

Extrinsic reference frames modify the neural substrates of object-location representations

Edgar Chan^{a,b}, Oliver Baumann^{a,*}, Mark A. Bellgrove^c, Jason B. Mattingley^{a,b}

^a The University of Queensland, Queensland Brain Institute, St Lucia 4072, Australia

^b The University of Queensland, School of Psychology, St Lucia 4072, Australia

^c Monash University, School of Psychology and Psychiatry, Clayton 3800, Australia

ARTICLE INFO

Article history:

Received 7 January 2013

Received in revised form

7 February 2013

Accepted 9 February 2013

Available online 17 February 2013

Keywords:

Allocentric

fMRI

Navigation

Reference frames

Spatial memory

Viewpoint

ABSTRACT

The ability to form spatial representations of object locations is an important component of successful spatial navigation. Evidence from behavioral studies suggests that environmental features that have a salient coordinate axis (e.g., a rectangular building or a geometrical room) may provide a reference frame for the encoding of object-location information. Here we used functional magnetic resonance imaging (fMRI) to determine the brain networks engaged when object-location representations are stored with respect to an extrinsic reference frame. Participants learned the layout of an object array in an active, virtual-navigation paradigm. A square mat positioned on the floor of the virtual arena acted as the extrinsic reference frame. Knowledge of the spatial arrangement of the object array was probed while participants underwent fMRI, using a spatial judgment task that required them to imagine orientations of the learned array that were either aligned or misaligned with the geometry of the mat. Consistent with previous findings, participants responded faster and were more accurate when the imagined orientation was aligned, as opposed to misaligned, with the extrinsic reference frame. Analysis of the fMRI data revealed important differences in brain activity between the two conditions. Significantly greater activity was observed in the aligned condition compared with the misaligned condition across a bilateral network of brain areas that included the inferior occipital gyri, inferior and middle temporal gyri, and fusiform gyri. By contrast, activity in the misaligned condition was significantly greater than in the aligned condition in bilateral dorsolateral prefrontal and anterior cingulate cortex, and in the right anterior prefrontal and anterior insular cortex. These results suggest that retrieval of spatial locations that are aligned with an extrinsic reference frame involve direct access to detailed and accurate representations within the ventral visual pathway, whereas spatial locations that are misaligned with this reference frame are only weakly represented and require active inferential processes through the recruitment of prefrontal cortical networks. Our findings are consistent with a “reference direction” account of spatial memory, which posits that inter-object spatial relationships are primarily encoded with respect to specified reference directions.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

In humans, successful spatial navigation depends largely upon our ability to accurately encode object-location information. When planning a journey or giving directions, we are often required to access a stored representation of relevant spatial information that is independent of our current location and viewing perspective, i.e., an *allocentric* representation. Although it is generally accepted that medial temporal brain regions, particularly the hippocampus, are important for the formation of allocentric representations, the manner in which this

information is stored and organized in the brain is less clear. While some theories postulate that stored allocentric representations are orientation-independent and akin to a “mental map” (Byrne, Becker, & Burgess, 2007; Sholl, 2001), others suggest that allocentric representations are based upon available cognitive heuristics such as extrinsic reference frames (Mou, McNamara, Valiquette, & Rump, 2004; Shelton & McNamara, 2001). In the current study, we used fMRI to determine the neural correlates of object-location memory after participants had learned a virtual object array during active spatial navigation. Our findings are consistent with theories suggesting that the geometry of an environmental landmark exerts a significant modulatory effect on object-location representations.

Spatial information acquired during navigation can be encoded with respect to one or more different frames of reference (Galati,

* Corresponding author. Tel.: +61 7 3346 3305; fax: +61 7 3346 6301.

E-mail addresses: o.baumann@uq.edu.au, oliverbaumann@yahoo.de (O. Baumann).

Pelle, Berthoz, & Committeri, 2010; Klatzky, 1998). In humans, a distinction is commonly drawn between an *egocentric* reference frame, which is established with respect to oneself, and an *allocentric* reference frame, which is defined with respect to other objects in the environment. Egocentric and allocentric reference frames are supported by distinct neural networks involving parietal and medial temporal regions, respectively (Byrne et al., 2007; Feigenbaum & Morris, 2004; Gramann, Muller, Eick, & Schonebeck, 2005; Gramann, Muller, Schonebeck, & Debus, 2006; Pizzamiglio, Guariglia, & Cosentino, 1998; Zaehle et al., 2007). In keeping with *cognitive map theory* (Tolman, 1948), which originated from rodent research and the discovery of hippocampal place cells (O'Keefe & Nadel, 1978), several models of human spatial memory have posited that allocentric representations are stored in an orientation-independent manner within the hippocampus (Byrne et al., 2007; Sholl, 2001). Object-to-object spatial relationships within an array are assumed to be represented with equal salience in memory. Retrieval of spatial information for navigation requires transformation from an allocentric to an egocentric reference frame via a single neural circuit involving posterior midline and parietal brain regions (e.g. Byrne et al., 2007).

As an alternative to the orientation-independent account outlined above, the *reference direction theory* proposes that allocentric object-location information is represented with respect to salient reference frames (Mou et al., 2004; Shelton & McNamara, 2001). Such reference frames can be defined by salient coordinate axes that are based on the geometric structure of the surrounding environment, such as the geometry of a room or a prominent landmark (i.e., *extrinsic* reference frames). Evidence supporting the reference direction theory comes from behavioral studies that have shown that judgments of learned-object locations are faster and more accurate when the retrieved perspective is *aligned* with the extrinsic reference frame (i.e., parallel to the defined axis) than when it is *misaligned* – the so-called “*alignment effect*” (Kelly & McNamara, 2008; McNamara, Rump, & Werner, 2003; Shelton & McNamara, 2001). Although the performance benefit for aligned versus misaligned heading perspectives is taken as a proxy for the increased demands of retrieving misaligned study views, the cognitive and neural bases of the alignment effect are yet to be fully explained.

