., 2014). This involves areas such as the vmPFC and MTL engaging very early, with the former in particular being notable in driving activity in other brain regions, including visual areas, a point I consider next.

The effective connectivity results position the vmPFC as the source of top-down influence on other regions that include the hippocampus, PHC, RSC, thalamus, and visual association cortex. It may facilitate the rapid extrapolation of scene information by activating knowledge relevant to the schema for the scene physically present (e.g. a park schema). The importance of the vmPFC in initiating the retrieval of schema-relevant knowledge is illustrated in vmPFC-damaged patients with confabulation, who are unable to inhibit irrelevant memory representations (Moscovitch and Melo, 1997; Gilboa et al., 2006; Gilboa and Marlatte, 2017), and require explicit cues in order to retrieve scenes from a past experience (Kurczek et al., 2015). The vmPFC-driven processing underlying the BE effect indicates that this brain area not only plays a role in mental scene construction during explicit tasks, as shown in Chapters 3 and 4, but is also involved in a faster, implicit form of scene construction. This accords with the cognitive model of scene and event construction I described in Chapter 1. In that architecture, the vmPFC initiates the construction of a single scene by activating relevant elements in other cortical areas, which are then conveyed to the hippocampus to build the scene image. I can now add that the vmPFC also seems to orchestrate the anticipation of the next view. This is further consistent with evidence that the vmPFC engages in predictive processing more generally, such as anticipating utility and reward (Bechara et al., 2000; Iigaya et al., 2020), and foreseeing future episodes (Addis et al., 2009; Bertossi and Ciaramelli, 2016; Bertossi et al., 2017), serving to direct decision-making.

Is there an alternative explanation for the difference I found between the BE and NoBE conditions? By using a rapid scene presentation, the placement and duration of eye fixations were closely matched between scenes where BE occurred, and those where it did not. Participants fixated solely on the central portion of the study pictures in both

conditions (and not at the boundaries), with no difference in looking time (fixation duration). My findings are therefore highly unlikely to be explained by oculomotor behaviour or low-level processing. On no trial did any participant indicate that they did not see a stimulus, suggesting that they paid attention throughout this short (~12 minute in total) MEG experiment. Moreover, there was also no difference in participants' confidence in their decisions for BE and NoBE trials. If they had for some reason been less attentive during BE trials, this might have been reflected in the confidence ratings (or indeed, the eye tracking data). For differences in attention to explain our findings, loss of attention would have had to occur systematically and on approximately half of trials for each participant, which seems very unlikely. I also demonstrated there were no consistent item-level effects on BE within the set of scene stimuli that were viewed by all participants. Since no scenes were identified as eliciting a consistently strong (or weak) BE effect, it is unlikely that systematic stimulus biases were present and contributed to the effects observed.

To summarise, this study represents the first investigation of the BE effect source localised at the millisecond level. I found that high-level brain regions associated with scene processing were involved in rapidly extrapolating beyond the view in scenes, driven by the vmPFC. In this way, BE may represent an important anticipatory mechanism that starts the process of linking multiple single scenes to form an evolving mental episode. In the next experiment I examine directly how sets of individual scenes are combined to form unfolding events.

6 Building events

6.1 Précis

So far my experiments have focussed on single scenes, starting with how scene imagery is built from multiple objects and a 3D space (Chapters 3 and 4), and subsequently characterising the automatic extrapolation beyond the borders of individual scenes (Chapter 5), which may contribute to the representation of episodic events. My findings are consistent with the proposal that the hippocampus and vmPFC, among a distributed set of brain areas, are centrally involved in the construction of scene imagery. I provided further evidence that this neural recruitment may be particularly attuned to scenes, over and above other multi-element mental representations, such as arrays of objects and a 2D space (Chapter 3). In this final experiment, I examined directly how a stream of scenes may be linked to form an unfolding, coherent episode. My aim was to establish the neural dynamics pertaining to scene-based event representations, distinct from events built from non-scene imagery. To do this, I created brief movies composed of either a series of static scenes or a series of abstract patterns which, when sequentially presented, showed a temporally evolving event. Using a predominantly ERF approach, I examined the transition from a single scene to the event as it unfolded. I sought to extend findings from a previous MEG study of events in the context of autobiographical memory retrieval (McCormick et al., 2020) to examine, in my case, newly encountered events, the role of scene imagery, and the preferential engagement, or otherwise, of the hippocampus and vmPFC.

6.2 Introduction

Our lived experience comprises a series of ongoing contiguous events. This allows us to construct a coherent autobiography, imbuing our lives with a sense of continuity and a capacity to recall experiences even when the original events occurred many decades ago. How the brain achieves this monumental feat is a central question in neuroscience. As noted throughout this thesis, scene imagery appears to dominate a great deal of our mental lives. A set of brain regions typically engaged during episodic cognitive processes, including autobiographical memory, imagining the future, and navigation, is similarly active when perceiving and imagining single scenes alone (Hassabis et al., 2007a; Epstein, 2008; Zeidman et al., 2015b; Çukur et al., 2016; Dalton et al., 2018; Robin et al., 2018; Robin and Olsen, 2019). It has therefore been proposed that generating scene imagery is a key feature of episodic memory processing (Spreng et al., 2009; Maguire et al., 2015; Zeidman and Maguire, 2016; Clark et al., 2019, 2020), a topic I discussed in Chapter 1.

In previous experiments I examined how we might internally generate a single snapshot of our surroundings – which I suggest is a basic building block of an episode. I found that both the hippocampus and vmPFC were engaged at the first stage of scene construction, with vmPFC driving hippocampal activity, indicating that the scene is set at an early point. My results so far converge with and extend previous MEG findings of hippocampal-vmPFC dialogue underpinning scene imagery (Kaplan et al., 2017; Barry et al., 2019a). It also aligns with fMRI work highlighting the preferential nature of hippocampal engagement during scenes, over and above other mental representations composed of multiple associated elements (Summerfield et al., 2010; Dalton et al., 2018; McCormick et al., 2021). In this chapter, I turn my attention to how event processing itself is initiated, where transitions occur from a single scene to an episode that

subsequently unfolds over time. One aspect of this question is whether scenes are indeed special units of events, or whether the same neural processing underpins unfolding mental representations even when content comprises non-scenes.

Previous neuropsychological studies of patients with damage to either the hippocampus or vmPFC highlight the converging roles of these regions in scene and event processing. Impairments surrounding autobiographical memory are observed following lesions to both regions (Viskontas et al., 2000; Steinvorth et al., 2005; Rosenbaum et al., 2008, 2009; Kurczek et al., 2015; Bertossi et al., 2016a, 2016b, 2017; McCormick et al., 2018a). However, hippocampal-damaged patients also struggle to mentally construct and visualise single scenes, even though they can describe, from their semantic knowledge, appropriate elements that would be within a scene (Hassabis et al., 2007b; Maguire and Mullally, 2013; McCormick et al., 2018a). In contrast, the basic ability of vmPFC patients to generate scene imagery appears intact, as long as very specific cues are provided (Kurczek et al., 2015). They are unable, however, to go beyond a single scene to describe an unfolding event (Bertossi et al., 2016a), and display more general deficits involving the initiation of endogenous processing (for a review see McCormick et al., 2018a).

The hippocampus, therefore, seems fundamental in constructing scene imagery, damage to which prevents patients from building events from sets of scene images during encoding, recall, or future thinking. The vmPFC, instead, may have a more general organisational or executive role in providing the appropriate event structure into which scenes are constructed and updated. These observations are the basis of the simple scene and event construction model outlined in Chapter 1 (Section 1.4), placing scenes at the

centre of events, over and above other types of representation (McCormick et al., 2018a; Ciaramelli et al., 2019).

Other accounts of hippocampal function, however, argue that the formation of associations between disparate elements in space relies upon the hippocampus irrespective of whether these are scenes (Cohen and Eichenbaum, 1993; Lee et al., 2005; Lee et al., 2012; Eichenbaum, 2006; Aly et al., 2013; Erez et al., 2016). In addition to relationships within space, these accounts posit that the hippocampus also mediates the formation of associations across time, since memory for temporal sequences is impaired in patients with hippocampal damage (Konkel et al., 2008). It remains unclear whether scene processing in particular engages the hippocampus and vmPFC during event processing, or whether these two brain regions are engaged by all manner of multi-component unfolding representations.

Tracking the transition from single scenes to an unfolding event is difficult to achieve with fMRI due to the temporal lag of the BOLD signal. The finer temporal resolution of evoked MEG responses, combined with their reasonable spatial resolvability, offers a means to establish the neural dynamics of events that evolve millisecond-by-millisecond. Surprisingly, no such MEG study had been published. While there have been some previous electrophysiological investigations of event processing, these have largely focussed upon establishing inter-subject synchronisation of brain activity during movie viewing (Lankinen et al., 2014, 2018), and the phenomenon of “event segmentation” (Honey et al., 2012; Silva et al., 2019). The latter refers to the discretisation of continuous events into distinct episodes (Zacks et al., 2007; Kurby and Zacks, 2008), a process in which fMRI studies have implicated the hippocampus and a distributed set of neocortical regions (Ben-Yakov and Dudai, 2011; Ben-Yakov et al.,

2013; Baldassano et al., 2017; Chen et al., 2017; Ben-Yakov and Henson, 2018; for a review see Brunec et al., 2018). However, none of these studies, whilst using naturalistic movies, addressed how event processing might be initiated.

McCormick et al. (2020) examined this question in relation to the reconstruction of past complex, naturalistic personal life events, using MEG. Due to the temporal resolution of this approach, they could test whether the hippocampus or vmPFC orchestrated event recall. Responses in the vmPFC preceded and drove engagement of the hippocampus, both during the initiation and subsequent elaboration of event recall, suggesting that the vmPFC plays an important role in coordinating the reconstruction of autobiographical events. In the current study, I sought to extend this previous work by using MEG to temporally resolve the emergence of neural responses evoked by novel events as they unfolded moment-by-moment. I devised visually simplified line-drawn animations, depicting either events built from a series of temporally related scenes (e.g. a child throwing a paper aeroplane and chasing after it in a library), or events built from abstract non-scene patterns that nonetheless depict an evolving narrative (e.g. diamond shapes expanding, contracting, and rotating in different directions).

I compared these two event types at successive time points over the course of movies using an ERF approach. I predicted that there would be an early signal for scene-based events within the first few scene frames, distinguishable from more abstracted forms of evolving events. I hypothesised this response would neither be related to scene imagery nor to the linking of visual information *per se*, but rather would be elicited by the combination of linking scenes together, over and above the linking of non-scenes. Given previous evidence of preferential hippocampal activity during scene construction (Hassabis et al., 2007a; Zeidman et al., 2015b; Bertossi et al., 2016a; Dalton et al., 2018;

Barry et al., 2019a, 2019b), and autobiographical memory retrieval (McCormick et al., 2020), I predicted that the hippocampus would be more engaged during events built from scenes – typical of everyday, naturalistic events. Due to the close matching of the event stimuli being directly compared, whether a difference between events would be evident throughout the clips, beyond the early response predicted, remained an open question.

With regard to the vmPFC, extant evidence for the co-activation of the hippocampus and vmPFC during the recall of autobiographical events (e.g. Addis et al., 2004; Robin et al., 2015; Inman et al., 2018; McCormick et al., 2018b, 2020), and that the vmPFC may facilitate elaboration beyond discrete atemporal scenes (van Kesteren et al., 2010a; Bertossi et al., 2016a; Gilboa and Marlatte, 2017; Barry et al., 2018), suggests that the vmPFC may also be more engaged during scene-based events. Although this region was preferentially activated during the explicit construction of single scenes when compared to non-scene arrays (Chapter 3), it is less clear whether the vmPFC specifically supports scene-based event processing, more so than other evolving situational contexts that do not involve scene imagery. Given that both stimuli in the current experiment were events, and the dearth of knowledge concerning how events unfold millisecond-by-millisecond, I kept an open mind as to whether preferential vmPFC involvement would be observed for scene-based events.

In this experiment I implemented a broadband approach (1 to 30 Hz), since the focus of this experiment was evoked activity. This frequency range was also used to perform source level analyses. Although activity within the theta band (4-8Hz) is often pinpointed as the means for disparate brain regions to communicate during episodic memory processes (Colgin, 2013; Fuentemilla et al., 2014; Garrido et al., 2015; Kaplan et al., 2017), there is also evidence for the role of alpha (9-12Hz) and beta (13-30Hz) power

in episodic memory (Hanslmayr and Staudigl, 2014). With the specific functional contributions of different frequency bands rather unclear, and the nature of the current study being novel, I elected to not restrict analyses to a single, narrow frequency band.

6.3 Materials and methods

6.3.1 *Participants*

Twenty one healthy participants (11 females; mean age 23.42 years, Std Dev = 4.51) took part in the study.

6.3.2 *Stimuli*

Short visual movies were created by hand using the animation program Stykz 1.0.2 (<https://www.stykz.net>), each composed of 16 individually drawn image frames presented in a sequence. This allowed me to control the content of each frame of a movie clip. An image comprised a combination of straight lines and circles which created simple line imagery that was easily interpretable. Images were all greyscale to keep the luminance contrast low, and this was invariant across frames and between conditions. A pixelated grey background for all frames was created in the image manipulation software program GIMP 2.8 (<https://www.gimp.org>).

There were two main stimulus types (Figure 34A, upper two panels). In one, each image frame within a movie was a simple scene (Pictures), while in the other, each image frame was composed of abstract shapes (Patterns). In both cases, small changes in the stimuli image-to-image showed a progression of activity that connected every image frame, resulting in two conditions called Pictures-Linked and Patterns-Linked. In essence,

each movie resembled a flip-book animation. Pictures-Linked movies contained temporally related scenes unfolding over time such that a scene-based event was perceived. A stick-figure character in the centre of the image performed an activity with a single object that appeared in every frame of the movie. Each movie clip portrayed a different stick-figure paired with a different object with which they interacted. The environment for each movie was a simple representation of an indoor (50% of clips) or outdoor (50% of clips) scene. Other background objects were included to give a sense of perspective and movement. Patterns-Linked movies contained temporally unfolding patterns that matched the evolving nature of Pictures-Linked movies, where each showed a novel abstract shape which underwent numerous simple mechanical changes, such as a rotation in a particular direction. The result was that a non-scene event was perceived.

There was one-to-one matching between the stimuli of the two conditions. For each image frame in a Patterns-Linked movie, the number of pixels composing the central shape was matched with the number of pixels composing the stick-figure and its paired object from a corresponding Pictures-Linked movie. Pictures-Linked background images (minus the stick-figure and object) were scrambled to form the individual backgrounds of Patterns-Linked image frames. The number of frames it took for a particular pattern's movement to unfold (e.g. completion of one rotation) corresponded to the same number of frames it took for a stick-figure to accomplish a sub-activity (e.g. the stick-figure skateboarded over a ramp), so that the pace of sub-events was matched between conditions. An average of 4 sub-events occurred per Linked movie. There were ten unique movies for each stimulus type (see https://www.fil.ion.ucl.ac.uk/~amonk/event_videos/ for dynamic examples of the movies).

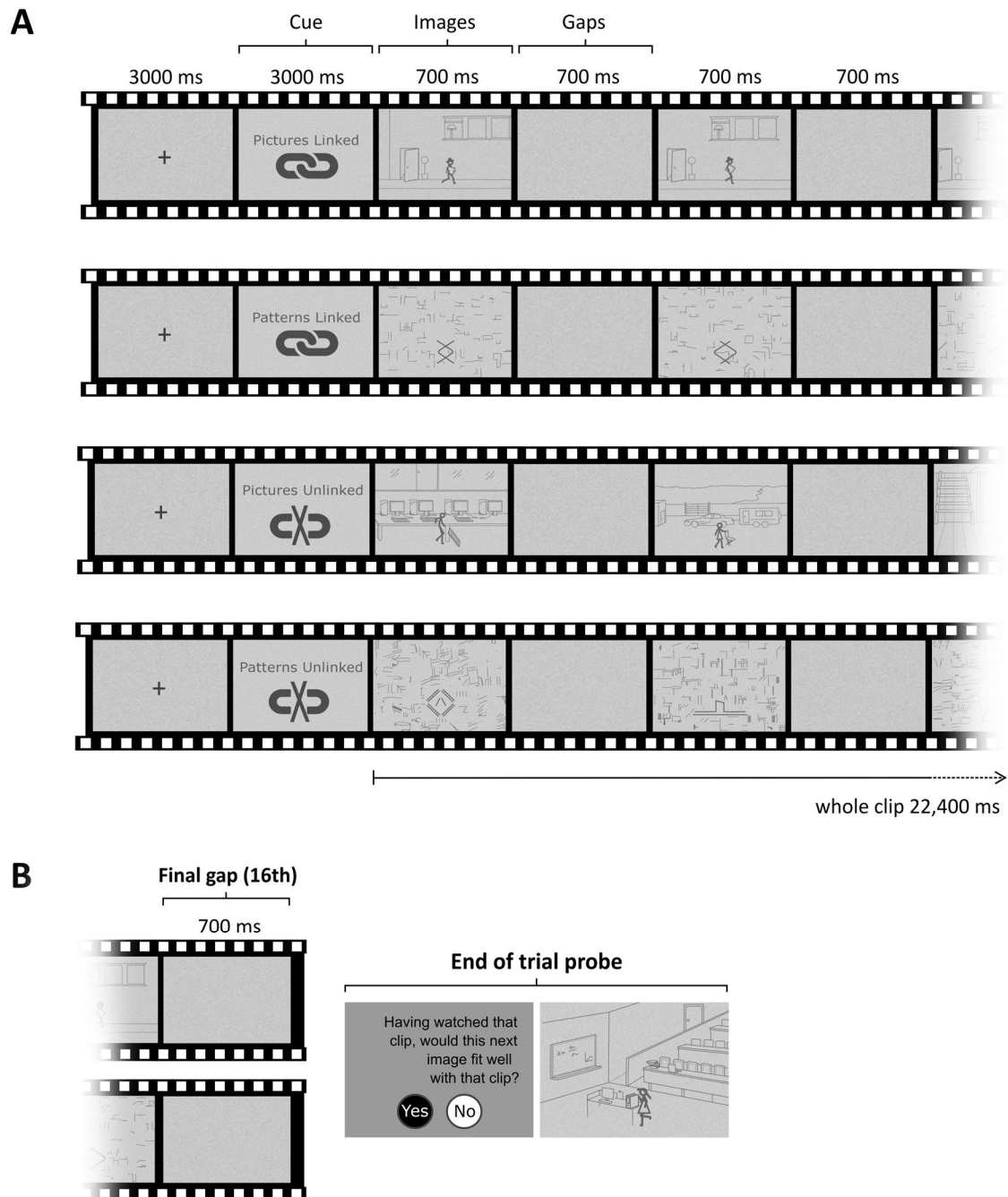


Figure 34. Experimental paradigm. (A) Schematic of the trial structure common to all movies, with examples of each condition. Each image frame was followed by a gap frame. (B) A probe question occasionally followed the completion of a clip to assess participants' engagement with the task.

There were two control conditions (Figure 34A, lower two panels), each with ten movie clips. Each movie was composed of a series of unique and separate unrelated image frames such that no evolving event could be conceived. Pictures-Unlinked movies contained separate scenes for each image frame, and in total there were 160 unique scenes, twenty different stick-figures and 152 unique objects. Patterns-Unlinked movies were composed of a series of unrelated abstract shapes. In total, there were 160 unique shapes. The same direct frame-to-frame matching procedure used for the Linked movies was applied to Unlinked movies in terms of the corresponding pixel count of central items and scrambled backgrounds of each image.

Every image frame was presented for 700 ms – a duration identified by piloting as long enough to comprehend the scene or pattern being viewed, and brief enough to minimise saccades and limit fixations to the centre of frames. Between each image, “gap” frames of the same duration were inserted, where no image was displayed and which consisted of only a pixilated grey background (see Figure 34). The pixilation served to mask the visual persistence of the preceding image. Since images were presented in a sequence, the primary function of gaps was to act as a temporal separator, so that individual images could be subjected to analysis independently. Gaps also ensured that images in Unlinked movies were clearly perceived as independent, and their inclusion in Linked movies ensured close matching. The 16 gaps matched the number of images in each movie clip, and each movie ended with a gap.

6.3.2.1 Pilot study

Careful pilot testing of a larger number of stimuli involved a separate group of participants ($n = 7$) performing the experiment, followed by a detailed debriefing session.

A key aim was to ensure that I only included in the main experiment those Patterns movies that were not interpreted as depicting real objects, scenes, or social events. I also confirmed that the gaps between images did not interrupt the naturalistic comprehension of Linked movies or their sense of unfolding, whilst adequately masking any visual persistence from the preceding image that might interfere with the perception of the next image.

Next, each individual movie was rated on: (1) Perceived Linking – how linked (or disconnected) images appeared to be; and (2) Thinking Ahead – how much of the time people found themselves thinking about what might happen next. A significant Friedman Test for Perceived Linking ($\chi^2_{(3)} = 18, p = 0.0004$) followed by Wilcoxon signed-rank tests found no significant difference between the two Linked conditions ($Z = -0.314, p = 0.753$), or between the two Unlinked conditions ($Z = -0.368, p = 0.713$). There was, as expected, a significant effect of Linking when comparing Pictures-Linked with Pictures-Unlinked ($Z = 2.371, p = 0.018$), and Patterns-Linked with Patterns-Unlinked ($Z = 2.366, p = 0.018$). Similarly, for perceived Thinking Ahead, a significant Friedman Test ($\chi^2_{(3)} = 17.735, p = 0.0005$) was followed by Wilcoxon signed-rank tests that revealed no significant difference between the two Linked conditions ($Z = -0.169, p = 0.866$), or between the two Unlinked conditions ($Z = -0.271, p = 0.786$). There was a significant effect of Thinking Ahead when comparing Pictures-Linked with Pictures-Unlinked ($Z = 2.371, p = 0.018$), and Patterns-Linked with Patterns-Unlinked ($Z = 2.366, p = 0.018$).

