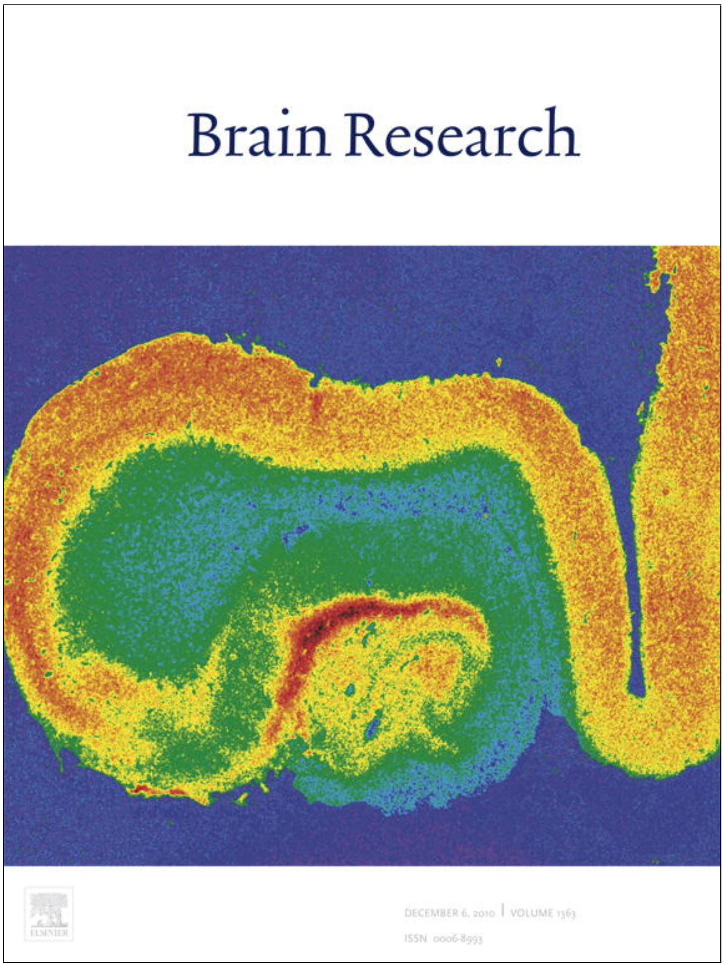
Brain Research
ELSEVIER
DECEMBER 6, 2010 | VOLUME 1363
ISSN 0006-8993

BRAIN RESEARCH 1363 (2010) 107–116

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

Neural mechanisms of recognizing scene configurations from multiple viewpoints

Chengli Xiao[a,b,*], Timothy P. McNamara[c], Shaozheng Qin[d], Weimin Mou[a,e]

[a]Institute of Psychology, Chinese Academy of Sciences, PR China
[b]Department of Psychology, Nanjing University, Nanjing 210093, PR China
[c]Department of Psychology, Vanderbilt University, USA
[d]F.C. Donders Centre for Cognitive Neuroimaging, Radboud University of Nijmegen, The Netherlands
[e]Department of Psychology, University of Alberta, Canada

ARTICLE INFO

Article history:
Accepted 18 September 2010
Available online 25 September 2010

Keywords:
Spatial reference direction
Anterior cingulate cortex
Right intraparietal sulcus
Scene recognition
fMRI

ABSTRACT

Using functional magnetic resonance imaging (fMRI), the present study examined the neural mechanisms involved in recognizing spatial configurations of a scene from multiple viewpoints. Prior to scanning, participants were instructed to learn a desktop array of seven objects relative to an intrinsic direction that was different from the participants' viewpoint. During scanning, participants recognized triplets of objects from the previously memorized scene and from a mirror reflection of the scene at different perspectives. Half of the triplets included two objects located along the instructed intrinsic direction (intrinsic triplets) and the other half did not (non-intrinsic triplets). Consistent with previous mental rotation studies, bilateral intraparietal sulcus and bilateral middle frontal gyrus showed increasing activation with the angular disparity between the test view and the study view. The right intraparietal sulcus was more activated to the non-intrinsic triplets than the intrinsic triplets. The anterior cingulate cortex was more deactivated in recognizing non-intrinsic triplets and novel views. These findings are consistent with the behavioral results that recognition was easier for intrinsic triplets than for non-intrinsic triplets and easier for the familiar view than for novel views (Mou et al., 2008a).