McNamara and colleagues have suggested that only inter-object spatial relationships that are specified with respect to a particular spatial reference system are explicitly encoded in memory (Mou et al., 2004; Shelton & McNamara, 2001). Retrieval of spatial relations from *misaligned* perspectives is thus assumed to rely upon mental transformation of the stored, aligned perspectives. This mental transformation process is thought to be the source of the observed speed and accuracy costs for misaligned perspectives. It should be noted, however, that the greater demands of retrieving misaligned views might also arise because misaligned spatial relations are explicitly represented in memory, but in a weaker or lower fidelity form. Larger judgement errors for misaligned perspectives could be due to increased noise associated with the relatively low-fidelity spatial representations of misaligned views, rather than to any mental transformation from an aligned view stored in memory. On this account, any increase in retrieval time for misaligned views would be due to increased demands on higher-level decision making processes. Thus, although this alternative account makes identical predictions in terms of behavioral effects (i.e., slower and more error-prone responses in the misaligned condition than the aligned condition), it assumes different underlying neural processes.

In the present study, we aimed to determine the neural substrates of the behavioral alignment effect. Participants learned the spatial layout of a 7-item object-array while actively navigating

a virtual environment that included a salient landmark that acted as an extrinsic reference frame. Memory for the array was tested 24 h after learning while participants underwent fMRI. During scanning, participants performed a spatial judgment task that required them to imagine taking a specific perspective within the array that was either aligned or misaligned with the extrinsic reference frame. We predicted greater activity in brain areas related to visuospatial retrieval, such as visual association areas and inferior occipito-temporal cortices (Roland & Gulyas, 1994; Slotnick, 2004), during recall of *aligned* compared with *misaligned* perspectives, reflecting efficient retrieval of accurate and detailed spatial representations. In contrast, the original account by McNamara and colleagues would not predict enhanced retrieval-related activity for the aligned condition, since this account assumes that the same representation is retrieved in the aligned and misaligned conditions. Furthermore, we predicted greater activity in brain areas associated with mental reasoning, such as the dorsolateral prefrontal cortex (Duncan & Owen, 2000), for retrieving *misaligned* versus *aligned* perspectives, consistent with the hypothesis that higher level reasoning processes would be required to infer inter-object relations from weak spatial representations arising from misaligned heading directions. By contrast, if retrieval from misaligned directions involves a spatial transformation process that translates object-location information encoded from aligned perspectives for retrieval of misaligned perspectives, we would expect elevated activity in brain regions commonly associated with mental transformation, such as the posterior parietal cortex (Gogos et al., 2010; Milivojevic, Hamm, & Corballis, 2009; Zacks, 2008).

2. Methods

2.1. Participants

Twenty healthy, adult volunteers gave their informed consent to participate in the study, which was approved by Human Research Ethics Committee of The University of Queensland. Data from one participant were excluded from the analysis due to artefacts. The remaining 19 participants (12 females) ranged in age from 17 to 31 years (mean age = 22 years). All participants were right-handed and had normal or corrected to normal vision.

2.2. Design and procedures

We developed an active navigation paradigm that was performed within a virtual arena to examine the influence of extrinsic reference frames on the neural encoding of object locations. Participants were required to learn the spatial layout of a seven-object array by actively navigating the virtual arena from a first-person perspective using a handheld joystick. A square “mat” positioned on the floor of the virtual arena acted as the extrinsic reference frame for the environment. Neural responses were measured using fMRI while participants performed a spatial judgment task that required them to imagine perspectives of the learned object-array that were either *aligned* or *misaligned* with the reference frame, and to point in the direction of a target object using the joystick.

In the *learning phase* participants were required to navigate within the virtual arena (radius = 23 virtual meters, m) with the goal of learning the spatial locations of the seven distinct objects. Common fruits were chosen as the objects as they contain no intrinsic spatial information and are visually familiar and distinctive. The objects were hidden beneath identical circular gray pedestals (radius = 0.7 m; height = 0.5 m) so that participants were only ever able to see one target object at a time during learning. The array was spatially arranged in a symmetrical configuration so that the objects could be perceptually grouped vertically, horizontally or diagonally (Fig. 1a). A square blue mat (18 m × 18 m) that encompassed the array provided the extrinsic reference frame to the environment; viewpoints parallel to the sides of the mat were defined as *aligned* viewpoints, and those diagonal to the sides of the mat were defined as *misaligned*. The arena was void of any other distinctive visual cues. The virtual environment was created using the Blender open source 3D content creation suite (The Blender Foundation, Amsterdam, Netherlands) and presented on a 21.5 inch LCD monitor. Fig. 1b provides a survey view of the virtual arena. (Note that participants were never shown this survey view during the study.)

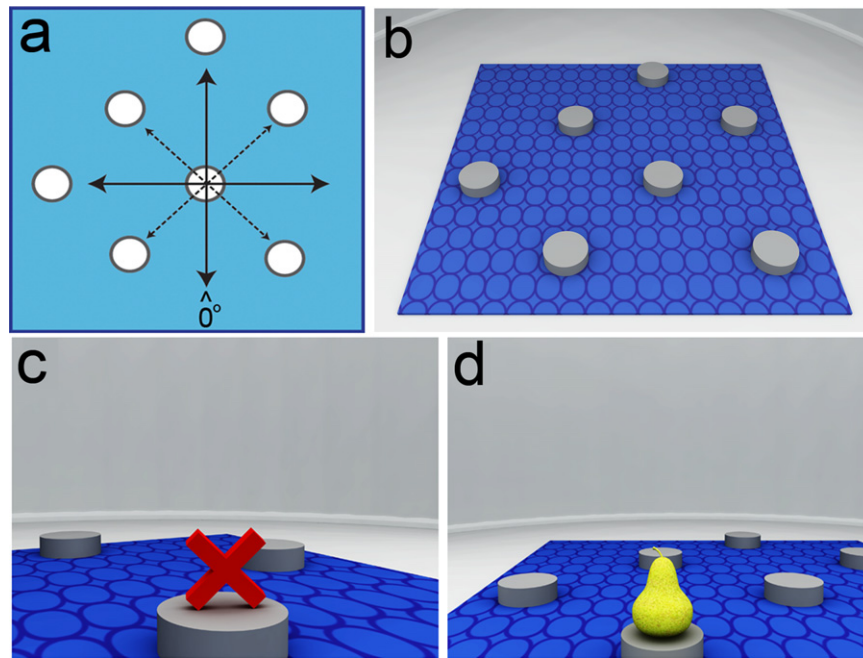


Fig. 1. Schematic of the virtual arena used in the study and snapshots of the arena as seen by participants during the learning phase. (a) Spatial arrangement of the object locations used in the study (white circles). With respect to the square mat, solid lines represent the aligned coordinate axes and dashed lines represent the misaligned coordinate axes. (b) Aerial perspective of the virtual arena through which participants navigated during the learning phase of the task. The gray pedestals marked the seven possible locations of the target objects. Participants never saw an aerial view of the environment during the experiment. (c) First person view of the arena, as experienced by the participants. When an incorrect location was probed during a trial, a red cross appeared immediately above the pedestal. (d) When participants arrived at the correct location, the probed target object (the pear in this example) appeared above the pedestal.