Participants were also asked to describe the event clips to ensure that each narrative was coherent and understood in a highly similar way across participants – e.g. in one clip a man delivers a gift to someone's house, participants broadly described this as “man exits a house and closes door...walks down the street...ties up the gift with

string...drops off the gift at a house.” Any inconsistencies alerted me to adjust the clip, improve its clarity, or not to include it in the final stimulus set.

6.3.3 Pre-scan training

Participants were trained prior to the MEG scan to ensure familiarity with the different conditions and the rate of presentation of movie frames. Specifically, participants were shown examples and told: “The movie clips are made up of a series of images. Some clips have images that are clearly linked to each other, and some have images that are not linked at all, so the images are completely unrelated to one another. The images can be either pictures with a stick-figure character, or abstract patterns.” Movies were preceded by one of four visual cues – Pictures-Linked, Patterns-Linked or, for the control conditions, Pictures-Unlinked and Patterns-Unlinked (Figure 34A). For Pictures-Linked movies, it was explained “...as you can see, there is a stick-figure character doing something. You’ll have noticed that the pictures are all linked together so that the clip tells a story.” For Patterns-Linked movies it was explained: “...for this type of clip, the patterns are all linked together, so one pattern leads to the next one in the clip. In this example clip the pattern moved outwards at first, then the crosses became larger, and then the circles increased in size, then the pattern changed again. The shape changed a bit step-by-step so that the clip portrays an evolving pattern.” Participants were instructed not to link the images in Unlinked movies and to treat each image frame separately when viewing them.

6.3.4 Task procedure

Each trial was preceded by a cue advising of the upcoming condition (e.g. Pictures-Linked) which was shown for 3000 ms. Each movie was 22400 ms in duration from the

appearance of the first image frame to the end of the final gap frame (Figure 34A). Individual image and gap frames were each 700 ms in duration. Participants then saw a fixation cross for 3000 ms before the next cue. To ensure participants attended to the movies throughout the scanning session, an occasional probe question was included (two trials per condition, Figure 34B). Following the final gap frame of a movie, a novel image was presented (either a Picture or a Pattern) and participants were asked whether this image fitted well with the movie clip they just saw. Of the two probe trials per condition, one was a ‘yes’ trial – that is, the image was congruent with the movie – and one was a ‘no’ trial – that is, the image was incongruent with the movie. The use of probe questions was also piloted, to ensure these were easy to understand. I chose this as a means to ensure participants maintained their attention throughout scanning blocks, without interfering with natural movie viewing.

Given the rate at which frames were presented, I sought to minimise a systematic relationship between spontaneous blinking and stimulus onset. Furthermore, fatigue is known to increase blink duration, which could result in participants missing individual frames, and increase the risk of significant head movement. Consequently, to ensure participants remained alert, the scanning session was split into five blocks each lasting approximately 6 minutes. During breaks between recordings, participants were instructed to blink and rest. Each recording block contained eight trials presented in a randomised order for each participant. They were instructed to maintain fixation in the centre of frames during the entire trial and to restrict eye movements to between-trial periods.

In summary, movies were visually very similar, with one-to-one matching between the two Linked and also the two Unlinked conditions. Common to all movies was the use of a central, item per image, the inclusion of interleaved gap frames, use of

simple line illustrations of scenes or patterns in greyscale, all of which were presented at the same frame rate of ~ 1.43 frames per second.

6.3.5 In-scanner eye tracking

An Eyelink 1000 Plus (SR Research) eye tracking system with a sampling rate of 2000 Hz to monitor task compliance and record eye tracking across the full screen. Some participants were excluded from eye tracking analyses due to insufficiently accurate calibration of the eye tracker during the scan, leaving 16 data sets for eye tracking analyses.

6.3.6 Post-scan surprise memory test

Following the experiment, participants completed a surprise free recall test for the event movies, since the principal aim was to examine the neural differences between Pictures-Linked and Patterns-Linked movies. Participants were asked to remember everything they could about what happened in each of these clips, unprompted. If they correctly recalled the simple story, they scored “1” for that clip, otherwise they scored “0”. The maximum score per participant and event condition was therefore 10 (as there were 10 clips per condition). Pictures-Linked and Patterns-Linked were compared using a paired-samples t-test.

The visit of each participant took approximately 2 hours in total, including training, set up in the MEG scanner and eye tracker calibration, and the experimental task (~ 30 minutes).

6.3.7 MEG data acquisition

MEG data were recorded using the whole-head 275 channel CTF Omega MEG system with a sampling rate of 1200 Hz, as described in Chapter 2.

6.3.8 MEG data pre-processing

MEG data were initially preprocessed using SPM12 (www.fil.ion.ucl.ac.uk/spm). Data were filtered with a 1 Hz high-pass filter, 48-52 Hz stop-band filter, and 98-102 Hz stop-band filter, to remove slow drifts in signals from the MEG sensors and power line interference. Epochs corresponding to the each movie cue were defined as -100 to 1000 ms relative to cue onset. Image frames were defined as -100 to 700 ms relative to image onset. Gap periods were defined as -100 to 700 ms relative to gap onset. Epochs were concatenated across trials for each condition, and across scanning sessions.

Prior to the calculation of ERFs, data were first low-pass filtered using a two-pass Butterworth filter, with a 6th order filter and frequency cut-off of 30 Hz. This filter causes a small degree of temporal blurring in either direction, although the effect is minimal. Following visual inspection of the data, an average of only 0.76 epochs were discarded on the basis of containing eye blinks and muscular artefacts. To baseline-correct, the activity during the first 1000 ms from the onset of the fixation period was averaged and subtracted from each cue, image or gap epoch (an explanation of this baseline-correction approach is given in Chapter 2, Section 2.6.2). The robust average was calculated to obtain an ERF per participant and condition. This averaging method down-weights outliers when computing the average and helps to suppress high-frequency artefacts to minimise

trial rejection, particularly helpful when only a small number of trials are available (Wager et al., 2005).

6.3.9 MEG data analysis

There are numerous ways in which this rich dataset could be analysed. The principal aim was to assess differences between the Pictures-Linked and Patterns-Linked conditions since my hypotheses concerned the construction of events built from scenes relative to those built from non-scenes. The other two non-event conditions therefore acted as control conditions. In order to make this key comparison, my focus was on the individual image frames that composed the movies. As previously mentioned, gaps were included in the design to provide temporal separation between images, so that brain activity associated with each movie image could be separately examined without interference or leakage from the previous image. When choosing an analytic approach, my main concern was to temporally resolve the dynamic unfolding of an event in real time in order to address my hypotheses. Looking across a wide time-window, such as the entire duration of the 22400 ms movie clips, or chunks over several seconds, might better facilitate the reconstruction of source space brain activity, and could increase the frequency resolution for a range of time-frequency analyses. However, this experiment was designed so that individual snapshots (image frames) of an event could be separately examined, permitting a more time-resolved approach. Therefore, to exploit the temporal resolution afforded by MEG to its fullest effect, I examined the data strategically on a time-point by time-point basis. Consequently, I explored both the evoked responses to particular images along the time course of the movies and then, as a secondary step, the likely sources of these responses. These steps are described below.

6.3.9.1 *Sensor level: ERFs*

ERFs were analysed using the FieldTrip toolbox (Oostenveld et al., 2011), implemented in Matlab R2018a, on the robust averaged data per condition. I sought to determine the timing and channel locations of statistically significant differences in evoked activity between Pictures-Linked compared to Patterns-Linked conditions in order to address the hypotheses concerning how events are built from scenes (Pictures) relative to those built from non-scenes (Patterns). I used a targeted sliding window approach to examine differences between conditions within potentially salient time-windows during movies – starting at the pre-movie cue, and then at key image frames, namely images 1, 2, 8, and 16. At image 1, only this first single image of a sequence was being viewed; at image 2, there was already the context set by the preceding first image; image 8 represented the mid-point of a sequence; and image 16 was the final image. This approach enabled me to sample across a long clip length, whilst also minimising multiple comparisons. I next performed a number of secondary contrasts involving the control conditions to ensure any differences observed between event conditions could not be explained by other factors.

Each pairwise comparison was performed using a non-parametric cluster-based permutation test, correcting for multiple comparisons across all MEG channels and time samples (Maris and Oostenveld, 2007), across the first 1000 ms of the cue and the entire 700 ms of images. This approach controls for the Type I error rate by identifying clusters of significant differences over time and sensors, rather than performing separate tests for each sample of time and space. I report only effects that survived this correction (FWE, $p < 0.05$). Cluster-level statistics are the sum of t-values within each cluster, and this was calculated by taking the maximum cluster-level statistic (positive and negative separately),

over 5000 random permutations of the observed data. The obtained p-value represents the probability under the null hypothesis (no difference between a pair of conditions) of observing a maximum greater or smaller than the observed cluster-level statistics.

I also examined a possible interaction involving all 4 conditions, using the same cluster-based permutation test, namely: (Pictures-Linked – Pictures-Unlinked) minus (Patterns-Linked – Patterns-Unlinked).

6.3.9.2 *Source level: beamforming*

Source reconstruction was performed using the DAiSS toolbox (<https://github.com/SPM/DAiSS>) implemented in SPM12 (www.fil.ion.ucl.ac.uk/spm). The LCMV beamformer algorithm (Van Veen et al., 1997) was used to generate maps of power differences between conditions of interest, as informed by the preceding ERF results. For each participant, the covariance matrix was estimated using a common spatial filter for all conditions. For consistency, this was performed within the same broadband frequency spectrum as the ERF analysis (1-30 Hz). Due to the narrow time window of interest (< 700 ms), this spatial filter was computed across the entire epoch (-700 to 700 ms) including both pre- and post-response windows, as this has the advantage of a more stable covariance matrix. Whole brain power images per condition and per participant were subsequently generated within only the shorter interval of interest identified by the ERF analysis. Coregistration to MNI space was performed using a 5 mm volumetric grid and was based on nasion, left and right preauricular fiducials. The forward model was computed using a single-shell head model (Nolte, 2003). This resulted in one weight-normalised image per participant within the interval of interest for each condition, which

were then smoothed using a 12 mm Gaussian kernel, and a t-contrast was performed at the second level.

6.3.9.3 Coherence

Coherence between any localised peak in either ROI (hippocampus, vmPFC) was performed using Dynamic Imaging of Coherent Sources (DICS; Gross et al., 2001) in the 1-30Hz frequency range, consistent with previous analyses steps, with a 3mm grid. This generated an image of sources coherent with the peak for each condition of interest. Images were smoothed using a 12mm Gaussian kernel and entered into a second-level paired t-test, thresholded at $p < 0.005$ uncorrected.

6.4 Results

6.4.1 Behavioural results

6.4.1.1 Probe trials

Participants correctly identified images as either being congruent or incongruent with the antecedent clip on an average of 93% of probe trials (Std Dev = 0.75), confirming they maintained their attention throughout the scanning session.

6.4.1.2 Eye tracking data

Eye tracking data were examined during the time-window where ERF analyses showed a significant difference between conditions. Fixation heat maps revealed the spatial pattern of fixations during image 1 (0-700ms; Figure 35) were highly similar across the conditions, and confirmed that participants maintained their focus on the centre of the images for

the whole duration of the image presentation. All four conditions were entered into a two-way repeated measures ANOVA to examine differences in fixations. No significant difference in fixation count was found between conditions during image 1 ($F_{(3,45)} = 0.535$, $p = 0.661$), and this included between Pictures-Linked and Patterns-Linked ($t_{(15)} = 0.141$, $p = 0.889$).

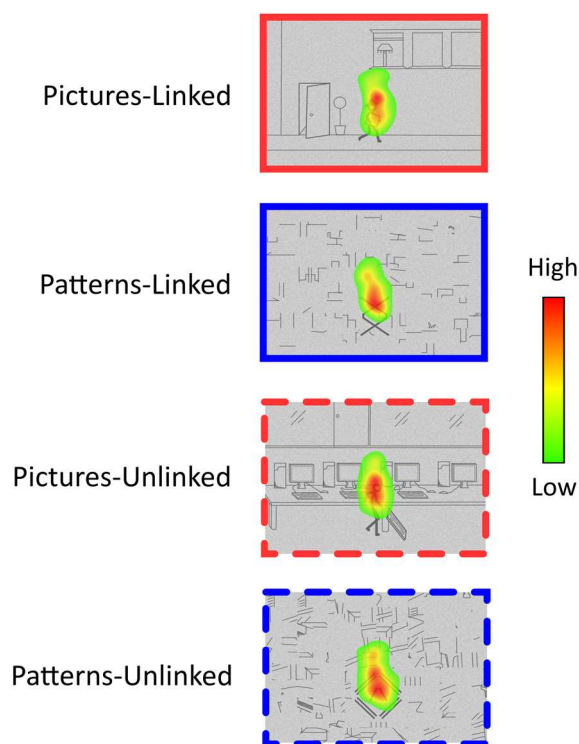


Figure 35. Eye fixations. Group eye fixation heat maps for each condition during the image 1 time-window ($n = 16$). Red in the heat map indicates higher fixation density, and green lower fixation density.

6.4.1.3 Memory for events

After scanning, participants engaged in a surprise free recall memory test for Pictures-Linked and Patterns-Linked movies. A paired samples t-test revealed no significant difference in recall (Pictures-Linked mean = 9.48, Std Dev = 1.21; Patterns-Linked mean

= 9.24, Std Dev = 0.89; $t_{(20)} = 0.7555$, $p = 0.459$). Participants, therefore, paid sufficient attention during both tasks to successfully encode stimuli, even though they were never instructed to memorise.

6.4.2 MEG results

6.4.2.1 ERFs

The primary focus was on differences between Pictures-Linked and Patterns-Linked conditions. I examined this contrast across all time windows of interest, from the cue preceding the movie, to the final image frame. I also examined the equivalent gap frames. The only significant difference I found was between Pictures-Linked and Patterns-Linked conditions at the very first image, involving one negative cluster emerging between 178-447 ms ($p = 0.0398$) and distributed across right fronto-temporal sensors (Figure 36A). This difference cannot be due to an effect of image type (i.e. Pictures) *per se*, as no difference was observed at image 1 between Pictures-Unlinked and Patterns-Unlinked ($p = 0.173$). In fact, it is clear from Figure 36A, that Pictures-Linked sits apart from the other three conditions, differing significantly from not only Patterns-Linked (as already noted), but also Patterns-Unlinked ($p = 0.013$), and approaches significance for Pictures-Unlinked ($p = 0.0678$). The other three conditions do not differ from one another: this includes Patterns-Linked and Patterns-Unlinked ($p = 0.3583$), and Patterns-Linked and Pictures-Unlinked ($p = 0.357$), showing that the effect found cannot be due to linking *per se*.

This pattern of results with Pictures-Linked driving the effect at image 1 would suggest there may be an interaction. When this was tested formally, I found there was indeed a significant interaction effect at image 1 ($p = 0.012$; time-window = 164-475 ms)

encompassing the same time-window within which Pictures-Linked and Patterns-Linked diverged (see Figure 36A).

Following these key comparisons, as a secondary ERF analysis I performed pairwise contrasts between the conditions across all the image epochs of interest (i.e. images 1, 2, 8, and 16). No differences were found.

In summary, the only difference found was for image 1 when Pictures-Linked was compared to the other conditions.

6.4.2.2 Source space power differences between Pictures-Linked and Patterns-Linked

I subsequently performed a beamformer analysis on the image 1 Pictures-Linked versus Patterns-Linked contrast, restricted to the same time window (178–447 ms) and frequency band (1–30 Hz) within which the significant difference in evoked responses was found. This analysis served to give a better indication of where in the brain this difference originated. Greater activation was principally found in the right hippocampus for Pictures-Linked relative to Patterns-Linked (peak $x, y, z = 32, -20, -16$; Figure 36B). Increases were also observed for this contrast in the right precuneus (12, -66, 56), visual association cortex (left peak = -22, -96, 10), left inferior frontal gyrus (-48, 26, 2), left premotor cortex (-48, -2, 38), and cerebellum (right peak = -30, -64, -58).

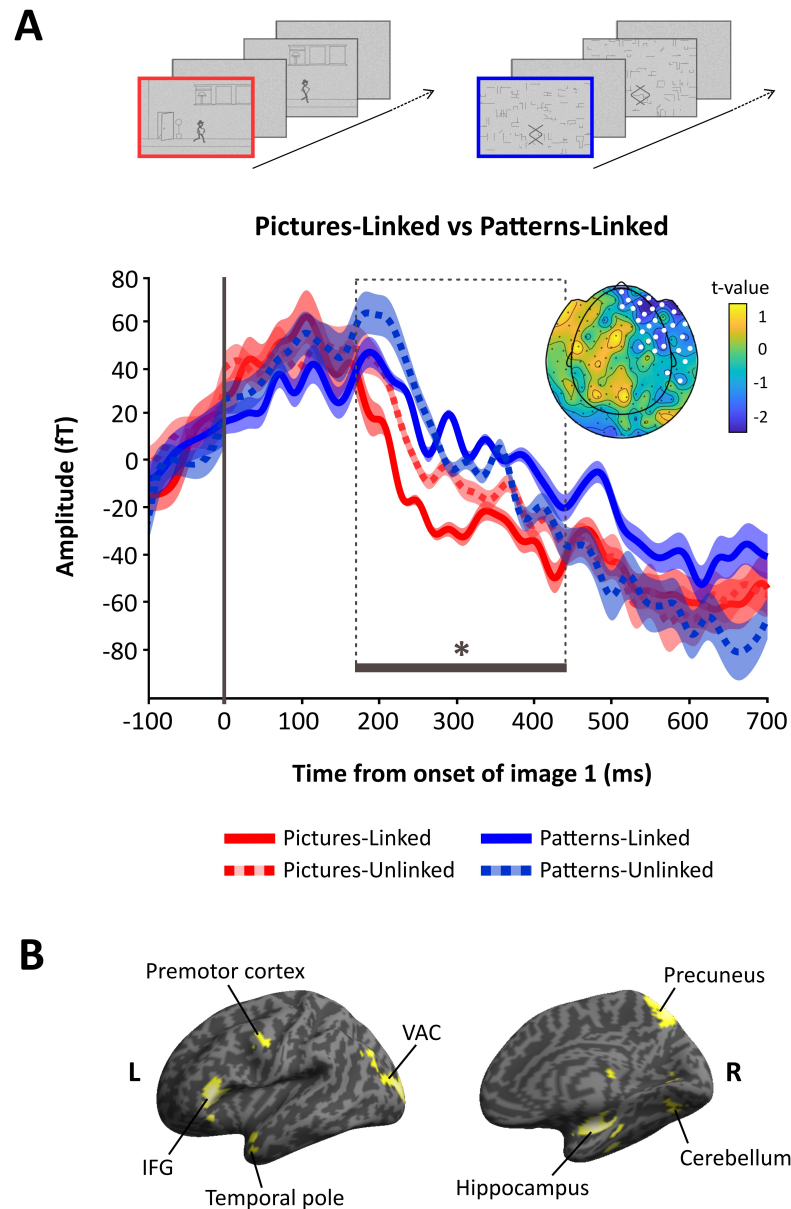


Figure 36. Image 1 evoked activity. (A) ERF result. A significant difference in the ERFs was found between Pictures-Linked (bold red line) and Patterns-Linked (bold blue line) for first image frames between 178-447 ms ($*p = 0.0398$), indicated by the dash line box. Displayed are the grand-averaged ERFs (shading indicates the SEM) for all four conditions, averaged over a right frontotemporal cluster (marked by white dots) within which the significant difference between Pictures-Linked and Patterns-Linked was observed. Displayed to the right of the ERF panel is the topographic distribution of the difference (t-values), displayed over the CTF151 helmet layout. Pictures-Unlinked (dash red line) and Patterns-Unlinked (dash blue line) did not significantly differ from one another, or from any other condition. **(B)** Source reconstruction of evoked activity at image 1 during the 178-447 ms interval, displayed on a rendered inflated cortical surface, thresholded at $p < 0.005$ uncorrected. L = left hemisphere, R = right hemisphere, IFG = inferior frontal gyrus, VAC = visual association cortex.

6.4.2.3 *Coherence*

Having established a difference in source level power during Pictures-Linked relative to Patterns-Linked at image 1 between 178–447 ms in the hippocampus, I explored whether this activity showed increased coherence with any other brain regions, but no effect was evident at the set threshold ($p < 0.005$).

6.5 Discussion

In this study I investigated the initiation and unfolding of events as they were experienced, comparing those built from successively linked scenes (Pictures-Linked) to those built from successively linked non-scene patterns (Patterns-Linked). This novel design allowed a close millisecond-by-millisecond examination of the transition from a single scene to an unfolding event. By using an ERF sliding window approach to the analysis, I strategically examined image frames across movies. One difference between scene and non-scene events emerged, within 178–447 ms of the onset of the first image frame across fronto-temporal sensors. Spatial analyses at the sensor level, however, provided limited information about the brain areas involved in generating this effect. This was therefore followed by source reconstruction of this difference, revealing that the largest change in broadband 1–30 Hz power was in the hippocampus during Pictures-Linked events.

Given that both conditions were events, I suggest that the observed Pictures-Linked effect characterises how events built from scenes are processed over and above general processing of unfolding dynamic imagery. Furthermore, for this effect to emerge at the very first image, I argue this reflects the onset of event construction. In fact, at image 1, the event is yet to unfold from this initial scene, so the neural dynamics captured

here may reflect the construction of a key building block destined for subsequent unfolding, suggesting events are established early on in processing.

A small number of MEG studies have previously investigated the neural correlates of visual event processing, which have implicated both the hippocampus and vmPFC during autobiographical memory recall (Fuentemilla 2014; Hebscher 2019; McCormick et al., 2020). In particular, the early hippocampal response I observed during the initiation of event construction is consistent with previous evidence from McCormick et al. (2020) for hippocampal involvement in the initiation of event recall. Interestingly I only found this difference in hippocampal involvement at the very start of clips. It is unclear why this was the case, but one might speculate that if the stage is set for the event within the first image, after this point each additional increment in information is relatively small. This perhaps serves to mirror our daily lives – as we perceive our surroundings, we do not notice large jumps from one view of a scene to the next, but rather, our experience is seamless. Since this gradual unfolding is a feature of both scene- and pattern-based events, this might explain the dissipation of the difference between these stimuli.