At the beginning of each trial participants were cued (1500 ms) to the identity of one of the seven target objects (apple, pear, grapes, cherries, banana, lemon, strawberry) and instructed to search for that target as quickly and directly as possible. Participants always commenced each new trial from the center of the arena. They used a joystick to navigate throughout the arena from a first-person perspective, with a maximum velocity of 3 m/s. Participants were able to ‘query’ the location of an object once they had moved within a 3 m radius of a pedestal, by pressing a button on the joystick. Each query resulted in visual feedback in the form of a red cross above the pedestal in the case of an incorrect response (Fig. 1c), or the appearance of the target object in the case of a correct response (Fig. 1d). The trial terminated when the target object was revealed and the next trial began immediately afterwards. Each of the seven target objects was cued in a random order without replacement before an object was cued again. The learning phase was discontinued when participants were able to accurately locate each target object without error in three of the last four trials. Participants were instructed to build a mental representation of the object locations, as this would help them during the test phase. Participants were never explicitly told to encode spatial locations with reference to the mat.

In the *retrieval phase* we measured neural responses associated with object-location memory with fMRI while participants completed a judgement of relative direction (JRD) task. On each trial, participants had to imagine themselves standing at the location of a given object within the arena, facing another object, and then to point to the location of a third object (see Fig. 2). Thus, for example, on a particular trial the participant might receive an instruction such as: “Imagine standing at the cherry, facing the banana; now point to the pear.” These instructions were delivered using pictures of the objects, rather than words, and participants were practiced with the format and sequence of the instructions beforehand. In the initial *imagined heading* screen participants viewed the two objects defining the imagined standing and facing positions; in the subsequent *pointing* screen they saw the target object toward which they had to point, as if they were inside the arena. Critically, each object-pair of the *imagined heading* screen was selected to elicit a viewpoint that was either *aligned* or *misaligned* with the extrinsic reference frame (the square mat), although this was never specified to participants during the task. The object pairs comprised only of objects that were directly adjacent to each other within the array.

Each trial began with a fixation cross (3000–5000 ms) followed by the critical *imagined heading* screen containing the two objects that defined an *aligned* or *misaligned* perspective. Participants were instructed to press a response button once they had recalled the relevant imagined location within the arena (maximum 10 s). As soon as participants pressed the button, a blank screen was presented for 500 ms, after which the *pointing* phase followed. In the pointing phase the participants were shown the target object toward which they had to point as accurately as possible, using a joystick held in their right hand (maximum 3 s).

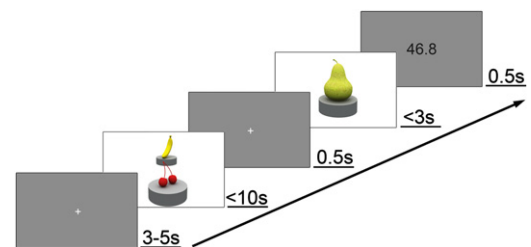


Fig. 2. Sequence of events during a typical trial of the test phase. Each trial began with a fixation cross (3000–5000 ms) followed by the critical *imagined heading* screen containing the two objects that defined an *aligned* or *misaligned* perspective. In the *imagined heading* screen, the larger fruit at the bottom of the screen represented the imagined standing location (the cherry in this example) and the smaller fruit at the top of the screen represented the imagined facing location (the banana). Participants were instructed to press a response button once they had recalled the relevant imagined location within the arena (maximum 10 s). As soon as participants pressed the button, a blank screen was presented for 500 ms. In the *pointing* display, participants had to indicate the imagined direction of the target object (the pear in this example), as if they were standing within the virtual arena (maximum allowed time 3 s). As soon as a response was made, a *feedback* screen (500 ms) provided participants with feedback on pointing accuracy (46.8° in this example).

As soon as a response was made, a *feedback* screen (500 ms) provided participants with information about the accuracy of their response in angular degrees (i.e., the absolute difference in the recorded pointing direction and the actual direction of the target object given the specified heading). The same imagined heading angle was never probed on two consecutive trials. In the scanner, participants completed a total of 2×80 unique trials (40 aligned, 40 misaligned), split into four runs. Trials were presented using Presentation software (Neurobehavioral Systems; <http://www.neurobs.com>). Response latency (i.e., the time in seconds that participants spent viewing the *imagined heading* screen) and pointing accuracy (i.e., angular error in degrees) were recorded for subsequent analysis.