However, the early hippocampal response poses somewhat of a challenge to fMRI studies of movie viewing that found the hippocampus responded later, towards the offset of events (event boundaries), which authors have speculated reflects event replay, aiding memory consolidation (Ben-Yakov and Dudai, 2011; Baldassano et al., 2017; for a review see Griffiths and Fuentemilla, 2020). The conclusions that can be drawn from this work remain unclear, since fMRI is less suitable for resolving rapid perceptual processes than MEG or EEG. The few EEG studies that have examined episodic memory encoding using movies are limited in spatial resolution to confirm hippocampal

involvement at these boundaries (Silva et al., 2019). The use of MEG in the future might help to extend fMRI event boundary research to characterise more precisely the temporal dynamics of hippocampal activity, a point I consider further in Chapter 7.

Why might the hippocampus be particularly engaged during Pictures-Linked movies? It is notable that this condition was set apart from all others in the ERF results. Previous studies have demonstrated the prevalence and importance of scene imagery to performing hippocampal-dependent tasks, including episodic memory, imagining the future, and spatial navigation (Clark et al., 2019, 2020). The influence of scene imagery in events is unsurprising given how it mirrors the way in which we experience our surroundings – as scenes. Consideration of the other brain regions associated with Pictures-Linked movies – these were the posterior parietal, inferior frontal, premotor, and cerebellar cortices – may shed further light on why events built from scenes are treated differently. These areas have been identified in numerous studies as part of a network of regions that process biological motion and the anticipation of incoming intentional movement (Battelli et al., 2003; Rizzolatti and Craighero, 2004; Saygin et al., 2004; Fraiman et al., 2014). In particular, this has been observed in the context of point-light displays, in which a small number of moving lights (e.g. at the joints of a moving person) are sufficient to interpret this as behaviour (e.g. dancing). This draws parallels with the stimuli in my study. The Pictures-Linked events were highly simplified portrayals of activities depicted by stick-figures, lines and circles to create simple scenes. Although 2D drawings, they evoked 3D unfolding events of real-world activities that were easily grasped by participants.

Scene- and pattern-based evolving stimuli may have been processed differently because abstract patterns were not perceived as intentional, biological stimuli, while

participants could automatically infer the actions performed in scene-based events. Indeed, through careful piloting I sought to exclude patterns that consistently evoked the sense of biological motion. The success of this was reflected in the descriptions provided by participants in the post-scan memory test. For example, a Patterns-Linked movie showing three overlapping diamond shapes was described as “diamond shapes gradually expanded outwards, then rotated clockwise”, while Pictures-Linked movies were described in terms of the intentionality of the stick-figure.

What alternative explanations of the difference between Pictures-Linked and Patterns-Linked movies might lie at the heart of the hippocampal effect that I identified – is it perhaps due to scene processing, linking, or both?

It could be argued that the effect of Pictures-Linked was simply the result of scene processing. If this were true, a difference ought to be observed also between movies of unrelated scenes and unrelated patterns, since the hippocampus is known to respond strongly to scenes relative to non-scene stimuli (Hassabis et al., 2007a; Graham et al., 2010; Zeidman et al., 2015b, 2015a; Hodgetts et al., 2016; Zeidman and Maguire, 2016; Dalton et al., 2018; Barry et al., 2019a, 2019b). However, no difference was observed between Pictures-Unlinked and Patterns-Unlinked. This would suggest that image type alone cannot explain the observed effect in the context of my tasks. Another possibility is that the existence of a linking mechanism accounts for my data, irrespective of whether images in movies depicted scenes or not. However, despite both involving linking, I found that Pictures-Linked more so than Patterns-Linked was associated with a change in hippocampal power. In addition, no significant differences between any other pairs of conditions, including between Patterns-Linked and Patterns-Unlinked, and Patterns-Linked and Pictures-Unlinked were evident, suggesting the effect was not solely explained

by the linking of images, further evidenced by the image type by linking interaction that I observed. My findings, therefore, suggest that rather than being responsible for relational binding (Cohen and Eichenbaum, 1993; Eichenbaum, 2006; Konkel et al., 2008; Staesina and Davachi, 2009; Olsen et al., 2012; Aly et al., 2013) – a process common to both event conditions – the hippocampus seems to instead be preferentially engaged by events built from scenes (Hassabis and Maguire, 2007; Maguire and Mullally, 2013; McCormick et al., 2018a; Barry and Maguire, 2019a, 2019b; Ciaramelli et al., 2019).

Despite the measures taken to closely match event stimuli in terms of their sense of unfolding, scenes could simply be more engaging or predictable than patterns in events. If true, then one might have expected event memory to differ, but it did not, and both types of movie clips were easily recollected as clear narratives. I might also have expected to observe differences in oculomotor behaviour, but none were evident, also an indication of similar attentional processes for the two conditions. Furthermore, events were very well matched in terms of their evolving nature, reflected by equally high ratings for Thinking Ahead and Linking measures during piloting. Consequently, I believe that these two conditions were comparable in terms of predictability, and participants were equally engaged by the movies. It seems also unlikely that the difference between the two event types can be explained by working memory load. If Pictures-Linked movies were easier to hold in mind, while Patterns-Linked were more effortful to process, I would have expected this to be reflected at later points in the clips, as memory load increased, but I did not find any such effects.

Having established the involvement of the hippocampus in unfolding events, it was surprising to not also find a relative change in engagement of the vmPFC, given the extant neuroscientific evidence for its recruitment during scene and event processing

(Hassabis et al., 2007a; Epstein, 2008; Bertossi et al., 2016b; Zeidman and Maguire, 2016; McCormick et al., 2018a; Barry et al., 2019a). Notably, at the sensor level I did find frontal (as well as temporal) sensors involved in the Pictures-Linked effect, similar to previous sensor level findings in the context of autobiographical memory retrieval (Hebscher et al., 2020). However, given that source reconstruction provides an improved estimation of the distribution of sensor level effects, the absence of frontally-located power changes requires unpacking.

While McCormick et al. (2020) compared event processing to a low level baseline condition (counting), my comparison was between two very closely matched event types. My scene and non-scene events comprised linked image sequences, and anticipation of future activity upon viewing the first image likely occurred in both, such as “he will lose the balloon” (Pictures-Linked), and “the triangles will move apart” (Patterns-Linked), as noted in the participant feedback provided during piloting of these stimuli. Since the vmPFC appears to have a more general anticipatory function (e.g. Igaya et al., 2020), involved in linking conceptual information, including more abstracted ideas (Roy et al., 2012), and is associated with a variety of imagery content (Ishai et al., 2000; Yomogida et al., 2004), this region may not respond preferentially to scene-based stimuli in this paradigm. Therefore, it is possible that I am not able to deduce the contributions of the vmPFC here because it is not attuned to scene-based representations alone, but is instead important for the processing of complex unfolding representations more broadly, helping to guide adaptive behaviour. Future studies may help to establish if this supposition is accurate.

In summary, this MEG study revealed early hippocampal engagement associated with the initiation of events built from scenes, over and above evolving abstract sequences

built from non-scene imagery. Together with the hippocampus, involvement of other regions, including posterior parietal, inferior frontal, premotor, and cerebellar cortices, may reflect the processing of biologically-relevant evolving information, which typifies the scene-rich episodes we encounter in our daily lives.

7 General discussion

7.1 Précis

In this final chapter I discuss the main findings of my four experiments within a broader context. These experiments are referred to in this chapter as: “Scenes/Arrays” (Chapter 3), where the step-by-step construction of scenes was contrasted with the construction of non-scene arrays; “SD/SA” (Chapter 4), where SD objects were contrasted with SA objects during scene construction; “BE” (Chapter 5), where extrapolation beyond the border of scenes occurred; and “Events”, where scene and non-scene event construction were compared while participants watched animated movies (Chapter 6). I begin by summarising briefly the main results from each experimental chapter, before proceeding to relate them to the simple cognitive model described in Chapter 1 (Section 1.4; McCormick et al., 2018a), and then widening my discussion to consider the implications of my findings more broadly. To conclude, I suggest directions for future research.

7.2 Summary of the main results

The broad aim of this thesis was to better characterise, on a finer temporal scale than previously achieved, the neural dynamics underlying two processes: (1) the construction of visual scene mental imagery, and (2) the initiation of dynamic events that unfold from a series of scenes. To do this, I conducted four MEG experiments that sought to shed light in particular on the nature of the involvement of the hippocampus and vmPFC, two regions proposed to interact to mentally represent scenes and unfolding events. My research was motivated by a simple model of scene and event construction within which

these two brain areas align hierarchically (McCormick et al., 2018a; Ciaramelli et al., 2019), as described in Chapter 1. For simplicity, I refer here to this model as the “scene construction model”, in recognition of its relationship with the scene construction theory (Hassabis and Maguire, 2007; Maguire and Mullally, 2013). Each experiment was designed to address a different aspect of the model – how a single scene is constructed, the preparation of a single scene for inclusion into an extended event representation, and the initiation of event construction.

In the Scenes/Arrays experiment, I found that the anterior hippocampus and vmPFC were preferentially engaged during the earliest stages of scene construction (3 objects and a 3D space), relative to array construction (3 objects and 2D space). This effect was restricted to the theta band (4-8 Hz). During the very first construction stage of scenes, the vmPFC drove hippocampal theta, whereas the anterior hippocampus drove vmPFC activity during the first stage of array construction. In the SD/SA experiment, SD objects were associated with greater engagement of the vmPFC and an area in the anterior temporal lobe, the STG, relative to SA objects. This was again characterised by changes in theta power, and the vmPFC drove activity in the STG.

The other two experiments focussed on evoked responses to the transition between a single scene and an unfolding event. In the BE study, the implicit extrapolation of scenes (the BE effect) emerged rapidly within 12.5-58 ms of viewing scenes, and was associated with power changes in the hippocampus, vmPFC, and other scene processing regions (including PHC, RSC, thalamus, and visual association cortex). Consistent with the previous experiments, this process was also driven by the vmPFC. Finally, in the Events experiment, within 178-447 ms of the onset of the very first movie frame there was a difference in evoked responses for scene-based events compared to abstract non-

scene events. The hippocampus showed power changes associated with early processing of scene-based events, over and above abstract non-scene events, as did the posterior parietal, inferior frontal, premotor and cerebellar cortices.

Having already discussed each set of results in their relevant chapter, here I consider the two central questions posed in Chapter 1: how do we construct individual scenes, and how do these sets of single scenes become unfolding events?

7.3 The mental construction of single scenes

As described in Chapter 1, for most people, scene imagery is prevalent across cognition. Our continuous perception of the world and recollection of important moments in our lives fundamentally involves scene imagery, and empirical work has confirmed its dominance in a variety of contexts (Hassabis et al., 2007a, 2007b; Epstein et al., 2017; Clark et al., 2019, 2020; Epstein and Baker, 2019), even if the scene imagery is vague or fleeting (Clark et al., 2020). In this thesis I used a previously established definition of a scene – a naturalistic three-dimensional spatially-coherent representation of the world, typically populated by objects, and viewed from an egocentric perspective (Maguire and Mullally, 2013; Dalton et al., 2018). This may appear to be at odds with other characterisations of spatial representations. For example, allocentric, viewer-independent spatial representations are held to guide our actions, particularly during navigation (Tolman, 1948; O'Keefe and Nadel, 1978; Burgess, 2006; Klinghammer et al., 2016; Chen et al., 2018; Chen and Crawford, 2020; Karimpur et al., 2020), as exemplified in the cognitive map theory (O'Keefe and Nadel, 1978). The egocentric framework I emphasise in this thesis, however, seems to me to have particular relevance for the first-person experiences we witness in real-time (Crawford et al., 2011; Maguire and Mullally, 2013;

Dalton et al., 2018), whereas an allocentric reference frame, and switching between allo- and egocentrism, may have greater utility for large-scale navigation (Burgess, 2006; Karimpur et al., 2020).

My thesis focussed upon establishing whether the hippocampus and vmPFC were preferentially recruited to build this egocentric scene imagery. While the hippocampus has been consistently implicated in this form of imagery (Hassabis and Maguire, 2007; Epstein, 2008; Mullally et al., 2012a; Zeidman et al., 2015b; Dalton et al., 2018; Robin et al., 2018; Barry et al., 2019a, 2019b; Robin and Olsen, 2019; McCormick et al., 2020), debate surrounds the exact nature of the constructive process with which this region is involved. Is it in the service of building scene imagery *per se*, as posited by the scene construction theory (Hassabis and Maguire, 2007; Maguire and Mullally, 2013; Clark and Maguire, 2016; Zeidman and Maguire, 2016), or is its function more general, building any form of mental representation comprising numerous associated components, as suggested, for example, by the relational theory (Cohen and Eichenbaum, 1993; Lee et al., 2005; Hassabis and Maguire, 2007; Schacter and Addis, 2007; Yonelinas et al., 2019)?

I note that the cognitive map theory also alludes to the unification of elements to build map-like representations of the environment to guide our decisions, particularly in the context of navigation (Tolman, 1948; O'Keefe and Nadel, 1978; Epstein et al., 2017). However, this account emphasises the allocentric nature of these representations. Given that the stimuli used in this thesis were egocentric, and my tasks had no navigational requirement, my findings do not directly speak to comparisons between the scene construction theory and cognitive mapping.

The scene construction model further posits that the vmPFC, a region closely connected to the hippocampus (Catani et al., 2012, 2013; Adnan et al., 2016), initiates and supports scene construction processes (McCormick et al., 2018a; Ciaramelli et al., 2019). The question regarding scene-specificity, therefore, extends to include this prefrontal region.

7.3.1 vmPFC-hippocampal collaboration is key to building scenes

The MEG experiments in this thesis provided a time-resolved window on the question of how scenes are built from their constituent elements. In the Scenes/Arrays experiment, by directly comparing the gradual construction of scenes with that of non-scene arrays, I conducted a more precise examination of the individual stages of these constructive processes than was previously possible using fMRI (Summerfield et al., 2010; Dalton et al., 2018). This experiment provided several novel insights. First, the anterior hippocampus and vmPFC were preferentially engaged during scene construction, and this occurred at the earliest stages of scene building. This response to scenes compared to the construction of non-scene arrays speaks against purely relational accounts of hippocampal processing, since scenes and arrays were closely matched in terms of the multiple elements of which they were composed, and constructive and associative processing.

DCM then revealed that the vmPFC drove hippocampal activity during the very first stage of scene construction, a relationship that was not evident during array construction. This aligns with the scene construction model's prediction that in order to facilitate the hippocampal construction of scenes, the vmPFC provides early input to initiate the process (McCormick et al., 2018a)

. Neuropsychological evidence in support of this idea was described in some detail in Chapter 1. Notably, impairments associated with hippocampal damage seem to primarily converge around the inability to mentally construct spatially coherent scene imagery (Hassabis et al., 2007b; Maguire and Mullally, 2013; Clark and Maguire, 2016; McCormick et al., 2018a). These patients can, however, name appropriate scene elements, and detect semantic violations in scenes (McCormick et al 2017), signalling their semantic knowledge is intact. In contrast, vmPFC-damaged patients are not impaired in constructing scenes provided they are guided by highly specific cues (Kurczek et al., 2015), suggesting a role for the vmPFC in the initiation of endogenous processing. This is further evidenced by their limited mind-wandering (Bertossi and Ciaramelli, 2016) and difficulties elaborating beyond a single scene to construct an extended event (Bertossi et al., 2016b). These findings also resonates with the results of my BE experiment, where the vmPFC again exerted causal influence on scene-processing regions, in this case to initiate the extrapolation of the scene beyond its visible borders, an implicit, additive form of scene construction that involves making predictions about what might be beyond the view (I will return to this experiment in the next section).

Related to the inability to muster information relevant to a scene or event, many patients with vmPFC damage also confabulate, characterised by unintentionally recounting false memories (Gilboa et al., 2006; Ghosh et al., 2014). This implies a fundamental inability to suppress information irrelevant to the current episode. In fact, these patients display deficits in the processing of schema-related information even when memory is not being tested (Melo et al., 1999; Ghosh et al., 2014; Warren et al., 2014). For example, in response to simple, single words such as “receptionist”, they could not produce commonly associated schema such as “going to the doctor’s” (Ghosh et al.,

2014). Of note, in healthy individuals, the vmPFC appears to provide a form of data reduction for incoming complex signals during learning, helping one to efficiently hone in on the most task-relevant information (Mante et al., 2013; Mack et al., 2020). My findings may therefore reflect the necessity of the vmPFC in providing early guidance in making relevant schematic knowledge available to the hippocampus to build scenes using appropriate elements.

If so, then the same vmPFC-driven early reinstatement of learned schemas was not elicited by the initiation of non-scene arrays, where the hippocampus instead drove the vmPFC, perhaps in pursuit of greater guidance due to a lack of coherence, akin to a prediction error (Friston and Kiebel, 2009). By “slowing down” scene construction into three stages, I was able to also observe the incorporation of each additional scene element subsequent to this initial vmPFC-directed “scene-setting”. The mutual entrainment of hippocampus and vmPFC observed at the remaining stages may reflect iterative feedback loops between them to facilitate the integration of each object to form a fully realised scene.

I propose that the vmPFC initiates the generation of a 3D scene representation, and that this may reflect its established function in the activation of a relevant schematic framework. This schema may then serve to guide other regions (particularly the hippocampus) to build the scene. For this to occur at the first stage of the construction paradigm in Scenes/Arrays, suggests that a single object in a 3D space was sufficient to elicit associated conceptual knowledge about the potential scene that was in the process of being constructed.

Indeed, objects appear to be an important source of information about the scene being imagined (or viewed), signalling to the observer the most likely category (Oliva and Torralba, 2006; Summerfield and Egner, 2009; Clarke and Tyler, 2015; Trapp and Bar, 2015). The influence of objects is highlighted in particular when little information is available to us about the scene context, and thus an object could help to resolve ambiguity (Chiou and Lambon Ralph, 2016). This is an interesting point, because the scene stimuli used for both the Scenes/Arrays and SD/SA experiments were deliberately composed of semantically *unrelated* objects, which may have further increased the engagement of schema-relevant regions, such the vmPFC. It was in the SD/SA experiment that the contribution of in-scene objects was explicitly examined, the results of which provided further indications of the likely influence of prior schematic knowledge during the construction of scene imagery, as I consider next.

7.3.2 In-scene SD objects may be a source of schematic knowledge

My second experiment revealed only two regions that differentiated between the processing of SD objects and SA objects during scene construction. In addition to the vmPFC, engagement was observed in the anterior temporal lobe, specifically in the STG. The anterior temporal lobe is known to facilitate processing of semantic and conceptual knowledge in support of visual object identification, particularly knowledge pertaining to the scenes in which objects are typically found (Patterson et al., 2007; Peelen and Caramazza, 2012; Lambon Ralph, 2014; Chiou and Lambon Ralph, 2016). Indeed, patients with semantic dementia, who have anterior temporal lobe atrophy, have impoverished conceptual knowledge, while perceptual knowledge is typically preserved (Campo et al., 2013; Guo et al., 2013).

The strong structural connectivity between the anterior temporal lobe and vmPFC via the uncinate fasciculus (see Section 1.3; Von Der Heide et al., 2013), as well as the direct input to the vmPFC from low-level visual areas via the fronto-occipital fasciculus (Catani et al., 2012), makes the vmPFC well placed to coordinate scene construction via feedback connections to regions in the anterior temporal lobe. Results from the DCM analysis of the SD/SA experiment extended those of the Scenes/Arrays study, by showing that the vmPFC again drove activity downstream during scene construction, in this case influencing the STG.

Engagement of the STG in response to SD objects, driven by activity in the vmPFC, may indicate that these kinds of objects are an important source of conceptual information. It could be that due to their inherent elicitation of surrounding 3D space, SD objects provide the key signal about the likely scene gist. For example, by seeing a park bench, an object which is typically defined by its confinement to the ground of a park, makes it more likely that the developing scene is that of a park. Mullally and Maguire (2013) found that their participants showed a clear preference for SD objects over SA objects, both when selecting objects to construct scenes, and when retaining objects during scene deconstruction. In providing greater access to useful conceptual information, SD objects may therefore help to anchor and define a scene. Consistent with the scene construction model (McCormick et al., 2018a), top-down modulation from the vmPFC may guide conceptual processing in areas downstream like STG, and this appropriate information could be what is conveyed to the hippocampus to build an image of the scene.

The relationship between scene SD objects and schema also speaks to the active inference scheme (Mirza et al., 2016). This statistical formulation recognises the world is

inherently an ambiguous place, requiring epistemic (explorative sampling) behaviour to accumulate the evidence available to form and repeatedly update probabilistic beliefs (internal models). Ultimately this serves to minimise ambiguity and prediction errors (Friston et al., 2010). In this account, scene construction is regarded as involving active visual foraging through multiple saccadic eye movements that inspect individual components of the scene, to deduce the likely scene category. In the face of uncertainty, when insufficient information exists to form a belief about the probable scene category, prior knowledge steps in to compensate (Jóhannesson et al., 2016; Mirza et al., 2016; Yang et al., 2016; Heins et al., 2020). In this way, SD objects may act as visual cues to categorise a scene, particularly when confronted with an ambiguous context. More broadly, as a scene representation develops, information is accumulated and updated about the scene, perhaps orchestrated under the supervision of the vmPFC.

The top-down influence of the vmPFC during the effortful construction of novel scenes that I observed in the Scenes/Arrays and SD/SA experiments, aligns with its role in coordinating autobiographical memory retrieval (McCormick et al., 2020), integrating existing knowledge to form schemas (van Kesteren et al., 2010a, 2010b; Navawongse and Eichenbaum, 2013; Kumaran et al., 2015), and reinstating schemas during new experiences (Melo et al., 1999; Tse et al., 2007; Benoit et al., 2014; Ghosh et al., 2014; Warren et al., 2014; Spalding et al., 2015). Collectively, these findings bolster the view that the vmPFC makes inferences based on the past, allowing us to anticipate the immediate future, and flexibly guide behaviour (Kveraga et al., 2007; Schlichting and Preston, 2015; Heins et al., 2020; Parr et al., 2020).

So far I have found support for the position that the hippocampus is preferentially recruited to build scenes, and that the vmPFC initiates and drives this process. I also

suggest that the vmPFC has access to prior knowledge that supports scene construction. However, the Scenes/Arrays and SD/SA paradigm involved a gradual and deliberate imagination process, allowing participants to move their eyes to fixate on and inspect each individual component as it was mentally generated. Beyond this explicit form of mental imagery, making inferences is a critical aspect also of spontaneous cognitive processing, for example when we view whole scenes. In the BE and Events experiments, I probed this idea further, going beyond the voluntary imagination of scenes to capture instead more implicit, involuntary processing that may underpin the transition from an individual scene to an event. These investigations also revealed the possible anticipatory function of the vmPFC and its operationalisation of schema during implicit imagery.

7.4 From single scenes to unfolding events

7.4.1 *The automatic extrapolation of single scenes*

Traditionally, visual stimulus recognition has been conceptualised as a bottom-up process, where visual input arrives at low-level regions, which is then fed forward to higher-level regions, after which semantic processing occurs (Bullier, 2001; Kveraga et al., 2007). This would suggest that we only understand what we are looking at after processing across multiple saccades. However, a purely feedforward visual perception framework is problematic when we consider that our natural experience of the world is often fleeting, yet still requires comprehension. Similarly, while the contribution of multiple eye movements to explore the rich detail of our world is clearly important, there is adaptive and survival value in apprehending the world even within a single fixation.

Previous behavioural studies have shown that as little as 13 ms exposure is needed to categorise scenes (Potter, 1975; Thorpe et al., 1996; VanRullen and Thorpe, 2001; Rousselet et al., 2005; Joubert et al., 2007; Castelhana and Henderson, 2008; Greene and Oliva, 2009; Potter et al., 2014), even when a scene is displayed in the peripheral vision for only 28 ms (Thorpe et al., 2001). Despite these observations, we cannot appreciate every aspect of our environment in a given moment, and during real-world scene and event perception, ambiguities cannot be resolved simply from the information afforded to us from the immediate sensory input. It would, therefore, be advantageous to have an anticipatory mechanism that extends our experience, making any ambiguities, occlusions or transitions between scene views imperceptible as we look around and move through our environment. This implicit form of prediction of what is currently not in view is exemplified by the BE effect, which I examined in my third experiment.

After briefly (250 ms) viewing photographs of scenes, participants often erroneously recall having seen a wider view than they witnessed, reflecting the tendency to extrapolate beyond the scene's borders during its first viewing (Intraub and Richardson, 1989). This speaks to the relationship between an implicit scene construction process underpinning visual perception and episodic memory encoding, since it may provide the crucial anticipatory mechanism through which individual scene views come to overlap and are thus seamlessly linked, enabling a continuous event to be experienced. It is notable that I found this extrapolation of scenes was facilitated by the high-level regions typically engaged by scene processing, including the vmPFC, hippocampus, PHC and RSC (e.g. Maguire, 2001; Svoboda et al., 2006; Park et al., 2007; McDermott et al., 2009; Spreng et al., 2009; Chadwick et al., 2013; Zeidman et al., 2015). I also found direct evidence of top-down processing, with the vmPFC exerting a driving influence over the

other brain regions, including the hippocampus. I suggest that this internal extension of a scene representation corresponds to an implicit additive scene construction process, supported by hippocampus and vmPFC, that occurred during BE trials over and above that associated with the scene processing that was also required during NoBE trials.

The vmPFC, therefore, not only appears to drive activity downstream during the construction of a single scene image, addressing the first part of the cognitive model outlined in Chapter 1 (Figure 14A-B; McCormick et al., 2018a), but may also prepare the scene for inclusion into a broader context (Figure 14C). As the difference in processing between trials where BE occurred and those where it did not results from subjective perception, my findings further support the conceptualisation of the vmPFC as a region that instigates anticipatory processing, making inferences about what lies beyond the view from the available sensory input based on prior experience (Friston, 2010; Friston et al., 2010; Ghosh and Gilboa, 2013; Ghosh et al., 2014; Spalding et al., 2015; Gilboa and Marlatte, 2017; Parr et al., 2020). This is consistent with deficits exhibited by patients with damage to either the vmPFC or hippocampus. vmPFC-damaged patients are impaired in the retrieval of semantic and schema-oriented knowledge (D'Argembeau and Mathy, 2011; Irish et al., 2012; Ghosh et al., 2014) and in guiding attention towards predicted rewards (Vaidya and Fellows, 2015), whilst their ability to construct scenes remains intact if appropriately cued (Kurczek et al., 2015). In contrast, hippocampal patients' deficits are centred around the generation of spatially coherent mental representations of scenes (Hassabis et al., 2007a, 2007b; Mullally et al., 2012a; Clark and Maguire, 2016; McCormick et al., 2018a), whereas their capacity to engage in high-level reasoning is preserved (Mullally and Maguire, 2014).

The vmPFC and hippocampus have also been implicated in episodic future simulation (Zeithamova et al., 2012; Benoit et al., 2014; Weilbacher and Gluth, 2016; Campbell et al., 2018; Schacter and Addis, 2020), and patients with either vmPFC or hippocampal damage exhibit impairments in this cognitive function (Bertossi et al., 2016a, 2016b; Palombo et al., 2018). BE may reflect a similar process of mental simulation, facilitated by vmPFC driving the hippocampus to predict beyond the edges of the scene with reference to schematic knowledge (Benoit et al., 2014; De Luca et al., 2018; McCormick et al., 2018a).

A particularly surprising finding from my BE experiment was the demonstration that the top-down processing by a high-order brain region (the vmPFC) occurred so rapidly (under ~ 58 ms of scene onset). This accords with extant behavioural studies that indicate the overall scene gist is quickly extracted (Greene and Oliva, 2009; Larson et al., 2014; Potter et al., 2014). It also implies that feedback mechanisms occur very early on during the visual processing that supports scene comprehension (Di Lollo et al., 2000; Lamme and Roelfsema, 2000; Bullier, 2001; Hochstein and Ahissar, 2002; Bar, 2004; Bar et al., 2006). As I discussed in the relevant experimental chapter (Section 5.5), it remains a matter of debate as to what exactly scene categorisation-related processing entails when performed so rapidly. Nevertheless, in terms of BE, sufficient higher-level processing to enable prediction of elements concordant with the scene in view must have occurred in order to generate a wider view of the scene.

This has some parallels with pattern completion, which involves the recall of a whole representation from a partial cue (Marr, 1971; Hopfield, 1982; McClelland et al., 1995; Norman and O'Reilly, 2003; Rolls, 2013), an important feature of episodic memory (Horner et al., 2015; Grande et al., 2019). Pattern completion may go some way towards

explaining the function of the hippocampus in BE, as the scene in view might trigger the retrieval of schematic knowledge (facilitated by the vmPFC) with which the hippocampus then constructs more of the scene with elements that are congruent with the information currently in view. It should be noted that pattern completion is associated with distinct hippocampal subfields (Rolls, 2013). However, these cannot be distinguished using MEG, and so my data cannot speak to this aspect of pattern completion.

The ability to efficiently extract the overall gist from complex visual scenes as they are experienced, and drawing upon previous experience of similar scenes to automatically flesh out the scene beyond its bounds, represents a highly adaptive process. With a driving influence from the vmPFC to high-level visual areas, inferences can be efficiently verified or refuted as the scene changes (Mack et al., 2016; Ahlheim and Love, 2018), allowing us to efficiently update our internal model of the world and plan our actions.

The findings from this thesis discussed so far appear to suggest an account of the vmPFC as a possible mediator of schema and conceptual knowledge, facilitating the construction of spatially coherent scenes by the hippocampus, along with other brain areas that support this construction. This may resonate with accumulating evidence for the vmPFC's possible role in coding schematic or abstract knowledge structures more generally (Constantinescu et al., 2016; Schuck et al., 2016; Wikenheiser and Schoenbaum, 2016; Mack et al., 2020). Contributions of the vmPFC to this type of processing have been largely examined in the context of learning over a longer time frame (Constantinescu et al., 2016; Schuck et al., 2016; Wikenheiser and Schoenbaum, 2016), but also during encoding (Miller and Cohen, 2001; Mack et al., 2020). It is unclear whether the vmPFC is responsible for forming conceptual representations, or if its activity reflects the retrieval of this knowledge from elsewhere. The scene construction model suggests its

hierarchically elevated position with respect to areas that build scene imagery (e.g. hippocampus) reflects processing that reactivates prior schema relevant to the current context, with suppression of irrelevant details (McCormick et al., 2018a; Ciaramelli et al., 2019).

Other perspectives, such as “SUSTAIN” (Love et al., 2004), suggest a similar flexible, goal-oriented function for the vmPFC that creates a knowledge structure, wherein only relevant elements are selected in response to the current visual input, that are fed to areas like the hippocampus to drive goal-directed action (Mack et al., 2020). I suggest that the neuronal effects observed during BE reflect the vmPFC drawing inferences, informing expected outcomes, and thus providing the hippocampus with the means to update internal mental representations to optimise goal-oriented behaviour. A similar process may facilitate navigation and episodic memory processing.

7.4.2 Linking scenes to make events

BE might represent the beginning of an event being constructed, allowing for a smooth transition beyond a single scene to incorporate successive scenes. In the Events experiment, I sought to directly examine how sets of scenes, compared to non-scenes, were linked to become unfolding events. In a similar way to the Scenes/Arrays experiment, I examined how attuned the hippocampus and vmPFC were to scene-based imagery, but in the context of an extended narrative. The result was the preferential response of the hippocampus to the very first frame of scene events, over and above events comprising abstract patterns, with no further differences detected between the two conditions throughout the remainder of the movie frames. This further highlights the importance of the hippocampus in processing scene imagery, and the establishment of

this imagery at a very early stage of construction, also corroborating what I found during BE in terms of early hippocampal engagement.

Notably, throughout the experiments in this thesis, scenes have consistently elicited responses from the hippocampus, in line with previous studies involving the construction of scene imagery (Hassabis et al., 2007a; Lee et al., 2012; Aly et al., 2013; Chadwick et al., 2013; Zeidman et al., 2015b; Clark et al., 2018; Dalton et al., 2018; Palombo et al., 2018; Barry et al., 2019a; McCormick et al., 2020). The building of scenes appears to be distinct from other closely matched processing, whether static non-scene arrays or evolving non-scene events. The additional involvement of brain areas known to support the interpretation and comprehension of biologically-relevant information (Battelli et al., 2003; Rizzolatti and Craighero, 2004; Saygin et al., 2004; Fraiman et al., 2014) during the scene-based movie events served to highlight the distinction between real-world events and abstract narratives.

To illustrate this point, scene-based event movies contained a stick-figure, with which pattern-based events were closely matched. In scenes, this central figure was an “agent”, acting intentionally within an identifiable context (e.g. opening an umbrella as it starts raining). By contrast, the corresponding pattern, although still an evolving narrative (i.e. one frame lead to the next to create a movement), had no intentional influence within a recognisable environment, and its movement was too abstract in nature to be interpreted in a behaviourally-relevant sense. This important difference conveys a key feature of what makes scenes unique stimuli, and not simply amalgamations of disparate items bound in space (Cohen and Eichenbaum, 1993; Addis and Schacter, 2012). Scenes represent a realistic depiction of our experience of the world, where the constellation of objects in space conveys meaning not present in abstract representations.

My examination of ongoing events resonates with other investigations of the paradoxical impression of our experience as continuous, whilst also being recalled as distinct self-contained episodes. For example, when we describe a personal event we easily provide a step-by-step narrative of what happened (as participants could do after viewing the event movies in my Events experiment). Similarly, in describing how to reach a destination, we may describe a sequence of turns and changes in direction (Feller et al., 2019; Shin and DuBrow, 2020). Early behavioural research on this topic focused on how people detect natural boundaries that mark the end of the current activity or context and the start of a new one, in tasks where participants attended to audio-visual events (such as a movie or verbal narrative). This was originally formalised as the “event segmentation theory” (Zacks and Tversky, 2001; Zacks et al., 2007; Swallow et al., 2009; for an extension of this theory, the Event Horizon Model, see Radvansky, 2012), which asserts that changes in brain activity triggered by “event boundaries” serves to update ongoing mental representations (Richmond and Zacks, 2017). I did not seek to examine the broad segmentation of events as defined by these researchers. Instead, my focus was on a finer unit of events, namely, the scenes of which they were composed, and how events built from these scene units were initiated in the brain. Boundaries in the “Zacks” (and colleagues) sense occurred in my stimuli only at the start and end of movies, as each movie depicted a single event. Nevertheless, it is worth considering what the neural processing at event boundaries reported by previous studies may signify, and how this might relate to my results.

Boundaries have important consequences for how events are organised in long-term memory. Temporal associations are more likely to be formed and recalled within rather than across events, for a variety of contexts, including perceptual boundaries (e.g.

Heusser et al., 2018), spatial boundaries separating rooms (e.g. Horner et al., 2016), and those representing turns during navigation (e.g. Brunec et al., 2020). fMRI work has consistently found that boundaries trigger increases in hippocampal activity (Ben-Yakov and Dudai, 2011; Ben-Yakov et al., 2013; Baldassano et al., 2017; Ben-Yakov and Henson, 2018). The nature of these boundary-evoked hippocampal responses is somewhat unclear, but may depend on task demands. For example, the hippocampus may attempt to maintain integration (pattern completion) despite a shift in context if the task is to recall a sequence, regardless of possible boundaries (see DuBrow and Davachi, 2016). In other cases, boundaries could represent a shift from pattern completion to pattern separation in the hippocampus to distinguish between overlapping information (Rolls, 2013; Clewett et al., 2019). Evidence for the latter in humans is most clearly evident in EEG studies that have documented replay of just-perceived events at boundaries (Sols et al., 2017), which resemble sequential reactivation at boundaries during rodent navigation (Foster and Wilson, 2007).

Hippocampal activity at the completion of movies could reflect working memory maintenance (see Leszczynski, 2011) because studies like that of Ben-Yakov and Dudai (2011) have used explicit memory tasks, where participants were instructed to attend to the event gist or details knowing they would be later tested. This may reflect not offline encoding processes, but rather replay of the newly formed event representations. This is a considerably different task to the movie watching in my experiment, where participants were instructed to simply attend to the narratives of clips, with no instruction to commit information to memory (which was then only tested in a surprise memory test post-scanning).

Similarly, the rodent navigation literature has shown that hippocampal activation at boundaries represents more complexity than simple reproduction of experience. It suggests that rats represent the currently navigated environment in segments, disproportionately representing the path that lies ahead at the start of a segment, and the path just completed as the animal approaches the end of the segment (Johnson and Redish, 2007; Gupta et al., 2010, 2012; for a review see Ólafsdóttir et al., 2018). This work implies that animals can re-experience journeys they have previously taken, as well as pre-play novel journeys not yet experienced, drawing parallels with our own ability to re-experience events as well as simulate future scenarios. I suggest this further reinforces the argument I have made that the hippocampus has an ongoing role in our experiences to update a flexible and changing internal model of the world.

The evoked hippocampal activity that I found cannot be attributed to endogenous event replay, since it occurred only at the first scene image frame, before the event had unfolded at all. Instead, I suggest my findings emphasise the role of the hippocampus in helping to create an internal model of the world, which can be subsequently updated in a flexible manner as the event unfolds. This can be used to simulate scenarios not yet experienced. As I found, this appears to be particular to sequentially presented scene-based imagery, over and above that of more general abstract sequential information.

The absence of the same hippocampal increase in engagement at the conclusion of the movies in my experiment may be because this traditionally defined boundary was simply not registered by my participants. This could have been because each brief movie clip depicted a single already discretised event, making this the only event boundary per stimulus. Furthermore, it is worth noting that if hippocampal activity at boundaries reflects a shift from pattern completion to pattern separation in order to encode new

information (Rolls, 2013; Clewett et al., 2019), without the explicit goal of remembering that chunk of information, the hippocampus may not be called upon at the event conclusion.

Another interesting difference between my experiment and previous event boundary movie studies concerns the cross-participant agreement about boundary placement. As I mentioned earlier, I could not examine the broad event boundaries as others have done because my stimuli were not constructed with that in mind. However, proponents of event segmentation theory also suggest that people register fine-grain boundaries nested within those of a temporally coarser grain, adding to the structure of events. For instance, if an actor is making a bed, a coarse-grain boundary marks the point at which all the pillow cases have been changed. Subsumed within this are the steps necessary to change individual pillow cases, such as picking up each pillow, and then putting on each pillow case (Zacks et al., 2001). Regardless of the granularity of boundaries, remarkable agreement in their locations between observers has been reported (e.g. Speer et al., 2003; Hasson et al., 2004; Zacks et al., 2007; Swallow et al., 2018).

I was interested in examining the neural underpinning of this finer level of segmentation within my movie events, and in fact conducted a separate behavioural experiment (not reported formally in this thesis, but described below) where participants viewed each event movie (both scene- and pattern-based). Although explicit instructions for event segmentation from previous studies are not typically provided in detail in the extant literature, participants are usually instructed to parse a movie into meaningful moments according to when they notice a change in the narrative, but “it is left to them to decide what constitutes a boundary and a meaningful event” (Zacks and Tversky, 2001; Zacks et al., 2001; Feller et al., 2019). I instructed participants in the same manner.

Participants first watched a movie from start to finish passively, and were told to only pay attention so as to understand what was taking place in the clip. During a second viewing, participants were instructed to segment the clip, by pausing the video (with a simple button press) and to explain to the experimenter their reason for doing so. An example of explanations given for boundaries during a scene-based event were: “she is skipping”, “she is opening her umbrella”, and “it is now raining”. For pattern-based events, examples included: “the shapes are rotating to the right”, “the shapes are expanding”, and “the shapes are separating”. Interestingly, the placement of these boundaries varied quite considerably across participants. For example, some participants noted the occurrence of the “raining” boundary at the first frame in which rain appeared (“started to rain”), by others it was indicated part way through the action (“it is raining”), and by others still, it was marked at the conclusion of the activity (“it rained”). It is unclear why this variability occurred, but the nature of stimuli may have contributed, as these were flip-book style cartoons, rather than cinematic movies as have been used previously. Given this lack of consistency across participants, I was unable to examine the MEG data according to event boundaries, but considering the clear influence event boundaries have on the organisation of associated information that form episodic memories, this would be interesting to pursue in a future experiment.

While the hippocampus was prominent in the comparison between scene-based and pattern-based events, interestingly, similar differential involvement of the vmPFC was not evident. Given its purported role in initiating the construction of scenes in concert with the hippocampus (De Luca et al., 2018; McCormick et al., 2018a; Barry et al., 2019a; Ciaramelli et al., 2019), one might have expected its contribution to be even greater during the construction of unfolding events. This is because such events entail a

number of scenes and thus multiple linking transitions, which would likely involve the retrieval of prior knowledge of how similar events typically unfold (McCormick et al., 2018a; Ciaramelli et al., 2019). This would also be in keeping with vmPFC-damaged patients being impaired at processing extended mental events, beyond a single scene (Bertossi et al., 2016b). Moreover, since the vmPFC is known to be involved in computing subjective value (Bartra et al., 2013; McNamee et al., 2013; Lin et al., 2016; Yamada et al., 2018), one might expect a larger vmPFC response to scene-based events since these may confer more useful and salient information to the viewer, containing as they do objects, people and scene contexts that might loosely bear some familiarity with people's own life experiences.

However, it is important to note that participants were passive observers of very short, simple activities in the scene-based movies. This is in contrast to tasks where participants are asked to imagine projecting themselves into the future in a novel scenario happening to them, which has been shown to involve the vmPFC (Benoit et al., 2014; Bertossi et al., 2017; Campbell et al., 2018; McCormick et al., 2018a). I therefore argue that the events in my experiment do not speak directly to this idea of “value” or future-oriented cognition as it relates to the self. The events built from scene imagery in my experiment were essentially devoid of social or personally relevant information, and simply depicted basic activities, while pattern-based events were too abstract to resemble any “real” activity.

An alternative explanation for the findings might be that the vmPFC is not particularly attuned to scene-based events, and may have a more general role, and therefore in the comparison of scene- and pattern-based events its engagement was cancelled out. There is evidence the vmPFC may not specifically represent visual aspects

of scenes, but instead the collection of constituent components of situational contexts (Aminoff and Tarr, 2015; Lieberman et al., 2019). These contexts include spatial, temporal, evaluative, and social aspects (Lieberman et al., 2019), and could include scenes (Ciaramelli et al., 2019) but also more abstracted and conceptual knowledge that helps the viewer to conceive the meaning of a situation to guide future actions (Roy et al., 2012). In this way, a scene- and a pattern-based event can both be representations of situations or schema, involving the abstraction of essential features to guide prospection, and grasp conceptual information and meaning.

However, there are also some distinctions between the Events experiment and the other experiments reported in this thesis that are worth noting. The vmPFC appears to be particularly called upon to initiate the active construction of novel scenes (relative to arrays), and to initiate implicit scene construction of new scene content in preparation for an event (during BE). It is possible that the initiation of internal construction of scene-based events during their *passive* viewing, as was the case in the Events experiment, did not rely upon the vmPFC to the same degree as the more active construction of content, where the need for endogenous elaboration was possibly greater. It has been shown that patients with bilateral vmPFC damage were able to discriminate between possible and impossible semantic scenes (e.g. an elephant with butterfly ears), even though an understanding of semantically impossible scenes would presumably require the re-activation of schemas congruent with the observed scene in order to understand the semantic violation (De Luca et al., 2019). The authors of that patient study suggested that their impossible scenes were sufficiently obvious that this schema reinstatement was unnecessary. The vmPFC may therefore not be needed for some aspects of scene processing. Overall, the absence of the vmPFC in the comparison between scene- and

pattern-based event unfolding cannot be determined from my data, but this represents another interesting avenue to pursue in future studies.