It is important to note that although the pointing component of the task provided an essential behavioral probe, the critical cognitive process of determining one's imagined heading occurs during the initial “heading” phase. We therefore optimized the design of our fMRI experiment to capture activity associated with the retrieval of stored heading information. We also anticipated that the motor response associated with the pointing task would introduce noise due to

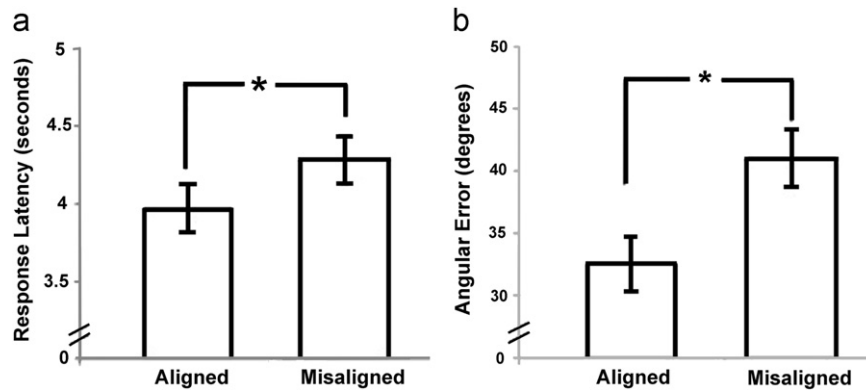


Fig. 3. Behavioral responses from the test phase of the study, obtained during fMRI scanning, showing performance for the aligned and misaligned conditions. (a) Mean response latency and (b) mean angular error. * $p < 0.01$.

movement artifacts and inter-participant variability. We therefore temporally separated the pointing period from the heading period by jittering the inter-period onset interval, and modeled the two periods separately. The heading and pointing periods were *self-paced*; thus the duration of both periods varied considerably from trial to trial (mean variance for the heading period = 2.47 s, mean variance for the pointing period = 0.18 s). Therefore, even though the duration of the blank screen between the two periods was constant, the inter-period onset interval varied from trial to trial (i.e., it contained a participant-determined jitter). In addition, we pseudo-randomized the order of the aligned and the misaligned conditions, to further increase the amount of jitter between the heading periods of aligned versus misaligned conditions. The sequencing of the trials, and the temporal jittering of rest periods, was optimized with optseq2 software (<http://surfer.nmr.mgh.harvard.edu/optseq/>).

Participants were trained on both the learning and retrieval components of the task one day prior to the fMRI session to ensure that they understood all task instructions and were able to complete the tasks within the designated time limits. The spatial configuration of the object array was identical for the training and fMRI session for each participant, to facilitate memory consolidation.

2.3. MRI acquisition and data analysis

Whole brain echo-planar imaging (EPI) was conducted on a 3T MR scanner (Trio; Siemens, Erlangen, Germany) with a 12-channel head coil. Thirty-three axial slices (slice thickness, 3.3 mm) were acquired in an interleaved order, using a gradient echo echo-planar T2*-sensitive sequence (repetition time, 2 s; echo time, 30 ms; flip angle, 90°; matrix, 64 × 64; field of view, 210 × 210 mm; voxel size (spacing), 3.3 × 3.3 × 3.3 mm). A T1-weighted structural MPRAGE was also acquired. A liquid crystal display projector (1024 × 768 resolution) back-projected the test-phase trials onto a screen positioned at the head of the scanner gantry. Participants lay on their backs within the bore of the magnet and viewed the stimuli via a 45° angled mirror that reflected the images displayed on the screen. The distance to the screen was 90 cm (12 cm from eyes to mirror) and the visible part of the screen encompassed approximately 22.0° × 16.4° of visual angle (35.5 × 26.0 cm). All participants were stabilized with tightly packed foam padding surrounding the head to minimize head movement.

Image processing and statistical analyses were performed using SPM5 (Wellcome Department of Imaging Neuroscience, UCL, London, UK). Functional data volumes were slice-time corrected and realigned to the first volume. A T2*-weighted mean image of the unsmoothed images was co-registered with the corresponding anatomical T1-weighted image from the same individual. The individual T1-image was used to derive the transformation parameters for stereotaxic space using the SPM5 segmentation algorithm (Ashburner & Friston, 2005), which was then applied to the individual co-registered EPI images. The voxel sizes of the normalized images were 2 mm³. Images were then smoothed with an 8-mm full-width half maximum (FWHM) isotropic Gaussian kernel. At the single-participant level, we applied a high-pass filter to remove baseline drifts. We anticipated that the motor response associated with the pointing task would likely introduce noise due to movement artifacts. To minimize the effect of the motor response in the *pointing* period on the neural correlates of the *heading* period we temporally separated the pointing period from the heading period by jittering the inter-period onset interval and modeling both periods separately as boxcar functions convolved with a hemodynamic response function (HRF). We modeled the exact duration of each individual heading and pointing period in SPM so that the height and duration of the corresponding hemodynamic response functions were scaled accordingly for every trial.

Our principal comparison involved the effect of *aligned* relative to *misaligned* imagined headings. Statistical parametric maps of the *t*-statistic were generated for each participant from linear contrasts of the conditions *Aligned* > *Misaligned*.

These contrasts of parameter estimates were then included in a second-level group analysis using single-sample *t*-tests on the contrast images obtained from each participant. Whole-brain statistical maps were assessed on a cluster-level, using a height threshold of $p = 0.005$. These were considered significant at $p < 0.05$ using a family-wise error correction (FWE).

3. Results

3.1. Behavioral results

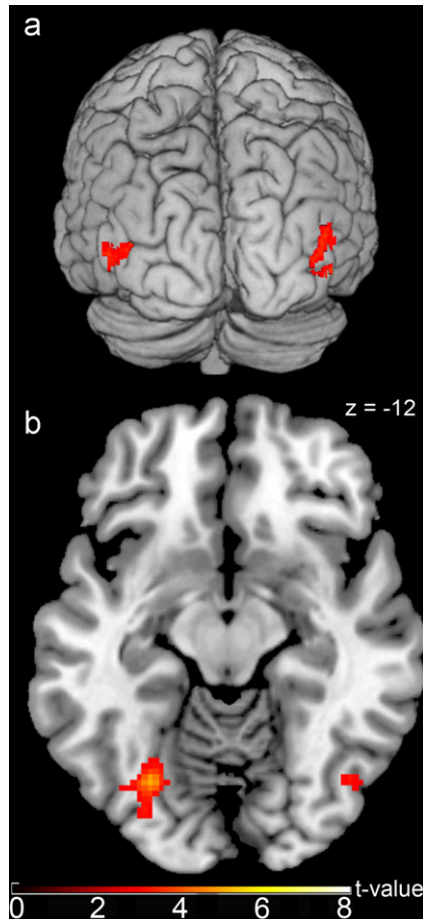
Participants completed an average of 72 learning trials (range: 60–88) within the virtual arena prior to the *test* phase within the scanner. In the learning phase, each target location was probed approximately seven times, and it took participants on average 8.48 seconds (SD = 1.75) to complete each trial. To examine whether there was any behavioral evidence of an alignment effect during the test phase, we measured and analyzed the time participants spent viewing the *imagined heading* screen and the absolute angular error participants made in their pointing response. Participants' mean response latency and angular error for the two alignment conditions are plotted in Fig. 3a and b, respectively. As predicted, participants spent less time viewing the *imagined heading* screen and were more accurate in their pointing responses for trials in which the heading orientation defined by the target objects was *aligned* as opposed to when it was *misaligned* with the extrinsic reference frame. These observations were confirmed statistically using a two-tailed, paired samples *t*-test, which showed significant differences in both response latency ($t(18) = -3.60$, $p < 0.01$) and angular error ($t(18) = -6.468$, $p < 0.01$) between the two alignment conditions.

3.2. Imaging results

Statistical analyses were carried out on the fMRI data to identify differences in BOLD activity between the aligned and misaligned conditions during the *imagined heading* period. According to McNamara and colleagues (Mou et al., 2004; Shelton & McNamara, 2001), the alignment effect arises from differences in retrieval of stored representations during the heading phase of the JRD task. Therefore, while the pointing component of the task provided a performance measure that allowed us to verify a behavioral alignment effect, the critical period for quantifying neural responses for aligned versus misaligned directions was the heading phase. Statistical parametric maps of the *t* statistic were generated from linear contrasts of aligned > misaligned, separately for each participant. These contrasts of parameter estimates were then included in a second-level group analysis using single-sample *t* tests on the contrast images obtained from each participant. This resulted in whole-brain, random effects for the comparison of aligned > misaligned and

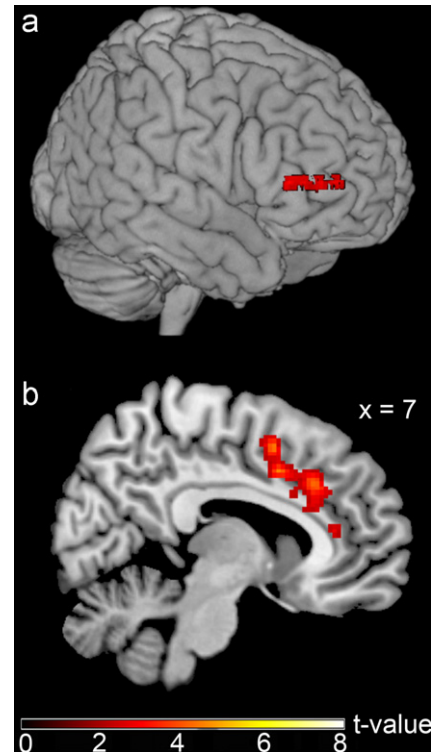
Table 1Summary of fMRI findings during the heading phase for the contrast Aligned > Misaligned and Misaligned > Aligned ($p \leq 0.05$, corrected for multiple comparisons).

Region	Hemisphere	Brodmann area	MNI coordinates			T-values/z-values of maxima (cluster size in number of voxels)
			x	y	z	
Aligned > Misaligned						
V2/V3/Fusiform gyrus	L	18/19/39	−28	−70	−12	5.08/3.95 (372)
V3/Fusiform gyrus/Angular gyrus	R	19/37/39	48	−72	−10	3.77/3.19 (259)
Misaligned > Aligned						
Dorsolateral prefrontal cortex/anterior cingulate cortex	L + R	9/24/32/33	4	8	50	5.18/4.00 (1191)
Anterior prefrontal cortex/anterior insula	R	10/13	46	30	10	4.96/3.89 (269)

**Fig. 4.** Mean BOLD activity for the contrast aligned > misaligned. Activations, displayed on an MNI normalized template brain, show significant bilateral activation in middle and inferior occipito-temporal regions. (a) Rendered posterior view and (b) axial view.

misaligned > aligned. Whole-brain statistical maps were assessed at the cluster-level, using a height threshold of $p = 0.005$, and were considered significant at $p < 0.05$ using a family-wise error correction (FWE).

We hypothesized greater activity in visuospatial brain regions during retrieval of *aligned* inter-object relationships. Activity in these regions should reflect fast and efficient retrieval of accurate and detailed spatial representations (Roland & Gulyas, 1994; Slotnick, 2004). In contrast, we hypothesized weaker activity in visuospatial brain regions during retrieval of *misaligned* inter-object relationships, reflecting retrieval of weak, noisy or otherwise low-fidelity

**Fig. 5.** Mean BOLD activity for the contrast misaligned > aligned. Activations, displayed on an MNI normalized template brain, show significant activation in anterior cingulate and right dorsolateral prefrontal regions. (a) Rendered lateral view and (b) sagittal view.

memory representations. Consistent with this, the results of the random-effects analysis of activity across the whole brain for the contrast aligned > misaligned revealed two significant activation clusters involving the inferior occipital, inferior and middle temporal, and fusiform gyri of the left and right hemispheres (Table 1; Fig. 4a,b).

We also hypothesized that higher level reasoning processes would be employed to infer inter-object relations from the weak spatial representations arising from misaligned heading directions. In line with this prediction, whole brain analysis for the contrast misaligned > aligned revealed distinct clusters of activation in bilateral dorsolateral prefrontal cortex and anterior cingulate, and in the right anterior prefrontal cortex extending to the right anterior insula (Table 1; Fig. 5a and b). No other activation clusters exceeded the statistical threshold. Finally, to assess the relative contributions of the aligned and misaligned conditions within the identified visual and frontal brain regions, we

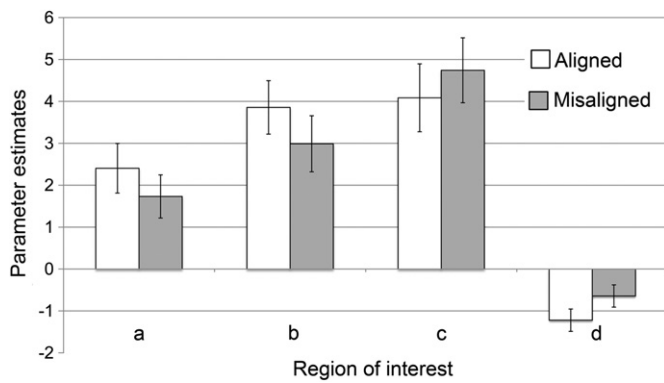


Fig. 6. Parameter estimates (beta values, ± 1 standard error) for the peak voxels from the analysis of aligned versus misaligned conditions (separately for the two conditions relative to the implicit baseline). (a) Fusiform gyrus ($-28, -70, -12$); (b) middle occipital gyrus ($48, -72, -10$); (c) medial frontal gyrus ($4, 8, 50$); (d) inferior frontal gyrus ($46, 30, 10$).

extracted the parameter estimates separately for both conditions relative to implicit baseline (Fig. 6).