7.5 Outstanding questions and future directions

Little is known about the fine temporal dynamics of scene and event processing, including autobiographical memory, navigation, and imagining the future. In particular, the flow of information between the hippocampus and vmPFC is poorly understood, with few electrophysiological investigations to date (Fuentemilla et al., 2014, 2018; Jafarpour et al., 2014; Kaplan et al., 2017; Hebscher et al., 2019, 2020; McCormick et al., 2020). This underexplored question is what I sought to address in my thesis. Overall, the results indicate that the hippocampus is involved in building scene representations, including the deliberate creation of novel scene imagery, the implicit extrapolation of scenes during perception, and the processing of scenes that form the basis of naturalistic events that unfold before us. Moreover, its contribution starts very early on in these processes. My findings are also consistent with the vmPFC acting as a supervisor in scene construction, driving the construction of scene imagery, aligning with the architecture of the simple cognitive model of scene and event construction proposed by McCormick et al. (2018a).

Numerous questions remain about the vmPFC in particular, including its absence from my comparison between scene- and pattern-based events, possibly due to its broader role in the processing of complex unfolding representations. I suggest the vmPFC may be providing a general anticipatory function, reinstating existing schemas and conceptual information which then guides processing elsewhere, such as the hippocampus during scene construction. I did not, however, formally test how schemas underpin the relationship between the hippocampus and vmPFC. My paradigms could

be adapted in this regard in future studies. For instance, schema-induced encoding could be explicitly manipulated to provide fresh insights about the influence exerted by prior and conceptual knowledge on memory formation, to better characterise the distinct contributions of the hippocampus and vmPFC.

Furthermore, while not the focus of this thesis, how might my findings relate to memory consolidation and retrieval? The influence exerted by the vmPFC over the hippocampus in my experiments during scene and event processing mirrors the vmPFC-driven relationship with the hippocampus found by McCormick et al. (2020) during autobiographical memory retrieval, which was notably true for both recent and remote memories. It would be interesting to explore such findings further by including a recall component to versions of my tasks, and manipulating the interval between encoding and retrieval.

An interesting feature of this thesis was the observation of decreases in oscillatory power within the theta band. This is, in fact, a feature of the human electrophysiological field more generally, and previous experiments involving scene construction and episodic memory have also been associated with decreases in theta power (Barry et al., 2019a; McCormick et al., 2020). However, both increases and decreases in theta, and in other low-frequencies, have been reported in human studies of episodic memory, presenting a complex picture of the role of theta oscillations in this processing. For instance, memory-related low-frequency power increases have been consistently found in spatial navigation and spatial working memory tasks (Ekstrom et al., 2005; Aghajan et al., 2017; Bohbot et al., 2017; Bush et al., 2017; Kaplan et al., 2017). However these power increases are less often associated with successful encoding during episodic memory, unlike power decreases, where such relationships have been reported

(for a review see Herweg et al., 2020). Notably, the majority of intracranial studies of human episodic memory show decreases in MTL theta (Greenberg et al., 2015; Long and Kahana, 2015; Solomon et al., 2017, 2019; Fellner et al., 2019) and these have been inversely associated with the fMRI BOLD signal (Fellner et al., 2016).

Whilst theta has been proposed to support associative memory, a common feature of spatial navigation and episodic memory, the mixed findings of power increases and decreases across associative forms of memory appears to go against this general-purpose mechanism (Herweg et al., 2020). Furthermore, where theta power increases (rather the usual decreases) have been reported in human intracranial studies, these tend to be short-lived and occur early, as well as being weaker and more spatially restricted (e.g. Burke et al., 2014; Long et al., 2014), perhaps reflecting underlying event-related activity, rather than sustained oscillatory activity (Herweg et al., 2020). While the mechanistic role of power attenuation is not entirely clear, the extant evidence indicates that it does have an active role in memory encoding (Hanslmayr et al., 2012), including when generating immediate novel scene imagery (Barry et al., 2019a) autobiographical memory retrieval (McCormick et al., 2020), and now the gradual construction of scene representations (Chapters 3 and 4). As the use of MEG increases for complex episodic memory processes, a clearer picture may begin to emerge.

New neuroimaging methods are currently in development which may also advance the application of MEG for the study of scene and event processing. In conventional SQUID-MEG or fMRI, participants must remain very still to mitigate problems with the accuracy of source reconstruction. Optically-pumped MEG technology (OP-MEG) permits the participant to move freely during tasks (Boto et al., 2018; Barry et al., 2019b). This offers the exciting possibility to directly study cognitive

processes such as autobiographical memory and navigation while the participant explores and engages with their environment in more naturalistic contexts. As OP-MEG still requires measurements to be made within a magnetically shielded room, this will necessarily take the form of VR (Roberts et al., 2019), as has already been conducted to some extent using EEG (e.g. Liang et al., 2018). These experiments can be designed to be immersive, approximating more closely how events are experienced “live”. This will provide a new dimension to examining the hippocampus-vmPFC relationship during encoding, particularly during events that are self-relevant to the participant. More immersive environments may have an influence on the degree of recruitment of these regions and the strength of their connectivity, as they could involve active decision-making or social interactions (Roy et al., 2012; D’Argembeau, 2013; Kurczek et al., 2015). OP-MEG technology also offers an exciting opportunity to study these online processes non-invasively with millisecond temporal precision (and without having to limit movement) in patients with hippocampal or vmPFC damage, to study how new memories are acquired, and indeed possibly fade, in populations where the scene “builder” (hippocampus) or “supervisor” (vmPFC) is absent.

Finally, in this thesis I have described the vmPFC as a single area. However, it is a complex region, encompassing several Brodmann areas (10, 14, 25, as well as parts of 32, 11, 12 and 13; Catani et al., 2012; McCormick et al., 2018a; Ciaramelli et al., 2019) and demonstrates clear heterogeneity in structure and function (Rempl-Clower and Barbas, 2000; Binder et al., 2009; Rushworth et al., 2011; Roy et al., 2012; Winecoff et al., 2013). However, its precise structural and functional organisation, including connectivity with MTL, remains unclear. This is at least partly due to the signal drop-out and distortion within the vmPFC in fMRI given its proximity to air-filled areas (such as eye sockets; see

Embleton et al., 2010; Halai et al., 2014), although in recent years advances in multi-echo acquisition methods has helped to mitigate this problem (e.g. Poser and Norris, 2009).

Overall, I found that anterior and dorsal vmPFC regions were engaged during scene-related processing which, as previous meta-analyses have suggested, may reflect cognitive processing that includes simulation, mentalisation, construction of internal models of experience, and self-projection (Roy et al., 2012). However, due to the poorer spatial resolution of MEG, confident conclusions regarding the involvement of precise sub-regions is difficult. Ultra-high resolution fMRI, for example at 7.0 Tesla, represents another exciting area for future experimentation, to hone in on the roles played by different sub-regions, and indeed the cortical layers, of the vmPFC in various forms of scene and event processing. This could also further illuminate the functions of specific hippocampal subfields. This type of data would serve as a useful complement to my MEG findings. For example, it would be interesting to consider whether the voluntary generation of scene imagery engages areas distinct from implicit forms of scene construction. It could also be instructive to examine predictive aspects of BE and schema processing, and how this relates to the cortical microcircuitry within the layers of the vmPFC.

In conclusion, the experiments in this thesis offer new insights into how the hippocampus and vmPFC cooperate to support the scene imagery that underpins so much of our cognition. In particular, the use of MEG allowed me to characterise their relationship with millisecond precision, and emphasised how key responses and interactions occur very early during scene and event construction. Clearly much remains to be understood. I hope the work reported here provides a useful stepping stone for

further discoveries about the cognitive functions that critically depend upon the dialogue between the hippocampus and vmPFC.

8 References

- Addis DR, McIntosh AR, Moscovitch M, Crawley AP, McAndrews MP (2004) Characterizing spatial and temporal features of autobiographical memory retrieval networks: A partial least squares approach. *Neuroimage* 23:1460–1471.
- Addis DR, Moscovitch M, McAndrews MP (2007) Consequences of hippocampal damage across the autobiographical memory network in left temporal lobe epilepsy. *Brain* 130:2327–2342.
- Addis DR, Pan L, Vu MA, Laiser N, Schacter DL (2009) Constructive episodic simulation of the future and the past: Distinct subsystems of a core brain network mediate imagining and remembering. *Neuropsychologia* 47:2222–2238.
- Addis DR, Schacter D (2012) The hippocampus and imagining the future: Where do we stand? *Front Hum Neurosci* 5:173.
- Adnan A, Barnett A, Moayedi M, McCormick C, Cohn M, McAndrews MP (2016) Distinct hippocampal functional networks revealed by tractography-based parcellation. *Brain Struct Funct* 221:2999–3012.
- Aggleton JP, Wright NF, Rosene DL, Saunders RC (2015) Complementary patterns of direct amygdala and hippocampal projections to the macaque prefrontal cortex. *Cereb Cortex* 25:4351–4373.
- Aghajan ZM, Schuette P, Fields TA, Tran ME, Siddiqui SM, Hasulak NR, Tcheng TK, Eliashiv D, Mankin EA, Stern J, Fried I, Suthana N (2017) Theta oscillations in the human medial temporal lobe during real-world ambulatory movement. *Curr Biol* 27:3743–3751.e3.
- Ahlheim C, Love BC (2018) Estimating the functional dimensionality of neural representations. *Neuroimage* 179:51–62.

-
- Aly M (2020) Brain dynamics underlying memory for lifetime experiences. *Trends Cogn Sci* 24:780–781.
- Aly M, Ranganath C, Yonelinas AP (2013) Detecting changes in scenes: The hippocampus is critical for strength-based perception. *Neuron* 78:1127–1137.
- Aminoff EM, Tarr MJ (2015) Associative processing is inherent in scene perception. *PLoS One* 10: e0128840.
- Amodio DM, Frith CD (2006) Meeting of minds: The medial frontal cortex and social cognition. *Nat Rev Neurosci* 7:268–277.
- Andrä W, Nowak H (2006) *Magnetism in medicine: A handbook* (Andrä W, Nowak H, eds). Wiley.
- Andrews-Hanna JR, Reidler JS, Sepulcre J, Poulin R, Buckner RL (2010) Functional-anatomic fractionation of the brain's default network. *Neuron* 65:550–562.
- Argyropoulos GPD, Loane C, Roca-Fernandez A, Lage-Martinez C, Gurau O, Irani SR, Butler CR (2019) Network-wide abnormalities explain memory variability in hippocampal amnesia. *Elife* 8:e46156.
- Attal Y, Schwartz D (2013) Assessment of subcortical source localization using deep brain activity imaging model with minimum norm operators: A MEG study. *PLoS One* 8:e59856.
- Auger SD, Maguire EA (2013) Assessing the mechanism of response in the retrosplenial cortex of good and poor navigators. *Cortex* 49:2904–2913.
- Auger SD, Maguire EA (2018) Retrosplenial cortex indexes stability beyond the spatial domain. *J Neurosci* 38:1472–1481.
- Auger SD, Mullally SL, Maguire EA (2012) Retrosplenial Cortex Codes for Permanent Landmarks. *PLoS One* 7:e43620.

-
- Auger SD, Zeidman P, Maguire EA (2015) A central role for the retrosplenial cortex in de novo environmental learning. *Elife* 4:e09031.
- Auger SD, Zeidman P, Maguire EA (2017) Efficacy of navigation may be influenced by retrosplenial cortex-mediated learning of landmark stability. *Neuropsychologia* 104:102–112.
- Backus AR, Schoffelen JM, Szebényi S, Hanslmayr S, Doeller CF (2016) Hippocampal-prefrontal theta oscillations support memory integration. *Curr Biol* 26:450–457.
- Bainbridge WA, Baker CI (2020) Boundaries Extend and Contract in Scene Memory Depending on Image Properties. *Curr Biol* 30:537-543.e3.
- Baldassano C, Chen J, Zadbood A, Pillow JW, Hasson U, Norman KA (2017) Discovering event structure in continuous narrative perception and memory. *Neuron* 95:709-721.e5.
- Bar M (2003) A cortical mechanism for triggering top-down facilitation in visual object recognition. *J Cogn Neurosci* 15:600–609.
- Bar M (2004) Visual objects in context. *Nature Rev Neuro* 5:617-29.
- Bar M, Kassam KS, Ghuman AS, Boshyan J, Schmid AM, Dale AM, Hamalainen MS, Marinkovic K, Schacter DL, Rosen BR, Halgre E (2006) Top-down facilitation of visual recognition. *PNAS* 103:449–454.
- Barense MD, Henson RN, Lee AC, Graham KS (2010) Medial temporal lobe activity during complex discrimination of faces, objects, and scenes: Effects of viewpoint. *Hippocampus* 20:389–401.
- Barnes GR, Hillebrand A (2003) Statistical flattening of MEG beamformer images. *Hum Brain Mapp* 18:1–12.

-
- Barry DN, Barnes GR, Clark IA, Maguire EA (2019a) The neural dynamics of novel scene imagery. *J Neurosci* 39:4375–4386.
- Barry DN, Chadwick MJ, Maguire EA (2018) Nonmonotonic recruitment of ventromedial prefrontal cortex during remote memory recall. *PLoS Biol* 16:e2005479.
- Barry DN, Maguire EA (2019a) Consolidating the case for transient hippocampal memory traces. *Trends Cogn Sci* 23:635–636.
- Barry DN, Maguire EA (2019b) Remote memory and the hippocampus: A constructive critique. *Trends Cogn Sci* 23:128–142.
- Barry DN, Tierney TM, Holmes N, Boto E, Roberts G, Leggett J, Bowtell R, Brookes MJ, Barnes GR, Maguire EA (2019b) Imaging the human hippocampus with optically-pumped magnetoencephalography. *Neuroimage* 203:116192.
- Bartra O, McGuire JT, Kable JW (2013) The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* 76:412–427.
- Battelli L, Cavanagh P, Thornton IM (2003) Perception of biological motion in parietal patients. *Neuropsychologia* 41:1808–1816.
- Beaty RE, Thakral PP, Madore KP, Benedek M, Schacter DL (2018) Core network contributions to remembering the past, imagining the future, and thinking creatively. *J Cogn Neurosci* 30:1939–1951.
- Bechara A (2004) The role of emotion in decision-making: evidence from neurological patients with orbitofrontal damage. *Brain Cogn* 55:30–40.
- Bechara A, Tranel D, Damasio H (2000) Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain* 123:2189–2202.

-
- Beer JS, John OP, Scabini D, Knight RT (2006) Orbitofrontal cortex and social behavior: Integrating self-monitoring and emotion-cognition interactions. *J Cogn Neurosci* 18:871–879.
- Ben-Yakov A, Dudai Y (2011) Constructing realistic engrams: poststimulus activity of hippocampus and dorsal striatum predicts subsequent episodic memory. *J Neurosci* 31:9032–9042.
- Ben-Yakov A, Eshel N, Dudai Y (2013) Hippocampal immediate poststimulus activity in the encoding of consecutive naturalistic episodes. *J Exp Psychol Gen* 142:1255–1263.
- Ben-Yakov A, Henson RN (2018) The hippocampal film editor: Sensitivity and specificity to event boundaries in continuous experience. *J Neurosci* 38:10057–10068.
- Benoit RG, Szpunar KK, Schacter DL (2014) Ventromedial prefrontal cortex supports affective future simulation by integrating distributed knowledge. *Proc Natl Acad Sci USA* 111:16550–16555.
- Berens SC, Horner AJ (2017) Theta rhythm: Temporal glue for episodic memory. *Curr Biol* 27:3143–3148.
- Berens SC, Joensen BH, Horner AJ (2020) Tracking the emergence of location-based spatial representations in human scene-selective cortex. *J Cogn Neurosci* (in press).
- Berman D, Golomb JD, Walther DB (2017) Scene content is predominantly conveyed by high spatial frequencies in scene-selective visual cortex. *PLoS One* 12:e0189828.
- Bertossi E, Aleo F, Braghittoni D, Ciaramelli E (2016a) Stuck in the here and now: Construction of fictitious and future experiences following ventromedial prefrontal damage. *Neuropsychologia* 81:107–116.

-
- Bertossi E, Candela V, De Luca F, Ciaramelli E (2017) Episodic future thinking following vmPFC damage: Impaired event construction, maintenance, or narration? *Neuropsychology* 31:337–348.
- Bertossi E, Ciaramelli E (2016) Ventromedial prefrontal damage reduces mind-wandering and biases its temporal focus. *Soc Cogn Affect Neurosci* 11:1783–1791.
- Bertossi E, Tesini C, Cappelli A, Ciaramelli E (2016b) Ventromedial prefrontal damage causes a pervasive impairment of episodic memory and future thinking. *Neuropsychologia* 90:12–24.
- Binder JR, Desai RH, Graves WW, Conant LL (2009) Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex* 19:2767–2796.
- Bohbot VD, Copara MS, Gotman J, Ekstrom AD (2017) Low-frequency theta oscillations in the human hippocampus during real-world and virtual navigation. *Nat Commun* 8:14415.
- Bohbot VD, Iaria G, Petrides M (2004) Hippocampal function and spatial memory: Evidence from functional neuroimaging in healthy participants and performance of patients with medial temporal lobe resections. *Neuropsychology* 18:418–425.
- Bonnici HM, Chadwick MJ, Lutti A, Hassabis D, Weiskopf N, Maguire EA (2012) Detecting representations of recent and remote autobiographical memories in vmPFC and hippocampus. *J Neurosci* 32:16982–16991.
- Bonnici HM, Maguire EA (2018) Two years later – Revisiting autobiographical memory representations in vmPFC and hippocampus. *Neuropsychologia* 110:159–169.
- Boto E, Holmes N, Leggett J, Roberts G, Shah V, Meyer SS, Muñoz LD, Mullinger KJ, Tierney TM, Bestmann S, Barnes GR, Bowtell R, Brookes MJ (2018) Moving magnetoencephalography towards real-world applications with a wearable system. *Nature* 555:657–661.

-
- Brandman T, Peelen MV (2017) Interaction between scene and object processing revealed by human fMRI and MEG decoding. *J Neurosci* 37:0582–17.
- Brookes MJ, Vrba J, Robinson SE, Stevenson CM, Peters AM, Barnes GR, Hillebrand A, Morris PG (2008) Optimising experimental design for MEG beamformer imaging. *Neuroimage* 39:1788–1802.
- Brookes MJ, Zumer JM, Stevenson CM, Hale JR, Barnes GR, Vrba J, Morris PG (2010) Investigating spatial specificity and data averaging in MEG. *Neuroimage* 49:525–538.
- Brunec IK, Bellana B, Ozubko JD, Man V, Robin J, Liu ZX, Grady C, Rosenbaum RS, Winocur G, Barense MD, Moscovitch M (2018a) Multiple scales of representation along the hippocampal anteroposterior axis in humans. *Curr Biol* 28:2129–2135.e6.
- Brunec IK, Moscovitch M, Barense MD (2018b) Boundaries shape cognitive representations of spaces and events. *Trends Cogn Sci* 22:637–650.
- Brunec IK, Ozubko JD, Ander T, Guo R, Moscovitch M, Barense MD (2020) Turns during navigation act as boundaries that enhance spatial memory and expand time estimation. *Neuropsychologia* 141:107437.
- Bullier J (2001) Integrated model of visual processing. *Brain Res Rev* 36:96–107.
- Burgess N (2006) Spatial memory: How egocentric and allocentric combine. *Trends Cogn Sci* 10:551–557.
- Burke JF, Sharan AD, Sperling MR, Ramayya AG, Evans JJ, Healey MK, Beck EN, Davis KA, Lucas TH, Kahana MJ (2014) Theta and high-frequency activity mark spontaneous recall of episodic memories. *J Neurosci* 34:11355–11365.

-
- Bush D, Bisby JA, Bird CM, Gollwitzer S, Rodionov R, Diehl B, McEvoy AW, Walker MC, Burgess N (2017) Human hippocampal theta power indicates movement onset and distance travelled. *Proc Natl Acad Sci USA* 114:12297–12302.
- Bush G, Luu P, Posner MI (2000) Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn Sci* 4:215–222.
- Byrne P, Becker S, Burgess N (2007) Remembering the past and imagining the future: A neural model of spatial memory and imagery. *Psychol Rev* 114:340–375.
- Campbell KL, Madore KP, Benoit RG, Thakral PP, Schacter DL (2018) Increased hippocampus to ventromedial prefrontal connectivity during the construction of episodic future events. *Hippocampus* 28:76–80.
- Campo P, Poch C, Toledano R, Igoa JM, Belinchón M, García-Morales I, Gil-Nagel A (2013) Anterobasal temporal lobe lesions alter recurrent functional connectivity within the ventral pathway during naming. *J Neurosci* 33:12679–12688.
- Castelhano MS, Henderson JM (2008) The influence of color on the activation of scene gist. *J Exp Psychol Hum* 34:660–675.
- Catani M, Dell’Acqua F, Thiebaut de Schotten M (2013) A revised limbic system model for memory, emotion and behaviour. *Neurosci Biobehav Rev* 37:1724–1737.
- Catani M, Dell’Acqua F, Vergani F, Malik F, Hodge H, Roy P, Valabregue R, Thiebaut de Schotten M (2012) Short frontal lobe connections of the human brain. *Cortex* 48:273–291.
- Chadwick MJ, Mullally SL, Maguire EA (2013) The hippocampus extrapolates beyond the view in scenes: An fMRI study of boundary extension. *Cortex* 49:2067–2079.
- Chang N, Pyles JA, Marcus A, Gupta A, Tarr MJ, Aminoff EM (2019) BOLD5000, a public fMRI dataset while viewing 5000 visual images. *Sci Data* 6:49.

-
- Chang WT, Jääskeläinen IP, Belliveau JW, Huang S, Hung AY, Rossi S, Ahveninen J (2015) Combined MEG and EEG show reliable patterns of electromagnetic brain activity during natural viewing. *Neuroimage* 114:49–56.
- Chen J, Leong YC, Honey CJ, Yong CH, Norman KA, Hasson U (2017) Shared memories reveal shared structure in neural activity across individuals. *Nat Neurosci* 20:115–125.
- Chen Y, Crawford JD (2020) Allocentric representations for target memory and reaching in human cortex. *Ann NY Acad Sci* 1464:142–155.
- Chen Y, Monaco S, Crawford JD (2018) Neural substrates for allocentric-to-egocentric conversion of remembered reach targets in humans. *Eur J Neurosci* 47:901–917.
- Chiou R, Lambon Ralph MA (2016) The anterior temporal cortex is a primary semantic source of top-down influences on object recognition. *Cortex* 79:75–86.
- Ciaramelli E, Braghittoni D, Di Pellegrino G (2012) It is the outcome that counts! Damage to the ventromedial prefrontal cortex disrupts the integration of outcome and belief information for moral judgment. *J Internat Neuropsychol Soc* 18:962–971.
- Ciaramelli E, De Luca F, Monk AM, McCormick C, Maguire EA (2019) What “wins” in VMPFC: Scenes, situations, or schema? *Neurosci Biobehav Rev* 100:208–210.
- Ciaramelli E, Gheiti S, Frattarelli M, Làdavas E (2006) When true memory availability promotes false memory: Evidence from confabulating patients. *Neuropsychologia* 44:1866–1877.
- Cicmil N, Bridge H, Parker AJ, Woolrich MW, Krug K (2014) Localization of MEG human brain responses to retinotopic visual stimuli with contrasting source reconstruction approaches. *Front Neurosci* 8:127.

-
- Clark IA, Hotchin V, Monk A, Pizzamiglio G, Liefgreen A, Maguire EA (2019) Identifying the cognitive processes underpinning hippocampal-dependent tasks. *J Exp Psychol Gen* 148:1861–1881.
- Clark IA, Kim M, Maguire EA (2018) Verbal paired associates and the hippocampus: The role of scenes. *J Cogn Neurosci* 30:1821–1845.
- Clark IA, Maguire EA (2016) Remembering preservation in hippocampal amnesia. *Annu Rev Psychol* 67:51–82.
- Clark IA, Monk AM, Maguire EA (2020) Characterizing strategy use during the performance of hippocampal-dependent tasks. *Front Psychol* 11:2119.
- Clarke A, Tyler LK (2015) Understanding what we see: How we derive meaning from vision. *Trends Cogn Sci* 19:677–687.
- Clewett D, DuBrow S, Davachi L (2019) Transcending time in the brain: How event memories are constructed from experience. *Hippocampus* 29:162–183.
- Coffey EBJ, Herholz SC, Chepesiuk AMP, Baillet S, Zatorre RJ (2016) Cortical contributions to the auditory frequency-following response revealed by MEG. *Nat Commun* 7:11070.
- Cohen NJ, Eichenbaum H (1993) Memory, amnesia, and the hippocampal system. MIT Press.
- Colgin LL (2013) Mechanisms and functions of theta rhythms. *Annu Rev Neurosci* 36:295–312.
- Colgin LL (2016) Rhythms of the hippocampal network. *Nat Rev Neurosci* 17:239–249.
- Concha L, Gross DW, Beaulieu C (2005) Diffusion tensor tractography of the limbic system. *Am J Neuroradiol* 26:2267–2274.

-
- Constantinescu AO, O'Reilly JX, Behrens TEJ (2016) Organizing conceptual knowledge in humans with a gridlike code. *Science* 352:1464–1468.
- Corkin S, Amaral DG, Gonzalez RG, Johnson KA, Hyman BT (1997) H.M.'s medial temporal lobe lesion: Findings from magnetic resonance imaging. *J Neurosci* 17:3964–3979.
- Crawford JD, Henriques DYP, Medendorp WP (2011) Three-dimensional transformations for goal-directed action. *Annu Rev Neurosci* 34:309–331.
- Crespo-García M, Zeiller M, Leupold C, Kreiselmeier G, Rampp S, Hamer HM, Dalal SS (2016) Slow-theta power decreases during item-place encoding predict spatial accuracy of subsequent context recall. *Neuroimage* 142:533–543.
- Cuffin BN (1990) Effects of head shape on EEG's and MEG's. *IEEE Trans Biomed Eng* 37:44–52.
- Çukur T, Huth AG, Nishimoto S, Gallant JL (2016) Functional subdomains within scene-selective cortex: Parahippocampal place area, retrosplenial complex, and occipital place area. *J Neurosci* 36:10257–10273.
- Cushing CA, Im HY, Adams RBJ, Ward N, Albohn DN, Steiner TG (2018) Neurodynamics and connectivity during facial fear perception: The role of threat exposure and signal congruity. *Sci Rep* 8:2776.
- Cutting JE (2005) Perceiving scenes in film and in the world. In: *Moving image theory: Ecological considerations* (Anderson JD, Anderson BF, eds), pp 9–27. Carbondale, IL: University of Southern Illinois Press.
- Czigler I, Intraub H, Stefanics G (2013) Prediction beyond the borders: ERP indices of boundary extension-related error. *PLoS One* 8:e74245.
- D'Argembeau A (2013) On the role of the ventromedial prefrontal cortex in self-processing: The valuation hypothesis. *Front Hum Neurosci* 7:372.

-
- D'Argembeau A, Mathy A (2011) Tracking the construction of episodic future thoughts. *J Exp Psychol Gen* 140:258–271.
- Dalal SS, Jerbi K, Bertrand O, Adam C, Ducorps A, Schwartz D, Garnero L, Baillet S, Martinerie J, Lachaux J-P (2013) Evidence for MEG detection of hippocampus oscillations and cortical gamma-band activity from simultaneous intracranial EEG. *Epilepsy Behav* 28:310–311.
- Dalton MA, Maguire EA (2017) The pre/parasubiculum: A hippocampal hub for scene-based cognition? *Curr Opin Behav Sci* 17:34–40.
- Dalton MA, McCormick C, Maguire EA (2019) Differences in functional connectivity along the anterior-posterior axis of human hippocampal subfields. *Neuroimage* 192:38–51.
- Dalton MA, Zeidman P, Barry DN, Williams E, Maguire EA (2017) Segmenting subregions of the human hippocampus on structural magnetic resonance image scans: An illustrated tutorial. *Brain Neurosci Adv* 1:239821281770144.
- Dalton MA, Zeidman P, McCormick C, Maguire EA (2018) Differentiable processing of objects, associations, and scenes within the hippocampus. *J Neurosci* 38:8146–8159.
- De Luca F, McCormick C, Ciaramelli E, Maguire EA (2019) Scene processing following damage to the ventromedial prefrontal cortex. *Neuroreport* 30:828–833.
- De Luca F, McCormick C, Mullally SL, Intraub H, Maguire EA, Ciaramelli E (2018) Boundary extension is attenuated in patients with ventromedial prefrontal cortex damage. *Cortex* 108:1–12.
- Del Cul A, Baillet S, Dehaene S (2007) Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biol* 5:e260.

-
- Di Lollo V, Enns JT, Rensink RA (2000) Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *J Exp Psychol Gen* 129:481–507.
- Dickinson CA, Intraub H (2008) Transsaccadic representation of layout: What is the time course of boundary extension? *J Exp Psychol Hum Percept Perform* 34:543–555.
- Dietrich A, Haider H (2015) Human creativity, evolutionary algorithms, and predictive representations: The mechanics of thought trials. *Psychon Bull Rev* 22:897–915.
- Dima DC, Perry G, Singh KD (2018) Spatial frequency supports the emergence of categorical representations in visual cortex during natural scene perception. *Neuroimage* 179:102–116.
- Doeller CF, Barry C, Burgess N (2010) Evidence for grid cells in a human memory network. *Nature* 463:657–661.
- Doeller CF, King JA, Burgess N (2008) Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. *Proc Natl Acad Sci USA* 105:5915–5920.
- DuBrow S, Davachi L (2016) Temporal binding within and across events. *Neurobiol Learn Mem* 134:107–114.
- Eichenbaum H (2006) *Memory binding in hippocampal relational networks*. NY: Oxford University Press Inc.
- Eichenbaum H, Cohen NJ (2014) Can we reconcile the declarative memory and spatial navigation views on hippocampal function? *Neuron* 83:764–770.
- Ekstrom AD, Caplan JB, Ho E, Shattuck K, Fried I, Kahana MJ (2005) Human hippocampal theta activity during virtual navigation. *Hippocampus* 15:881–889.

-
- Ekstrom AD, Ranganath C (2017) Space, time, and episodic memory: The hippocampus is all over the cognitive map. *Hippocampus* 28:680–687.
- Embleton K V, Haroon HA, Morris DM, Ralph MAL, Parker GJM (2010) Distortion correction for diffusion-weighted MRI tractography and fMRI in the temporal lobes. *Hum Brain Mapp* 31:1570–1587.
- Epstein R, Kanwisher N (1998) A cortical representation the local visual environment. *Nature* 392:598–601.
- Epstein RA (2008) Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends Cogn Sci* 12:388–396.
- Epstein RA, Baker CI (2019) Scene perception in the human brain. *Annu Rev Vis Sci* 5:373–397.
- Epstein RA, Higgins JS (2007) Differential parahippocampal and retrosplenial involvement in three types of visual scene recognition. *Cereb Cortex* 17:1680–1693.
- Epstein RA, Patai EZ, Julian JB, Spiers HJ (2017) The cognitive map in humans: Spatial navigation and beyond. *Nat Neurosci* 20:1504–1513.
- Erez J, Cusack R, Kendall W, Barense MD (2016) Conjunctive coding of complex object features. *Cereb Cortex* 26:2271–2282.
- Fabre-Thorpe M (2011) The characteristics and limits of rapid visual categorization. *Front Psychol* 2:243.
- Faivre N, Dubois J, Schwartz N, Mudrik L (2019) Imaging object-scene relations processing in visible and invisible natural scenes. *Sci Rep* 9:1–13.
- Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1:1–47.

-
- Feller DP, Eerland A, Ferretti TR, Magliano JP (2019) Aspect and narrative event segmentation. *Collabra: Psychol* 5:12.
- Fellner MC, Gollwitzer S, Rampp S, Kreiselmeier G, Bush D, Diehl B, Axmacher N, Hamer H, Hanslmayr S (2019) Spectral fingerprints or spectral tilt? Evidence for distinct oscillatory signatures of memory formation. *PLoS Biol* 17:e3000403.
- Fellner MC, Volberg G, Wimber M, Goldhacker M, Greenlee MW, Hanslmayr S (2016) Spatial mnemonic encoding: Theta power decreases and medial temporal lobe BOLD increases co-occur during the usage of the method of loci. *eNeuro* 3:1–16.
- Foster DJ, Wilson MA (2007) Hippocampal theta sequences. *Hippocampus* 17:1093–1099.
- Fraiman D, Saunier G, Martins EF, Vargas CD (2014) Biological motion coding in the brain: Analysis of visually driven EEG functional networks. *PLoS One* 9:e84612.
- Friston K (2009) Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biol* 7:e1000033.
- Friston K (2010) The free-energy principle: A unified brain theory? *Nat Rev Neurosci* 11:127–138.
- Friston K, Kiebel S (2009) Predictive coding under the free-energy principle. *Philos Trans R Soc B Biol Sci* 364:1211–1221.
- Friston KJ, Daunizeau J, Kilner J, Kiebel SJ (2010) Action and behavior: A free-energy formulation. *Biol Cybern* 102:227–260.
- Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. *Neuroimage* 19:1273–1302.

-
- Friston KJ, Li B, Daunizeau J, Stephan KE (2011) Network discovery with DCM. *Neuroimage* 56:1202–1221.
- Friston KJ, Litvak V, Oswal A, Razi A, Stephan KE, Van Wijk BCM, Ziegler G, Zeidman P (2016) Bayesian model reduction and empirical Bayes for group (DCM) studies. *Neuroimage* 128:413–431.
- Fuentemilla L, Barnes GR, Düzel E, Levine B (2014) Theta oscillations orchestrate medial temporal lobe and neocortex in remembering autobiographical memories. *Neuroimage* 85:730–737.
- Fuentemilla L, Palombo DJ, Levine B (2018) Gamma phase-synchrony in autobiographical memory: Evidence from magnetoencephalography and severely deficient autobiographical memory. *Neuropsychologia* 110:7–13.
- García-Pacios J, Garcés P, Del Río D, Maestú F (2015) Early detection and late cognitive control of emotional distraction by the prefrontal cortex. *Sci Rep* 5:1–11.
- Garrido MI, Barnes GR, Kumaran D, Maguire EA, Dolan RJ (2015) Ventromedial prefrontal cortex drives hippocampal theta oscillations induced by mismatch computations. *Neuroimage* 120:362–370.
- Garrido MI, Kilner JM, Kiebel SJ, Stephan KE, Friston KJ (2007) Dynamic causal modelling of evoked potentials: A reproducibility study. *Neuroimage* 36:571–580.
- Ghosh VE, Gilboa A (2013) What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia* 53:104–114.
- Ghosh VE, Moscovitch M, Colella BM, Gilboa A (2014) Schema representation in patients with ventromedial PFC lesions. *J Neurosci* 34:12057–12070.
- Gilboa A, Alain C, Stuss DT, Melo B, Miller S, Moscovitch M (2006) Mechanisms of spontaneous confabulations: A strategic retrieval account. *Brain* 129:1399–1414.

-
- Gilboa A, Marlatte H (2017) Neurobiology of schemas and schema-mediated memory. *Trends Cogn Sci* 21:618–631.
- Gottesman CV, Intraub H (2002) Surface construal and the mental representation of scenes. *J Exp Psychol Hum Percept Perform* 28:589–599.
- Graham KS, Barense MD, Lee ACH (2010) Going beyond LTM in the MTL: A synthesis of neuropsychological and neuroimaging findings on the role of the medial temporal lobe in memory and perception. *Neuropsychologia* 48:831–853.
- Grande X, Berron D, Horner AJ, Bisby JA, Düzel E, Burgess N (2019) Holistic recollection via pattern completion involves hippocampal subfield CA3. *J Neurosci* 39:8100–8111.
- Greenberg JA, Burke JF, Haque R, Kahana MJ, Zaghoul KA (2015) Decreases in theta and increases in high frequency activity underlie associative memory encoding. *Neuroimage* 114:257–263.
- Greene MR, Hansen BC (2020) Disentangling the independent contributions of visual and conceptual features to the spatiotemporal dynamics of scene categorization. *J Neurosci* 40:5283–5299.
- Greene MR, Oliva A (2009) Recognition of natural scenes from global properties: Seeing the forest without representing the trees. *Cogn Psychol* 58:137–176.
- Gregory RL (1980) Perceptions as hypotheses. *Philos Trans R Soc London B, Biol Sci* 290:181–197.
- Griffiths BJ, Fuentemilla L (2020) Event conjunction: How the hippocampus integrates episodic memories across event boundaries. *Hippocampus* 30:162–171.
- Groen IIA, Silson EH, Baker CI (2017) Contributions of low- and high-level properties to neural processing of visual scenes in the human brain. *Philos Trans R Soc B Biol Sci* 372:20160102.

-
- Gross J, Baillet S, Barnes GR, Henson RN, Hillebrand A, Jensen O, Jerbi K, Litvak V, Maess B, Oostenveld R, Parkkonen L, Taylor JR, van Wassenhove V, Wibral M, Schoffelen JM (2013) Good practice for conducting and reporting MEG research. *Neuroimage* 65:349–363.
- Gross J, Kujala J, Hämäläinen M, Timmermann L, Schnitzler A, Salmelin R (2001) Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proc Natl Acad Sci USA* 98:694–699.
- Guderian S, Schott BH, Richardson-Klavehn A, Düzel E (2009) Medial temporal theta state before an event predicts episodic encoding success in humans. *Proc Natl Acad Sci USA* 106:5365–5370.
- Guo CC, Gorno-Tempini ML, Gesierich B, Henry M, Trujillo A, Shany-Ur T, Jovicich J, Robinson SD, Kramer JH, Rankin KP, Miller BL, Seeley WW (2013) Anterior temporal lobe degeneration produces widespread network-driven dysfunction. *Brain* 136:2979–2991.
- Gupta AS, van der Meer MAA, Touretzky DS, Redish AD (2010) Hippocampal replay is not a simple function of experience. *Neuron* 65:695–705.
- Gupta AS, van der Meer MAA, Touretzky DS, Redish AD (2012) Segmentation of spatial experience by hippocampal theta sequences. *Nat Neurosci* 15:1032–1039.
- Halai AD, Welbourne SR, Embleton K, Parkes LM (2014) A comparison of dual gradient-echo and spin-echo fMRI of the inferior temporal lobe. *Hum Brain Mapp* 35:4118–4128.
- Halder T, Talwar S, Jaiswal AK, Banerjee A (2019) Quantitative evaluation in estimating sources underlying brain oscillations using current source density methods and beamformer approaches. *eNeuro* 6:ENEURO.0170–19.2019.

-
- Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV (1993) Magnetoencephalography theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev Mod Phys* 65:413–497.
- Hansen PC, Kringelbach ML, Salmelin R (2010) *MEG: Introduction to methods*. New York: Oxford University Press.
- Hanslmayr S, Gross J, Klimesch W, Shapiro KL (2011) The role of alpha oscillations in temporal attention. *Brain Res Rev* 67:331–343.
- Hanslmayr S, Staudigl T (2014) How brain oscillations form memories – A processing based perspective on oscillatory subsequent memory effects. *Neuroimage* 85:648–655.
- Hanslmayr S, Staudigl T, Fellner M-C (2012) Oscillatory power decreases and long-term memory: The information via desynchronization hypothesis. *Front Hum Neurosci* 6:1–20.
- Harel A, Kravitz DJ, Baker CI (2013) Deconstructing visual scenes in cortex: Gradients of object and spatial layout information. *Cereb Cortex* 23:947–957.
- Harnett NG, Ference EW, Knight AJ, Knight DC (2018) White matter microstructure varies with post-traumatic stress severity following medical trauma. *Brain Imaging Behav* 14:1012–1024.
- Hartley T, Lever C, Burgess N, O'Keefe J (2013) Space in the brain: How the hippocampal formation supports spatial cognition. *Philos Trans R Soc B Biol Sci* 369:20120510–20120510.
- Hassabis D, Kumaran D, Maguire EA (2007a) Using imagination to understand the neural basis of episodic memory. *J Neurosci* 27:14365–14374.

-
- Hassabis D, Kumaran D, Vann SD, Maguire EA (2007b) Patients with hippocampal amnesia cannot imagine new experiences. *Proc Natl Acad Sci USA* 104:1726–1731.
- Hassabis D, Maguire EA (2007) Deconstructing episodic memory with construction. *Trends Cogn Sci* 11:299–306.
- Hasson U, Nir Y, Levy I, Fuhrmann G, Malach R (2004) Intersubject synchronization of cortical activity during natural vision. *Science* 303:1634–1640.
- Hayes SM, Salat DH, Verfaellie M (2012) Default network connectivity in medial temporal lobe amnesia. *J Neurosci* 32:14622–14630.
- Hebscher M, Ibrahim C, Gilboa A (2020) Precuneus stimulation alters the neural dynamics of autobiographical memory retrieval. *Neuroimage* 210:116575.
- Hebscher M, Meltzer JA, Gilboa A (2019) A causal role for the precuneus in network-wide theta and gamma oscillatory activity during complex memory retrieval. *Elife* 8:e43114.
- Heins RC, Mirza MB, Parr T, Friston K, Kagan I, Pooresmaeili A (2020) Deep active inference and scene construction. *Front Artif Intell* 3:509354.
- Henderson JM, Larson CL, Zhu DC (2008) Full Scenes produce more activation than close-up scenes and scene-diagnostic objects in parahippocampal and retrosplenial cortex: An fMRI study. *Brain Cogn* 66:40–49.
- Henson RN, Greve A, Cooper E, Gregori M, Simons JS, Geerligs L, Erzinçlioğlu S, Kapur N, Browne G (2016) The effects of hippocampal lesions on MRI measures of structural and functional connectivity. *Hippocampus* 26:1447–1463.
- Herweg NA, Kahana MJ (2018) Spatial representations in the human brain. *Front Hum Neurosci* 12:297.

-
- Herweg NA, Solomon EA, Kahana MJ (2020) Theta oscillations in human memory. *Trends Cogn Sci* 24:208–227.
- Heusser AC, Ezzyat Y, Shiff I, Davachi L (2018) Perceptual boundaries cause mnemonic trade-offs between local boundary processing and across-trial associative binding. *J Exp Psychol Learn Mem Cogn* 44:1075–1090.
- Hillebrand A, Barnes GR (2002) A quantitative assessment of the sensitivity of whole-head MEG to activity in the adult human cortex. *Neuroimage* 16:638–650.
- Hillebrand A, Barnes GR (2005) Beamformer analysis of MEG data. *Int Rev Neurobiol* 68:149–171.
- Hochstein S, Ahissar M (2002) View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron* 36:791–804.
- Hodgetts CJ, Shine JP, Lawrence AD, Downing PE, Graham KS (2016) Evidencing a place for the hippocampus within the core scene processing network. *Hum Brain Mapp* 37:3779–3794.
- Hodgetts CJ, Voets NL, Thomas AG, Clare S, Lawrence AD, Graham KS (2017) Ultra-high-field fMRI reveals a role for the subiculum in scene perceptual discrimination. *J Neurosci* 37:3150–3159.
- Holliday I, Barnes G, Hillebrand A, Singh K (2003) Accuracy and applications of group MEG studies using cortical source locations estimated from participants' scalp surfaces. *Hum Brain Mapp* 20:142–147.
- Honey CJ, Thesen T, Donner TH, Silbert LJ, Carlson CE, Devinsky O, Doyle WK, Rubin N, Heeger DJ, Hasson U (2012) Slow cortical dynamics and the accumulation of information over long timescales. *Neuron* 76:423–434.
- Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci USA* 79:2554–2558.