4. Discussion

The principal aim of the present study was to examine the influence of extrinsic reference frames on the neural representation of object-location information. Our central finding is that the neural areas recruited for object-location retrieval are modulated as a function of alignment with an extrinsic reference frame. Contrary to the notion that allocentric spatial representations in long-term memory are orientation-independent (Byrne et al., 2007; Sholl, 2001), our findings are instead consistent with the theory that learned inter-object spatial relationships are organized around available salient extrinsic reference frames (Mou et al., 2004; Shelton & McNamara, 2001). More broadly, our findings imply that retrieval of imagined headings that are aligned with an extrinsic reference frame involve direct access to detailed and accurate representations within the ventral visual pathway, whereas spatial perspectives that are misaligned with this reference frame are only weakly represented and require active inferential processes through the recruitment of prefrontal cortical networks.

We found significantly greater activation in bilateral inferior and middle occipito-temporal regions when participants imagined perspectives within the object-array that were aligned with the extrinsic reference frame available to them during navigation-based learning of the virtual array. Although activity within posterior ventral pathways has typically been associated with visual object identification (Kohler, Kapur, Moscovitch, Winocur, & Houle, 1995), recent investigations have suggested that these regions are also important for the retrieval and processing of visual memories (Slotnick, 2004) and object features (Pietrini et al., 2004). In addition, it has been shown that activation within the ventral pathways might be important for the retrieval of spatial location information (Thompson, Slotnick, Burrage, & Kosslyn, 2009). Greater activation in inferior and middle occipito-temporal regions for the aligned condition is likely to reflect the efficient retrieval of accurate and detailed representations of inter-object relations. Previous studies have suggested that posterior ventral brain regions might be important in the retrieval of stored spatial information, particularly with reference to environmental landmarks. For example, in a recent fMRI study (Morgan, MacEvoy, Aguirre, & Epstein, 2011), activity in occipito-temporal regions was found when participants were shown

photographs of familiar environmental landmarks. In that study, participants reported at post-test that the photographs prompted them to visualize themselves standing at the depicted location. Committeri et al. (2004) examined neural differences between environment-referenced (landmark-based) spatial representations and object-referenced (object-based) spatial representations and found specific involvement of ventral occipital regions, including the lingual and fusiform gyri, only when participants made spatial judgments in the environment-referenced condition. Furthermore, lesions of these areas are associated with specific impairments in retrieval of otherwise intact spatial representations, which in turn causes topographical disorientation (Barrash, 1998).

In contrast to the aligned condition, imagining perspectives from within the object-array that were misaligned with the extrinsic reference frame activated a network that included bilateral anterior cingulate (ACC) and dorsolateral prefrontal cortex (DLPFC), as well as the right anterior prefrontal cortex and anterior insula. These frontal regions are frequently active under conditions of uncertainty and mental effort (Botvinick, Cohen, & Carter, 2004), cognitive control (Manenti, Cotelli, Calabria, Maioli, & Miniussi, 2010; MacDonald, Cohen, Stenger, & Carter, 2000) and the coordination of information processing (Ramnani & Owen, 2004). Enhanced activity within these regions in the misaligned condition might therefore imply the operation of higher-order reasoning. Similar regions of activation were found in other studies that used spatial judgment tasks requiring visual imagery (Kukolja, Marshall, & Fink, 2006; Lee et al., 2005). Our data, however, are not consistent with a parietal-mediated transformation process, as predicted by the classical model of reference direction (Mou et al., 2004; Shelton & McNamara, 2001).

Consistent with prior studies (Kelly & McNamara, 2008; Shelton & McNamara, 2001; Valiquette, McNamara, & Labrecque, 2007), our behavioral data suggest that participants were both faster and more accurate in retrieving learned object-locations when their imagined heading was aligned with the extrinsic reference frame than when it was misaligned with this reference frame. The classical model by McNamara and colleagues proposed these differences in retrieval speed and accuracy arise because misaligned inter-object relationships have to be *inferred* from the stored representation via a process of mental transformation (Mou et al., 2004; Shelton & McNamara, 2001). An alternative view, however, is that differences in retrieval accuracy might be due to the fact that misaligned inter-object relationships are more weakly represented in memory, as indexed by lower activity in occipital visual areas. Our results further suggest that participants' slower responses during retrieval of misaligned perspectives are due to the engagement of frontally mediated inferential processes. This interpretation is in line with findings from several previous studies that have related mental effort (Botvinick et al., 2004) and cognitive control (Manenti et al., 2010; MacDonald et al., 2000) to activity in the dorsolateral prefrontal cortex and the anterior cingulate cortex.

Taken together, our findings are consistent with the suggestion that spatial relationships (bearings or directions) between objects in an environment are represented in memory in terms of reference directions specified by environmental reference frames (Mou et al., 2004; Shelton & McNamara, 2001). We have provided the first evidence for distinct neural networks underpinning the classical alignment effect, by demonstrating that retrieval of spatial locations that are aligned with an extrinsic reference frame involve direct access to detailed and accurate representations via the ventral visual pathways, whereas spatial locations that are misaligned with this reference frame are only weakly represented and require active inferential processes through the recruitment of prefrontal cortical networks. On the other hand,

our data are inconsistent with the notion that judgments from misaligned perspectives involve the retrieval of representations from aligned perspectives, which are subsequently transformed via a mental rotation process.