-
- Horner AJ, Bisby JA, Bush D, Lin WJ, Burgess N (2015) Evidence for holistic episodic recollection via hippocampal pattern completion. *Nat Commun* 6:7462.
- Horner AJ, Bisby JA, Wang A, Bogus K, Burgess N (2016) The role of spatial boundaries in shaping long-term event representations. *Cognition* 154:151–164.
- Hubbard TL, Hutchison JL, Courtney JR (2010) Boundary extension: Findings and theories. *Q J Exp Psychol* 63:1467–1494.
- Hullett PW, Hamilton LS, Mesgarani N, Schreiner CE, Chang EF (2016) Human superior temporal gyrus organization of spectrotemporal modulation tuning derived from speech stimuli. *J Neurosci* 36:2014–2026.
- Iigaya K, Hauser TU, Kurth-Nelson Z, O'Doherty JP, Dayan P, Dolan RJ (2020) The value of what's to come: Neural mechanisms coupling prediction error and the utility of anticipation. *Sci Adv* 6:eaba3828.
- Inman CS, James GA, Vytal K, Hamann S (2018) Dynamic changes in large-scale functional network organization during autobiographical memory retrieval. *Neuropsychologia* 110:208–224.
- Intraub H (2010) Rethinking scene perception: A multisource model. *Psychol Learn Motiv: Adv Res Theory* 52:231–264.
- Intraub H (2020) Searching for boundary extension. *Curr Biol* 30:R1463–R1464.
- Intraub H, Bender RS, Mangels JA (1992) Looking at pictures but remembering scenes. *J Exp Psychol Learn Mem Cogn* 18:180–191.
- Intraub H, Bodamer JL (1993) Boundary extension: Fundamental aspect of pictorial representation or encoding artifact? *J Exp Psychol Learn Mem Cogn* 19:1387–1397.

-
- Intraub H, Dickinson CA (2008) False memory 1/20th of a second later: What the early onset of boundary extension reveals about perception. *Psychol Sci* 19:1007–1014.
- Intraub H, Gottesman CV, Willey EV, Zuk IJ (1996) Boundary extension for briefly glimpsed photographs: Do common perceptual processes result in unexpected memory distortions? *J Mem Lang* 35:118–134.
- Intraub H, Hoffman JE, Wetherhold CJ, Stoehs SA (2006) More than meets the eye: The effect of planned fixations on scene representation. *Percept Psychophys* 68:759–769.
- Intraub H, Morelli F, Gagnier KM (2015) Visual, haptic and bimodal scene perception: Evidence for a unitary representation. *Cognition* 138:132–147.
- Intraub H, Richardson M (1989) Wide-angle memories of close-up scenes. *J Exp Psychol Learn Mem Cogn* 15:179–187.
- Irish M, Addis DR, Hodges JR, Piguet O (2012) Considering the role of semantic memory in episodic future thinking: Evidence from semantic dementia. *Brain* 135:2178–2191.
- Ishai A, Ungerleider LG, Haxby JV (2000) Distributed neural systems for the generation of visual images. *Neuron* 28:979–990.
- Jafarpour A, Fuentemilla L, Horner AJ, Penny W, Duzel E (2014) Replay of very early encoding representations during recollection. *J Neurosci* 34:242–248.
- Javadi A-H, Emo B, Howard LR, Zisch FE, Yu Y, Knight R, Pinelo Silva J, Spiers HJ (2017) Hippocampal and prefrontal processing of network topology to simulate the future. *Nat Commun* 8:14652.
- Jóhannesson ÓI, Thornton IM, Smith IJ, Chetverikov A, Kristjánsson Á (2016) Visual foraging with fingers and eye gaze. *Iperception* 7:1–18.

-
- Johnson A, Redish AD (2007) Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J Neurosci* 27:12176–12189.
- Josephs E, Konkle T (2020) Large-scale dissociations between views of objects, scenes, and reachable-scale environments in visual cortex. *Proc Natl Acad Sci USA* 117:29354–29362.
- Joubert OR, Rousselet GA, Fize D, Fabre-Thorpe M (2007) Processing scene context: Fast categorization and object interference. *Vision Res* 47:3286–3297.
- Julian JB, Ryan J, Epstein RA (2017) Coding of object size and object category in human visual cortex. *Cereb Cortex* 27:3095–3109.
- Kahan J, Foltynie T (2013) Understanding DCM: Ten simple rules for the clinician. *Neuroimage* 83:542–549.
- Kaplan R, Bush D, Bisby JA, Horner AJ, Meyer SS, Burgess N (2017) Medial prefrontal–medial temporal theta phase coupling in dynamic spatial imagery. *J Cogn Neurosci* 29:507–519.
- Kaplan R, Doeller CF, Barnes GR, Litvak V, Düzel E, Bandettini PA, Burgess N (2012) Movement-related theta rhythm in humans: Coordinating self-directed hippocampal learning. *PLoS Biol* 10:1–13.
- Karakaş S (2020) A review of theta oscillation and its functional correlates. *Int J Psychophysiol* 157:82–99.
- Karimpur H, Kurz J, Fiehler K (2020) The role of perception and action on the use of allocentric information in a large-scale virtual environment. *Exp Brain Res* 238:1813–1826.
- Kass RE, Raftery AE (1995) Bayes factors. *J Am Stat Assoc* 90:773.

-
- Kauffmann L, Ramanoël S, Guyader N, Chauvin A, Peyrin C (2015) Spatial frequency processing in scene-selective cortical regions. *Neuroimage* 112:86–95.
- Kiebel SJ, Garrido MI, Moran R, Chen CC, Friston KJ (2009) Dynamic causal modeling for EEG and MEG. *Hum Brain Mapp* 30:1866–1876.
- 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.
- Klein SB, Loftus J (2002) Memory and temporal experience: The effects of episodic memory loss on an amnesic patient's ability to remember the past and imagine the future. *Soc Cogn* 20:353–379.
- Klinghammer M, Schütz I, Blohm G, Fiehler K (2016) Allocentric information is used for memory-guided reaching in depth: A virtual reality study. *Vision Res* 129:13–24.
- Koenigs M, Young L, Adolphs R, Tranel D, Cushman F, Hauser M, Damasio A (2007) Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature* 446:908–911.
- König P, Wilming N, Kietzmann TC, Ossandon JP, Onat S, Ehinger B V., Gameiro RR, Kaspar K (2016) Eye movements as a window to cognitive processes. *J Eye Mov Res* 9:1–16.
- Konkel A, Warren DE, Duff MC, Tranel DN, Cohen NJ (2008) Hippocampal amnesia impairs all manner of relational memory. *Front Hum Neurosci* 2:15.
- Kopelman MD, Stanhope N, Kingsley D (1999) Retrograde amnesia in patients with diencephalic, temporal lobe or frontal lesions. *Neuropsychologia* 37:939–958.
- Kravitz DJ, Peng CS, Baker CI (2011) Real-world scene representations in high-level visual cortex: It's the spaces more than the places. *J Neurosci* 31:7322–7333.

-
- Kumaran D, Hassabis D, Spiers HJ, Vann SD, Vargha-Khadem F, Maguire EA (2007) Impaired spatial and non-spatial configural learning in patients with hippocampal pathology. *Neuropsychologia* 45:2699–2711.
- Kumaran D, Warren DE, Tranel D (2015) Damage to the ventromedial prefrontal cortex impairs learning from observed outcomes. *Cereb Cortex* 25:4504–4518.
- Kurby CA, Zacks JM (2008) Segmentation in the perception and memory of events. *Trends Cogn Sci* 12:72–79.
- Kurczek J, Wechsler E, Ahuja S, Jensen U, Cohen NJ, Tranel D, Duff M (2015) Differential contributions of hippocampus and medial prefrontal cortex to self-projection and self-referential processing. *Neuropsychologia* 73:116–126.
- Kveraga K, Ghuman AS, Bar M (2007) Top-down predictions in the cognitive brain. *Brain Cogn* 65:145–168.
- Kveraga K, Ghuman AS, Kassam KS, Aminoff EA, Hamalainen MS, Chaumon M (2011) Early onset of neural synchronization in the contextual associations network. *Proc Natl Acad Sci USA* 108:3389–3394.
- Lambon Ralph MA (2014) Neurocognitive insights on conceptual knowledge and its breakdown. *Philos Trans R Soc B Biol Sci* 369:20120392.
- Lamme VAF (2006) Towards a true neural stance on consciousness. *Trends Cogn Sci* 10:494–501.
- Lamme VAF, Roelfsema PR (2000) The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci* 23:571–579.
- Lankinen K, Saari J, Hari R, Koskinen M (2014) Intersubject consistency of cortical MEG signals during movie viewing. *Neuroimage* 92:217–224.

-
- Lankinen K, Saari J, Hlushchuk Y, Tikka P, Parkkonen L, Hari R, Koskinen M (2018) Consistency and similarity of MEG- and fMRI-signal time courses during movie viewing. *Neuroimage* 173:361–369.
- Larson AM, Freeman TE, Ringer RV, Loschky LC (2014) The spatiotemporal dynamics of scene gist recognition. *J Exp Psychol Hum Percept Perform* 40:471–487.
- Lee AC, Brodersen KH, Rudebeck SR (2013) Disentangling spatial perception and spatial memory in the hippocampus: A univariate and multivariate pattern analysis fMRI study. *J Cogn Neurosci* 25:534–546.
- Lee AC, Yeung LK, Barense MD (2012) The hippocampus and visual perception. *Front Hum Neurosci* 6:91.
- Lee ACH, Buckley MJ, Pegman SJ, Spiers H, Scahill VL, Gaffan D, Bussey TJ, Davies RR, Kapur N, Hodges JR, Graham KS (2005) Specialization in the medial temporal lobe for processing of objects and scenes. *Hippocampus* 15:782–797.
- Leopold A, Krueger F, Dal monte O, Pardini M, Pulaski SJ, Solomon J, Grafman J (2012) Damage to the left ventromedial prefrontal cortex impacts affective theory of mind. *Soc Cogn Affect Neurosci* 7:871–880.
- Leszczynski M (2011) How does hippocampus contribute to working memory processing? *Front Hum Neurosci* 5:168.
- Levens SM, Larsen JT, Bruss J, Tranel D, Bechara A, Mellers BA (2014) What might have been? The role of the ventromedial prefrontal cortex and lateral orbitofrontal cortex in counterfactual emotions and choice. *Neuropsychologia* 54:77–86.
- Levine B, Svoboda E, Hay JF, Winocur G, Moscovitch M (2002) Aging and autobiographical memory: dissociating episodic from semantic retrieval. *Psychol Aging* 17:677–689.

-
- Li FF, Iyer A, Koch C, Perona P (2007) What do we perceive in a glance of a real-world scene? *J Vis* 7:1–29.
- Li FF, VanRullen R, Koch C, Perona P (2002) Rapid natural scene categorization in the near absence of attention. *Proc Natl Acad Sci USA* 99:9596–9601.
- Liang M, Starrett MJ, Ekstrom AD (2018) Dissociation of frontal-midline delta-theta and posterior alpha oscillations: A mobile EEG study. *Psychophysiology* 55:e13090.
- Lieberman MD, Cunningham WA (2009) Type I and Type II error concerns in fMRI research: Re-balancing the scale. *Soc Cogn Affect Neurosci* 4:423–428.
- Lieberman MD, Straccia MA, Meyer ML, Du M, Tan KM (2019) Social, self, (situational), and affective processes in medial prefrontal cortex (MPFC): Causal, multivariate, and reverse inference evidence. *Neurosci Biobehav Rev* 99:311–328.
- Lin WJ, Horner AJ, Burgess N (2016) Ventromedial prefrontal cortex: Adding value to autobiographical memories. *Sci Rep* 6:28630.
- Long NM, Burke JF, Kahana MJ (2014) Subsequent memory effect in intracranial and scalp EEG. *Neuroimage* 84:488–494.
- Long NM, Kahana MJ (2015) Successful memory formation is driven by contextual encoding in the core memory network. *Neuroimage* 119:332–337.
- Love BC, Medin DL, Gureckis TM (2004) SUSTAIN: A network model of category learning. *Psychol Rev* 111:309–332.
- Luck A J (2014) *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.
- MacEvoy SP, Epstein RA (2011) Constructing scenes from objects in human occipitotemporal cortex. *Nat Neurosci* 14:1323–1329.

-
- Mack M, Love B, Preston A (2016) Dynamic updating of hippocampal object representations reflects new conceptual knowledge. *Proc Natl Acad Sci USA* 113:13203–13208.
- Mack ML, Preston AR, Love BC (2020) Ventromedial prefrontal cortex compression during concept learning. *Nat Commun* 11:46.
- Macmillan NA, Creelman CD (1990) Response bias: Characteristics of detection theory, threshold theory, and “nonparametric” indexes. *Psychol Bull* 107:401–413.
- Macmillan NA, Creelman CD (1991) *Detection theory: A user’s guide*. New York: Cambridge University Press.
- Maess B, Schröger E, Widmann A (2016) High-pass filters and baseline correction in M/EEG analysis. Commentary on: “How inappropriate high-pass filters can produce artefacts and incorrect conclusions in ERP studies of language and cognition.” *J Neurosci Methods* 266:164–165.
- Magliano JP, Zacks JM (2011) The impact of continuity editing in narrative film on event segmentation. *Med J Aust* 194:180–185.
- Maguire EA (2001) Neuroimaging studies of autobiographical event memory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 356:1441–1451.
- Maguire EA, Intraub H, Mullally SL (2015) Scenes, spaces, and memory traces: What does the hippocampus do? *Neurosci* 22:432–439.
- 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.

-
- Mante V, Sussillo D, Shenoy KV, Newsome WT (2013) Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* 503:78–84.
- Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164:177–190.
- Marr D (1971) Simple memory: A theory for archicortex. *Philos Trans R Soc Lond B Biol Sci* 262:23–81.
- Mayes AR, Holdstock JS, Isaac CL, Montaldi D, Grigor J, Gummer A, Cariga P, Downes JJ, Tsivilis D, Gaffan D, Gong Q, Norman KA (2004) Associative recognition in a patient with selective hippocampal lesions and relatively normal item recognition. *Hippocampus* 14:763–784.
- Mazziotta J et al. (2001) A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philos Trans R Soc B Biol Sci* 356:1293–1322.
- McClelland JL, McNaughton BL, O'Reilly RC (1995) Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev* 102:419–457.
- McCormick C, Barry DN, Jafarian A, Barnes GR, Maguire EA (2020) vmPFC drives hippocampal processing during autobiographical memory recall regardless of remoteness. *Cereb Cortex* 00:1–16.
- McCormick C, Ciaramelli E, De Luca F, Maguire EA (2018a) Comparing and contrasting the cognitive effects of hippocampal and ventromedial prefrontal cortex damage: A review of human lesion studies. *Neuroscience* 374:295–318.
- McCormick C, Dalton MA, Zeidman P, Maguire EA (2021) Characterising the hippocampal response to perception, construction and complexity. *Cortex* (in press).

-
- McCormick C, Moscovitch M, Valiante TA, Cohn M, McAndrews MP (2018b) Different neural routes to autobiographical memory recall in healthy people and individuals with left medial temporal lobe epilepsy. *Neuropsychologia* 110:26–36.
- McCormick C, Rosenthal CR, Miller TD, Maguire EA (2016) Hippocampal damage increases deontological responses during moral decision making. *J Neurosci* 36:12157–12167.
- McCormick C, Rosenthal CR, Miller TD, Maguire EA (2017) Deciding what is possible and impossible following hippocampal damage in humans. *Hippocampus* 27:303–314.
- McCormick C, Rosenthal CR, Miller TD, Maguire EA (2018c) Mind-wandering in people with hippocampal damage. *J Neurosci* 38:2745–2754.
- McDermott KB, Szpunar KK, Christ SE (2009) Laboratory-based and autobiographical retrieval tasks differ substantially in their neural substrates. *Neuropsychologia* 47:2290–2298.
- McNamee D, Rangel A, O'Doherty JP (2013) Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nat Neurosci* 16:479–485.
- Melo B, Winocur G, Moscovitch M (1999) False recall and false recognition: An examination of the effects of selective and combined lesions to the medial temporal lobe/diencephalon and frontal lobe structures. *Cogn Neuropsychol* 16:343–359.
- Meyer SS, Rossiter H, Brookes MJ, Woolrich MW, Bestmann S, Barnes GR (2017) Using generative models to make probabilistic statements about hippocampal engagement in MEG. *Neuroimage* 149:468–482.
- Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24:167–202.

-
- Miller TD, Chong TTJ, Davies AMA, Johnson MR, Irani SR, Husain M, Ng TWC, Jacob S, Maddison P, Kennard C, Gowland PA, Rosenthal CR (2020) Human hippocampal CA3 damage disrupts both recent and remote episodic memories. *Elife* 9:e41836.
- Mills T, Lalancette M, Moses SN, Taylor MJ, Quraan MA (2012) Techniques for detection and localization of weak hippocampal and medial frontal sources using beamformers in MEG. *Brain Topogr* 25:248–263.
- Mirza MB, Adams RA, Mathys CD, Friston KJ (2016) Scene construction, visual foraging, and active inference. *Front Comput Neurosci* 10:56.
- Moran R, Pinotsis DA, Friston K (2013) Neural masses and fields in dynamic causal modeling. *Front Comput Neurosci* 7:57.
- Moran RJ, Stephan KE, Seidenbecher T, Pape HC, Dolan RJ, Friston KJ (2009) Dynamic causal models of steady-state responses. *Neuroimage* 44:796–811.
- Morgan LK, MacEvoy SP, Aguirre GK, Epstein RA (2011) Distances between real-world locations are represented in the human hippocampus. *J Neurosci* 31:1238–1245.
- Moscovitch M (1989) Confabulation and the frontal systems: Strategic versus associative retrieval in neuropsychological theories of memory. In: *Varieties of memory and consciousness: Essays in honor of Endel Tulving*. (Roediger HL, Craik FI, eds), pp 133–160. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Moscovitch M, Melo B (1997) Strategic retrieval and the frontal lobes: Evidence from confabulation and amnesia. *Neuropsychologia* 35:1017–1034.
- Moulton ST, Kosslyn SM (2011) Imagining predictions: Mental imagery as mental emulation. *Philos Trans R Soc B Biol Sci* 364:1273–1280.
- Mullally SL, Hassabis D, Maguire EA (2012a) Scene construction in amnesia: An fMRI study. *J Neurosci* 32:5646–5653.

-
- Mullally SL, Intraub H, Maguire EA (2012b) Attenuated boundary extension produces a paradoxical memory advantage in amnesic patients. *Curr Biol* 22:261–268.
- Mullally SL, Maguire EA (2011) A new role for the parahippocampal cortex in representing space. *J Neurosci* 31:7441–7449.
- Mullally SL, Maguire EA (2013) Exploring the role of space-defining objects in constructing and maintaining imagined scenes. *Brain Cogn* 82:100–107.
- Mullally SL, Maguire EA (2014) Counterfactual thinking in patients with amnesia. *Hippocampus* 24:1261–1266.
- Murakami S, Okada Y (2006) Contributions of principal neocortical neurons to magnetoencephalography and electroencephalography signals. *J Physiol* 575:925–936.
- Murdaugh DL, Deshpande HD, Kana RK (2016) The impact of reading intervention on brain responses underlying language in children with autism. *Autism Res* 9:141–154.
- Navawongse R, Eichenbaum H (2013) Distinct pathways for rule-based retrieval and spatial mapping of memory representations in hippocampal neurons. *J Neurosci* 33:1002–1013.
- Nicolás B, Sala-Padró J, Cucurell D, Santurino M, Falip M, Fuentemilla L (2021) Theta rhythm supports hippocampus-dependent integrative encoding in schematic/semantic memory networks. *Neuroimage* 226:117558.
- Nolte G (2003) The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Phys Med Biol* 48:3637–3652.

-
- Norman KA, O'Reilly RC (2003) Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychol Rev* 110:611–646.
- Northoff G, Bermpohl F (2004) Cortical midline structures and the self. *Trends Cogn Sci* 8:102–107.
- 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 University Press.
- Ólafsdóttir HF, Bush D, Barry C (2018) The role of hippocampal replay in memory and planning. *Curr Biol* 28:R37–R50.
- Oliva A (2005) Gist of the scene. In: *Neurobiology of Attention*. (Itti L, Rees G, Tsotsos JK, eds), pp 251–256. Academic Press.
- Oliva A, Torralba A (2006) Building the gist of a scene: The role of global image features in recognition. *Prog Brain Res* 155:23–36.
- Olsen RK, Carr VA, Daugherty AM, La Joie R, Amaral RSC, Amunts K, Augustinack JC, Bakker A, Bender AR, Berron D, Boccardi M, Bocchetta M, Burggren AC, Chakravarty MM, Chetelat G, de Flores R, DeKraaker J, Ding S-L, Wisse LEM (2019) Progress update from the hippocampal subfields group. *Alzheimer's Dement Diagnosis, Assess Dis Monit* 11:439–449.
- 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.