In the current study, an extrinsic reference frame was established using a landmark that contained salient and well-defined coordinate axes. We chose an extrinsic landmark because previous fMRI studies have demonstrated that environmental landmark properties can modulate neural activity during allocentric spatial processing (Bird, Capponi, King, Doeller, & Burgess, 2010; Doeller, King, & Burgess, 2008; Janzen & van Turenhout, 2004). Other behavioral studies have demonstrated that reference frames can also be established using only intrinsic properties of the object array, such as the spatial layout (Mou & McNamara, 2002) or verbal instructions (Greenauer & Waller, 2008). Using fMRI, Xiao, McNamara, Qin, and Mou (2010) investigated how an intrinsic reference frame might modulate scene recognition for an object array. In that study, participants were asked to learn the spatial layout of objects placed on a table in specified columns, thus creating an intrinsic reference frame, for a subsequent fMRI scene-recognition task. Recognition of scenes containing groups of objects that were misaligned with the intrinsic reference frame was associated with increased activity in the right intraparietal sulcus and deactivation of the anterior cingulate. No specific areas of activity were associated with scenes containing objects within the same column, despite better behavioral performance in this condition.

It is unclear why the patterns of neural activity we observed are different from those reported by Xiao et al. (2010), but it is worth noting that the tasks used by us and by Xiao et al. were very different, both at encoding and retrieval. Perhaps scene recognition and spatial imagery involve separable and distinct retrieval processes (Valiquette & McNamara, 2007). Alternatively, intrinsic and extrinsic reference frames may have differential effects on the neural representation of object-location information despite yielding similar behavioral alignment effects. Future research should investigate how reference frames might modulate neural activity with different forms of encoding and retrieval tasks. Furthermore, while our experiment revealed the neural correlates of retrieving aligned versus misaligned perspectives, it was not optimized for identifying any differences in pointing-related activity. According to McNamara and colleagues (Mou et al., 2004; Shelton & McNamara, 2001), the alignment effect arises from differences in retrieval of stored representations during the heading phase of the JRD task. It is possible, however, that further processes during the pointing phase contribute to the alignment effect. In addition to motor-related activations, the neural correlates during the pointing phase are likely to be associated with the computation of egocentric object vectors, which might differ for aligned and misaligned perspectives. To enable a complete understanding of the neural underpinnings of the alignment effect, future studies should endeavor to dissociate the computation of object vectors from the manual response component. Finally, it remains an open question whether extrinsic reference frames, in addition to biasing memory representations, also exert an influence on navigation behavior. Future studies should therefore examine whether human participants show a preference for navigating along pathways that are aligned with salient landmarks. Such a finding would shed further light on the origin of the representational bias caused by environmental geometry.

In summary, we have shown that the presence of an extrinsic reference frame during navigation can bias the organization of allocentric object-location representations, and modulate associated neural systems involved in subsequent retrieval processes. Our findings provide an important addendum to the widely held reference-frame account (Mou et al., 2004; Shelton & McNamara,

2001) by showing that heading information that is aligned with an extrinsic reference frame involves detailed and accurate representations within the ventral visual pathway, whereas headings misaligned with this reference frame are only weakly represented and require active inferential processes through the recruitment of prefrontal cortical networks. More broadly, our findings may contribute to a better understanding of topographical difficulties experienced in both neurological patient groups and healthy individuals. Lesions in identified occipito-temporal regions may cause impairments in the direct access and processing of relevant spatial information. This in turn could result in greater reliance on frontal brain networks for making spatial judgments relevant for navigation. In addition, impairments in the use of reference frames to organize new spatial information may contribute to symptoms seen in topographical disorientation, such as difficulty in remembering spatial relationships between relevant landmarks (Aguirre & D'Esposito, 1999; Barrash, 1998; Brunsdon, Nickels, & Coltheart, 2007). Further imaging and neuropsychological studies will be necessary to better understand the various contributions extrinsic reference frames have on spatial navigation in daily life.

Acknowledgments

This work was supported by an Australian Research Council (ARC) Thinking Systems Grant. J.B.M. was supported by an Australian Research Council Laureate Fellowship (FL110100103) and O.B. was supported by an Australian Research Council Discovery Early Career Researcher Award (DE120100535). We gratefully acknowledge the Thinking Systems Team for their support, and in particular Mark Wakabayashi for programming the virtual environment used in the study.

References

- Aguirre, G. K., & D'Esposito, M. (1999). Topographical disorientation: A synthesis and taxonomy. *Brain*, 122, 1613–1628.
- Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *Neuroimage*, 26, 839–851.
- Barrash, J. (1998). A historical review of topographical disorientation and its neuroanatomical correlates. *Journal of Clinical and Experimental Neuropsychology*, 20, 807–827.
- Bird, C. M., Capponi, C., King, J. A., Doeller, C. F., & Burgess, N. (2010). Establishing the boundaries: The hippocampal contribution to imagining scenes. *Journal of Neuroscience*, 30, 11688–11695.
- Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Science*, 8, 539–546.
- Brunsdon, R., Nickels, L., & Coltheart, M. (2007). Topographical disorientation: Towards an integrated framework for assessment. *Neuropsychological Rehabilitation*, 17, 34–52.
- Byrne, P., Becker, S., & Burgess, N. (2007). Remembering the past and imagining the future: A neural model of spatial memory and imagery. *Psychological Review*, 114, 340–375.
- Committeri, G., Galati, G., Paradis, A. L., Pizzamiglio, L., Berthoz, A., & LeBihan, D. (2004). Reference frames for spatial cognition: Different brain areas are involved in viewer-, object-, and landmark-centered judgments about object location. *Journal of Cognitive Neuroscience*, 16, 1517–1535.
- Doeller, C. F., King, J. A., & Burgess, N. (2008). Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. *Proceedings of the National Academy of Science of the United States of America*, 105, 5915–5920.
- Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neuroscience*, 23, 475–483.
- Feigenbaum, J. D., & Morris, R. G. (2004). Allocentric versus egocentric spatial memory after unilateral temporal lobectomy in humans. *Neuropsychology*, 18, 462–472.
- Galati, G., Pelle, G., Berthoz, A., & Committeri, G. (2010). Multiple reference frames used by the human brain for spatial perception and memory. *Experimental Brain Research*, 206, 109–120.
- Gogos, A., Gavrilescu, M., Davison, S., Searle, K., Adams, J., Rossell, S. L., et al. (2010). Greater superior than inferior parietal lobule activation with increasing rotation angle during mental rotation: An fMRI study. *Neuropsychologia*, 48, 529–535.
- Gramann, K., Muller, H. J., Eick, E. M., & Schonebeck, B. (2005). Evidence of separable spatial representations in a virtual navigation task. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 1199–1223.