-
- Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869.
- Palombo DJ, Hayes SM, Peterson KM, Keane MM, Verfaellie M (2018) Medial temporal lobe contributions to episodic future thinking: Scene construction or future projection? *Cereb Cortex* 28:447–458.
- Papadelis C, Leonardelli E, Staudt M, Braun C (2012) Can magnetoencephalography track the afferent information flow along white matter thalamo-cortical fibers? *Neuroimage* 60:1092–1105.
- 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.
- Park S, Intraub H, Yi DJ, Widders D, Chun MM (2007) Beyond the edges of a view: Boundary extension in human scene-selective visual cortex. *Neuron* 54:335–342.
- Parkkonen L, Fujiki N, Mäkelä JP (2009) Sources of auditory brainstem responses revisited: Contribution by magnetoencephalography. *Hum Brain Mapp* 30:1772–1782.
- Parr T, Rikhye RV, Halassa MM, Friston KJ (2020) Prefrontal computation as active inference. *Cereb Cortex* 30:682–695.
- Patai EZ, Javadi AH, Ozubko JD, O’Callaghan A, Ji S, Robin J, Grady C, Winocur G, Rosenbaum RS, Moscovitch M, Spiers HJ (2019) Hippocampal and retrosplenial goal distance coding after long-term consolidation of a real-world environment. *Cereb Cortex* 29:2748–2758.
- Patterson K, Nestor PJ, Rogers TT (2007) Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat Rev Neurosci* 8:976–987.

-
- Peelen MV, Caramazza A (2012) Conceptual object representations in human anterior temporal cortex. *J Neurosci* 32:15728–15736.
- Pitcher D, Ianni G, Ungerleider LG (2019) A functional dissociation of face-, body- and scene-selective brain areas based on their response to moving and static stimuli. *Sci Rep* 9:8242.
- Poch C, Fuentemilla L, Barnes GR, Duzel E (2011) Hippocampal theta-phase modulation of replay correlates with configural-relational short-term memory performance. *J Neurosci* 31:7038–7042.
- Popov T, Oostenveld R, Schoffelen JM (2018) FieldTrip made easy: An analysis protocol for group analysis of the auditory steady state brain response in time, frequency, and space. *Front Neurosci* 12:711.
- Poppenk J, Evensmoen HR, Moscovitch M, Nadel L (2013) Long-axis specialization of the human hippocampus. *Trends Cogn Sci* 17:230–240.
- Poser BA, Norris DG (2009) Investigating the benefits of multi-echo EPI for fMRI at 7 T. *Neuroimage* 45:1162–1172.
- Potter MC (1975) Meaning in visual search. *Science* 187:965–966.
- Potter MC, Wyble B, Haggmann CE, McCourt ES (2014) Detecting meaning in RSVP at 13 ms per picture. *Atten, Percep, Psychophys* 76:270–279.
- Pscherer C, Bluschke A, Prochnow A, Eggert E, Mückschel M, Beste C (2020) Resting theta activity is associated with specific coding levels in event-related theta activity during conflict monitoring. *Hum Brain Mapp* 41:5114–5127.
- Pu Y, Cheyne DO, Cornwell BR, Johnson BW (2018) Non-invasive investigation of human hippocampal rhythms using magnetoencephalography: A review. *Front Neurosci* 12:1–16.

Radvansky GA (2012) Across the event horizon. *Psychol Sci* 21:269–272.

Ramkumar P, Hansen BC, Pannasch S, Loschky LC (2016) Visual information representation and rapid-scene categorization are simultaneous across cortex: An MEG study. *Neuroimage* 134:295–304.

Ranganath C, Ritchey M (2012) Two cortical systems for memory-guided behaviour. *Nat Rev Neurosci* 13:713–726.

Razi A, Seghier ML, Zhou Y, McColgan P, Zeidman P, Park HJ, Sporns O, Rees G, Friston KJ (2017) Large-scale DCMs for resting-state fMRI. *Netw Neurosci* 1:222–241.

Rempl-Clower NL, Barbas H (2000) Laminar pattern of connections between prefrontal and anterior temporal cortices in the rhesus monkey is related to cortical structure and function. *Cereb Cortex* 10:851–865.

Richmond LL, Zacks JM (2017) Constructing experience: Event models from perception to action. *Trends Cogn Sci* 21:1-962–980.

Riggs L, Moses SN, Bardouille T, Herdman AT, Ross B, Ryan JD (2009) A complementary analytic approach to examining medial temporal lobe sources using magnetoencephalography. *Neuroimage* 45:627–642.

Ritchey M, Montchal ME, Yonelinas AP, Ranganath C (2015) Delay-dependent contributions of medial temporal lobe regions to episodic memory retrieval. *Elife* 4:e05025.

Rizzolatti G, Craighero L (2004) The mirror-neuron system. *Annu Rev Neurosci* 27:169–192.

Roberts G, Holmes N, Boto E, Leggett J, Hill RM, Shah V, Alexander N, Rea M, Vaughan R, Maguire EA, Kessler K, Beebe S, Fromhold M, Barnes GR, Bowtell R, Brookes

-
- MJ (2019) Towards magnetoencephalography in a virtual reality environment. *Neuroimage* 199:408–417.
- Robin J, Buchsbaum BR, Moscovitch M (2018) The primacy of spatial context in the neural representation of events. *J Neurosci* 38:2755–2765.
- Robin J, Hirshhorn M, Rosenbaum RS, Winocur G, Moscovitch M, Grady CL (2015) Functional connectivity of hippocampal and prefrontal networks during episodic and spatial memory based on real-world environments. *Hippocampus* 25:81–93.
- Robin J, Olsen RK (2019) Scenes facilitate associative memory and integration. *Learn Mem* 26:252–261.
- Rodriguez PF (2010) Neural decoding of goal locations in spatial navigation in humans with fMRI. *Hum Brain Mapp* 31:391–397.
- Rolls ET (2013) The mechanisms for pattern completion and pattern separation in the hippocampus. *Front Syst Neurosci* 7:74.
- Rosenbaum RS, Gilboa A, Levine B, Winocur G, Moscovitch M (2009) Amnesia as an impairment of detail generation and binding: Evidence from personal, fictional, and semantic narratives in K.C. *Neuropsychologia* 47:2181–2187.
- Rosenbaum RS, Moscovitch M, Foster JK, Schnyer DM, Gao F, Kovacevic N, Verfaellie M, Black SE, Levine B (2008) Patterns of autobiographical memory loss in medial- temporal lobe amnesic patients. *J Cogn Neurosci* 21:1490–1506.
- Rosenbaum RS, Priselac S, Köhler S, Black SE, Gao F, Nadel L, Moscovitch M (2000) Remote spatial memory in an amnesic person with extensive bilateral hippocampal lesions. *Nat Neurosci* 3:1044–1048.
- Rousselet GA, Joubert OR, Fabre-Thorpe M (2005) How long to get to the “gist” of real-world natural scenes? *Vis cogn* 12:852–877.

-
- Roy M, Shohamy D, Wager TD (2012) Ventromedial prefrontal-subcortical systems and the generation of affective meaning. *Trends Cogn Sci* 16:147–156.
- Rudebeck SR, Filippini N, Lee ACH (2013) Can complex visual discrimination deficits in amnesia be attributed to the medial temporal lobe? An investigation into the effects of medial temporal lobe damage on brain connectivity. *Hippocampus* 23:7–13.
- Rushworth MFS, Noonan MAP, Boorman ED, Walton ME, Behrens TE (2011) Frontal cortex and reward-guided learning and decision-making. *Neuron* 70:1054–1069.
- Ruzich E, Crespo-García M, Dalal SS, Schneiderman JF (2019) Characterizing hippocampal dynamics with MEG: A systematic review and evidence-based guidelines. *Hum Brain Mapp* 40:1353–1375.
- Sakata H, Kusunoki M (1992) Organization of space perception: Neural representation of three-dimensional space in the posterior parietal cortex. *Curr Opin Neurobiol* 2:170–174.
- Sans-Dublanc A, Mas-Herrero E, Marco-Pallarés J, Fuentemilla L (2017) Distinct neurophysiological mechanisms support the online formation of individual and across-episode memory representations. *Cereb Cortex* 27:4314–4325.
- Saygin AP, Wilson SM, Hagler DJ, Bates E, Sereno MI (2004) Point-light biological motion perception activates human premotor cortex. *J Neurosci* 24:6181–6188.
- Schacter DL, Addis DR (2007) On the constructive episodic simulation of past and future events. *Behav Brain Sci* 30:330–332.
- Schacter DL, Addis DR (2020) Memory and imagination: Perspectives on constructive episodic simulation. In *The Cambridge Handbook of the Imagination* (ed. Abrahams A), pp. 119–131. Cambridge University Press.

-
- Schacter DL, Addis DR, Hassabis D, Martin VC, Spreng RN, Szpunar KK (2012) The future of memory: Remembering, imagining, and the brain. *Neuron* 76:677–694.
- Schlichting ML, Preston AR (2015) Memory integration: Neural mechanisms and implications for behavior. *Curr Opin Behav Sci* 1:1–8.
- Schoenbaum G, Takahashi Y, Liu TL, McDannald MA (2011) Does the orbitofrontal cortex signal value? *Ann NY Acad Sci* 1239:87–99.
- Schuck NW, Cai MB, Wilson RC, Niv Y (2016) Human orbitofrontal cortex represents a cognitive map of state space. *Neuron* 91:1402–1412.
- Scoville WB, Milner B (1957) Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatry* 20:11–21.
- Seamon JG, Schlegel SE, Hiester PM, Landau SM, Blumenthal BF (2002) Misremembering pictured objects: People of all ages demonstrate the boundary extension illusion. *Am J Psychol* 115:151–167.
- Shigeto H, Morioka T, Hisada K, Nishio S, Ishibashi H, Kira DI, Tobimatsu S, Kato M (2002) Feasibility and limitations of magnetoencephalographic detection of epileptic discharges: Simultaneous recording of magnetic fields and electrocorticography. *Neurol Res* 24:531–536.
- Shin YS, DuBrow S (2020) Structuring memory through inference-based event segmentation. *Top Cogn Sci* (in press).
- Silva M, Baldassano C, Fuentemilla L (2019) Rapid memory reactivation at movie event boundaries promotes episodic encoding. *J Neurosci* 39:8538–8548.
- Solomon EA, Kragel JE, Sperling MR, Sharan A, Worrell G, Kucewicz M, Inman CS, Lega B, Davis KA, Stein JM, Jobst BC, Zaghoul KA, Sheth SA, Rizzuto DS, Kahana MJ (2017) Widespread theta synchrony and high-frequency desynchronization underlies enhanced cognition. *Nat Commun* 8:1704.

-
- Solomon EA, Stein JM, Das S, Gorniak R, Sperling MR, Worrell G, Inman CS, Tan RJ, Jobst BC, Rizzuto DS, Kahana MJ (2019) Dynamic theta networks in the human medial temporal lobe support episodic memory. *Curr Biol* 29:1100–1111.
- Sols I, DuBrow S, Davachi L, Fuentemilla L (2017) Event boundaries trigger rapid memory reinstatement of the prior events to promote their representation in long-term memory. *Curr Biol* 27:3499–3504.e4.
- Spalding KN, Jones SH, Duff MC, Tranel D, Warren DE (2015) Investigating the neural correlates of schemas: Ventromedial prefrontal cortex is necessary for normal schematic influence on memory. *J Neurosci* 35:15746–15751.
- Spanò G, Intraub H, Edgin JO (2017) Testing the “boundaries” of boundary extension: Anticipatory scene representation across development and disorder. *Hippocampus* 27:726–739.
- Spanò G, Pizzamiglio G, McCormick C, Clark IA, De Felice S, Miller TD, Edgin JO, Rosenthal CR, Maguire EA (2020) Dreaming with hippocampal damage. *Elife* 9:e56211.
- Speer NK, Reynolds JR, Swallow KM, Zacks JM (2009) Reading stories activates neural representations of visual and motor experiences. *Psychol Sci* 20:989–999.
- Speer NK, Swallow KM, Zacks JM (2003) Activation of human motion processing areas during event perception. *Cogn Affect Behav Neurosci* 3:335–345.
- Spiers HJ (2020) The hippocampal cognitive map: One space or many? *Trends Cogn Sci* 24:168–170.
- Spiers HJ, Maguire EA (2006) Thoughts, behaviour, and brain dynamics during navigation in the real world. *Neuroimage* 31:1826–1840.

-
- Spreng RN, Mar RA, Kim ASN (2009) The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. *J Cogn Neurosci* 21:489–510.
- Squire LR (1992) Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol Rev* 99:195–231.
- Squire LR, Van der Horst AS, McDuff SGR, Frascino JC, Hopkins RO, Mauldin KN (2010) Role of the hippocampus in remembering the past and imagining the future. *Proc Natl Acad Sci USA* 107:19044–19048.
- St-Laurent M, Moscovitch M, Tau M, McAndrews MP (2011) The temporal unraveling of autobiographical memory narratives in patients with temporal lobe epilepsy or excisions. *Hippocampus* 21:409–421.
- Staresina BP, Davachi L (2009) Mind the gap: Binding experiences across space and time in the human hippocampus. *Neuron* 63:267–276.
- Staudigl T, Hanslmayr S (2013) Theta oscillations at encoding mediate the context-dependent nature of human episodic memory. *Curr Biol* 23:1101–1106.
- Steinvorth S, Levine B, Corkin S (2005) Medial temporal lobe structures are needed to re-experience remote autobiographical memories: Evidence from H.M. and W.R. *Neuropsychologia* 43:479–496.
- Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ (2009) Bayesian model selection for group studies. *Neuroimage* 46:1004–1017.
- Stephan KE, Penny WD, Moran RJ, den Ouden HEM, Daunizeau J, Friston KJ (2010) Ten simple rules for dynamic causal modeling. *Neuroimage* 49:3099–3109.
- Strange BA, Witter MP, Lein ES, Moser EI (2014) Functional organization of the hippocampal longitudinal axis. *Nat Rev Neurosci* 15:655–669.

-
- Summerfield C, Egner T (2009) Expectation (and attention) in visual cognition. *Trends Cogn Sci* 13:403–409.
- Summerfield JJ, Hassabis D, Maguire EA (2010) Differential engagement of brain regions within a “core” network during scene construction. *Neuropsychologia* 48:1501–1509.
- Svoboda E, McKinnon MC, Levine B (2006) The functional neuroanatomy of autobiographical memory: A meta-analysis. *Neuropsychologia* 44:2189–2208.
- Swallow KM, Kemp JT, Candan Simsek A (2018) The role of perspective in event segmentation. *Cognition* 177:249–262.
- Swallow KM, Zacks JM, Abrams RA (2009) Event boundaries in perception affect memory encoding and updating. *J Exp Psychol Gen* 138:236–257.
- Tan ES (2018) A psychology of the film. *Palgrave Commun* 4:82.
- Teng E, Squire LR (1999) Memory for places learned long ago is intact after hippocampal damage. *Nature* 400:675–677.
- Teyler TJ, DiScenna P (1986) The hippocampal memory indexing theory. *Behav Neurosci* 100:147–154.
- Teyler TJ, Rudy JW (2007) The hippocampal indexing theory and episodic memory: Updating the index. *Hippocampus* 17:1158–1169.
- Thompson PM, Schwartz C, Lin RT, Khan AA, Toga AW (1996) Three-dimensional statistical analysis of sulcal variability in the human brain. *J Neurosci* 16:4261–4274.
- Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. *Am J Ophthalmol* 122:608–609.

-
- Thorpe SJ, Gegenfurtner KR, Fabre-Thorpe M, Bülthoff HH (2001) Detection of animals in natural images using far peripheral vision. *Eur J Neurosci* 14:869–876.
- Tolman EC (1948) Cognitive maps in rats and men. *Psychol Rev* 55:189–208.
- Tranel D, Jones RD (2006) Knowing “what” and knowing “when.” *J Clin Exp Neuropsychol* 28:43–66.
- Trapp S, Bar M (2015) Prediction, context, and competition in visual recognition. *Ann NY Acad Sci* 1339:190–198.
- Troiani V, Stigliani A, Smith ME, Epstein RA (2014) Multiple object properties drive scene-selective regions. *Cereb Cortex* 24:883–897.
- Tse D, Langston RF, Kakeyama M, Bethus I, Spooner PA, Wood ER, Witter MP, Morris RGM (2007) Schemas and memory consolidation. *Science* 316:76–82.
- Tulving E (1985) Memory and consciousness. *Can Psychol* 26:1–12.
- Tulving E (2002) Episodic memory: From mind to brain. *Annu Rev Psychol* 53:1–25.
- Vaidya AR, Fellows LK (2015) Ventromedial frontal cortex is critical for guiding attention to reward-predictive visual features in humans. *J Neurosci* 35:12813–12823.
- van Kesteren MTR, Beul SF, Takashima A, Henson RN, Ruiter DJ, Fernández G (2013) Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: From congruent to incongruent. *Neuropsychologia* 51:2352–2359.
- van Kesteren MTR, Fernandez G, Norris DG, Hermans EJ (2010a) Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proc Natl Acad Sci USA* 107:7550–7555.

-
- van Kesteren MTR, Rijpkema M, Ruiter DJ, Fernández G (2010b) Retrieval of associative information congruent with prior knowledge is related to increased medial prefrontal activity and connectivity. *J Neurosci* 30:15888–15894.
- van Kesteren MTR, Ruiter DJ, Fernández G, Henson RN (2012) How schema and novelty augment memory formation. *Trends Neurosci* 35:211–219.
- Van Veen BD, Van Drongelen W, Yuchtman M, Suzuki A (1997) Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans Biomed Eng* 44:867–880.
- Vann SD, Aggleton JP, Maguire EA (2009) What does the retrosplenial cortex do? *Nat Rev Neurosci* 10:792–802.
- VanRullen R, Thorpe SJ (2001) Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artificial objects. *Perception* 30:655–668.
- Viskontas IV, McAndrews MP, Moscovitch M (2000) Remote episodic memory deficits in patients with unilateral temporal lobe epilepsy and excisions. *J Neurosci* 20:5853–5857.
- Von Der Heide RJ, Skipper LM, Klobusicky E, Olson IR (2013) Dissecting the uncinate fasciculus: Disorders, controversies and a hypothesis. *Brain* 136:1692–1707.
- Wager TD, Keller MC, Lacey SC, Jonides J (2005) Increased sensitivity in neuroimaging analyses using robust regression. *Neuroimage* 26:99–113.
- Warren DE, Duff MC, Jensen U, Tranel D, Cohen NJ (2012) Hiding in plain view: Lesions of the medial temporal lobe impair online representation. *Hippocampus* 22:1577–1588.
- Warren DE, Jones SH, Duff MC, Tranel D (2014) False recall is reduced by damage to the ventromedial prefrontal cortex: Implications for understanding the neural correlates of schematic memory. *J Neurosci* 34:7677–7682.

-
- Weilbacher RA, Gluth S (2016) The interplay of hippocampus and ventromedial prefrontal cortex in memory-based decision making. *Brain Sci* 7:4.
- Wikenheiser AM, Schoenbaum G (2016) Over the river, through the woods: Cognitive maps in the hippocampus and orbitofrontal cortex. *Nat Rev Neurosci* 17:513–523.
- Winecoff A, Clithero JA, Carter RMK, Bergman SR, Wang L, Huettel SA (2013) Ventromedial prefrontal cortex encodes emotional value. *J Neurosci* 33:11032–11039.
- Wisse LEM, Gerritsen L, Zwanenburg JJM, Kuijf HJ, Luijten PR, Biessels GJ, Geerlings MI (2012) Subfields of the hippocampal formation at 7 T MRI: In vivo volumetric assessment. *Neuroimage* 61:1043–1049.
- Wolbers T, Büchel C (2005) Dissociable retrosplenial and hippocampal contributions to successful formation of survey representations. *J Neurosci* 25:3333–3340.
- Wolbers T, Wiener JM, Mallot HA, Büchel C (2007) Differential recruitment of the hippocampus, medial prefrontal cortex, and the human motion complex during path integration in humans. *J Neurosci* 27:9408–9416.
- Wolters CH, Anwender A, Tricoche X, Weinstein D, Koch MA, MacLeod RS (2006) Influence of tissue conductivity anisotropy on EEG/MEG field and return current computation in a realistic head model: A simulation and visualization study using high-resolution finite element modeling. *Neuroimage* 30:813–826.
- Woollett K, Maguire EA (2010) The effect of navigational expertise on wayfinding in new environments. *J Environ Psychol* 30:565–573.
- Yamada H, Louie K, Tymula A, Glimcher PW (2018) Free choice shapes normalized value signals in medial orbitofrontal cortex. *Nat Commun* 9:162.

-
- Yang SCH, Lengyel M, Wolpert DM (2016) Active sensing in the categorization of visual patterns. *Elife* 5:e12215.
- Yomogida Y, Sugiura M, Watanabe J, Akitsuki Y, Sassa Y, Sato T, Matsue Y, Kawashima R (2004) Mental visual synthesis is originated in the fronto-temporal network of the left hemisphere. *Cereb Cortex* 14:1376–1383.
- Yonelinas AP (2013) The hippocampus supports high-resolution binding in the service of perception, working memory and long-term memory. *Behav Brain Res* 254:34–44.
- Yonelinas AP, Ranganath C, Ekstrom AD, Wiltgen BJ (2019) A contextual binding theory of episodic memory: Systems consolidation reconsidered. *Nat Rev Neurosci* 20:364–375.
- Zacks JM, Speer NK, Swallow KM, Braver TS, Reynolds JR (2007) Event perception: A mind-brain perspective. *Psychol Bull* 133:273–293.
- Zacks JM, Tversky B (2001) Event structure in perception and conception. *Psychol Bull* 127:3–21.
- Zacks JM, Tversky B, Iyer G (2001) Perceiving, remembering, and communicating structure in events. *J Exp Psychol Gen* 130:29–58.
- Zedelius CM, Schooler JW (2016) The richness of inner experience: Relating styles of daydreaming to creative processes. *Front Psychol* 6:2063.
- Zeidman P, Lutti A, Maguire EA (2015a) Investigating the functions of subregions within anterior hippocampus. *Cortex* 73:240–256.
- Zeidman P, Maguire EA (2016) Anterior hippocampus: The anatomy of perception, imagination and episodic memory. *Nat Rev Neurosci* 17:173–182.

Zeidman P, Mullally SL, Maguire EA (2015b) Constructing, perceiving, and maintaining scenes: Hippocampal activity and connectivity. *Cereb Cortex* 25:3836–3855.

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.

Zeithamova D, Dominick AL, Preston AR (2012) Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron* 75:168–179.