- Gramann, K., Muller, H. J., Schonebeck, B., & Debus, G. (2006). The neural basis of ego- and allocentric reference frames in spatial navigation: Evidence from spatio-temporal coupled current density reconstruction. *Brain Research*, 11181, 116–129.
- Greenauer, N., & Waller, D. (2008). Intrinsic array structure is neither necessary nor sufficient for nonegocentric coding of spatial layouts. *Psychonomic Bulletin and Review*, 15, 1015–1021.
- Janzen, G., & van Turenhout, M. (2004). Selective neural representation of objects relevant for navigation. *Nature Neuroscience*, 7, 673–677.
- Kelly, J. W., & McNamara, T. P. (2008). Spatial memories of virtual environments: How egocentric experience, intrinsic structure, and extrinsic structure interact. *Psychonomic Bulletin and Review*, 15, 322–327.
- Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In: C. Freska, & C. Havbel (Eds.), *Spatial cognition. An interdisciplinary approach to representing and processing spatial knowledge* (pp. 1–17). Heidelberg: Springer.
- Kohler, S., Kapur, S., Moscovitch, M., Winocur, G., & Houle, S. (1995). Dissociation of pathways for object and spatial vision: A PET study in humans. *Neuroreport*, 6, 1865–1868.
- Kukolja, J., Marshall, J. C., & Fink, G. R. (2006). Neural mechanisms underlying spatial judgements on seen and imagined visual stimuli in the left and right hemifields in men. *Neuropsychologia*, 44, 2846–2860.
- Lee, T. M., Liu, H. L., Hung, K. N., Pu, J., Ng, Y. B., Mak, A. K., et al. (2005). The cerebellum's involvement in the judgment of spatial orientation: A functional magnetic resonance imaging study. *Neuropsychologia*, 43, 1870–1877.
- Manenti, R., Cotelli, M., Calabria, M., Maioli, C., & Miniussi, C. (2010). The role of the dorsolateral prefrontal cortex in retrieval from long-term memory depends on strategies: A repetitive transcranial magnetic stimulation study. *Neuroscience*, 166, 501–507.
- MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, 288, 1835–1838.
- McNamara, T. P., Rump, B., & Werner, S. (2003). Egocentric and geocentric frames of reference in memory of large-scale space. *Psychonomic Bulletin and Review*, 10, 589–595.
- Milivojevic, B., Hamm, J. P., & Corballis, M. C. (2009). Functional neuroanatomy of mental rotation. *Journal of Cognitive Neuroscience*, 21, 945–959.
- Morgan, L. K., MacEvoy, S. P., Aguirre, G. K., & Epstein, R. A. (2011). Distances between real-world locations are represented in the human hippocampus. *Journal of Neuroscience*, 31, 1238–1245.
- Mou, W., & McNamara, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, 28, 162–170.
- Mou, W., McNamara, T. P., Valiquette, C. M., & Rump, B. (2004). Allocentric and egocentric updating of spatial memories. *Journal of Experimental Psychology: Learning Memory and Cognition*, 30, 142–157.
- O'Keefe, J., & Nadel, J. (1978). *The hippocampus as a cognitive map*. Oxford: Clarendon Press.
- Pietrini, P., Furey, M. L., Ricciardi, E., Gobbi, M. I., Wu, W. H., Cohen, L., et al. (2004). Beyond sensory images: Object-based representation in the human ventral pathway. *Proceedings of the National Academy of Science of the United States of America*, 101, 5658–5663.
- Pizzamiglio, L., Guariglia, C., & Cosentino, T. (1998). Evidence for separate allocentric and egocentric space processing in neglect patients. *Cortex*, 34, 719–730.
- Ramnani, N., & Owen, A. M. (2004). Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging. *Nature Reviews Neuroscience*, 5, 184–194.
- Roland, P. E., & Gulyas, B. (1994). Visual imagery and visual representation. *Trends in Neuroscience*, 17, 281–287.
- Sholl, M. J. (2001). The role of a self-reference system in spatial navigation. In: D. R. Montello (Ed.), *COSIT*, 2001. Berlin: Springer.
- Shelton, A. L., & McNamara, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology*, 43, 274–310.
- Slotnick, S. D. (2004). Visual memory and visual perception recruit common neural substrates. *Behavioral and Cognitive Neuroscience Reviews*, 3, 207–221.
- Thompson, W. L., Slotnick, S. D., Burrage, M. S., & Kosslyn, S. M. (2009). Two forms of spatial imagery: Neuroimaging evidence. *Psychological Science*, 20, 1245–1253.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55, 189–208.
- Valiquette, C., & McNamara, T. P. (2007). Different mental representations for place recognition and goal localization. *Psychonomic Bulletin and Review*, 14, 676–680.
- Valiquette, C. M., McNamara, T. P., & Labrecque, J. S. (2007). Biased representations of the spatial structure of navigable environments. *Psychological Research*, 71, 288–297.
- Xiao, C., McNamara, T. P., Qin, S., & Mou, W. (2010). Neural mechanisms of recognizing scene configurations from multiple viewpoints. *Brain Research*, 1363, 107–116.
- Zacks, J. M. (2008). Neuroimaging studies of mental rotation: A meta-analysis and review. *Journal of Cognitive Neuroscience*, 20, 1–19.
- Zaehle, T., Jordan, K., Wustenberg, T., Baudewig, J., Dechent, P., & Mast, F. W. (2007). The neural basis of the egocentric and allocentric spatial frame of reference. *Brain Research*, 1137, 92–103.