1. Introduction

The ability to recognize previously visited places is critical for people to maintain spatial orientation and to reorient to the environment. Studies have shown that scene recognition is viewpoint dependent (View effect). People recognize a scene faster and with higher accuracy from an experienced view than from a novel view (Diwadkar and McNamara, 1997; Shelton and McNamara, 2004; Simons and Wang, 1998). These findings have been taken as evidence that scene recognition relies on visual memory (e.g., visual "snapshot") and that a mental transformation process is used to match the test scene with the visual snapshot representation in memory.

A recent study (Mou et al., 2008a) found that scene recognition was not only viewpoint dependent, but also dependent upon the spatial reference direction of the scene. In this study, participants were instructed to learn a layout of objects along an intrinsic axis that was different from their

---

[*] *Corresponding author*. Department of Psychology, School of Social and Behavioral Sciences, Nanjing University, 22 Hankou Road, Nanjing 210093, PR China. Fax: +86 25 83594343.

E-mail address: xiaocl@nju.edu.cn (C. Xiao).

0006-8993/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2010.09.068

viewing direction. For example, as shown in Fig. 1, participants viewed the layout from the direction of 315° but were asked to learn the locations of objects according to the columns in the 0°–180° direction (e.g., clip–hat; glue–wood–ball; lock–candle; angular directions were defined arbitrarily). Therefore, 315° corresponded to the viewing direction and 0°–180° corresponded to the assumed intrinsic reference direction. Participants were then given a scene recognition task in which they had to distinguish triplets of objects in the layout at different views from mirror images of those test scenes. Similar to the View effect, participants recognized scenes fastest at the study view (view of 315°) and recognition latency increased linearly with the angular distance between the test view and the study view. Furthermore, a novel finding was that participants were faster to recognize intrinsic triplets of objects, which contained two objects parallel to the instructed intrinsic direction (0°–180° direction, e.g. lock–candle–hat), than non-intrinsic triplets of objects, which did not contain two objects parallel to the instructed intrinsic direction (e.g. hat–ball–lock). This pattern occurred for all test views, and is referred to as reference direction dependent (RDD) performance. This RDD effect was also observed when participants recognized a scene after briefly viewing it for 3 s (Mou et al., 2008b) and when participants recognized a previously viewed shape (Li et al., 2009).

A reference direction model of spatial memory (Mou and McNamara, 2002; Mou et al., 2008a; Mou et al., 2004) was used to explain these findings. According to this model, interobject spatial relations are represented in memory with respect to the instructed spatial reference direction. For example, in Fig. 2A, the angular directions from object 1 to object 3 (a_{13}), from object 1 to object 2 (a_{12}), and from object 1 to object 4 (a_{14}) are specified with respect to the intrinsic reference direction. The angular directions with respect to other directions are not represented and must be inferred if needed (e.g., Klatzky, 1998). These inferential processes introduce observable costs in terms of latency and error. This conceptual model predicts that people will be better assessing the angular direction from object 1 to object 3 with respect to the direction from object 1 to object 2, which is parallel to the reference direction, than with respect to the direction from object 1 to object 4. The model also proposes that viewing direction at study is represented with respect to the spatial reference direction. If the test view is different from the study view, the reference direction of the test view will be oriented differently from the reference direction in the mental representation. Therefore, people needed to align these two reference directions. Referring to Fig. 2, the reference direction defined by objects 1 and 2 at test (Fig. 2B) must be aligned with the reference direction defined by objects 1 and 2 at study (Fig. 2A). Processing costs are introduced during the alignment of the reference directions.

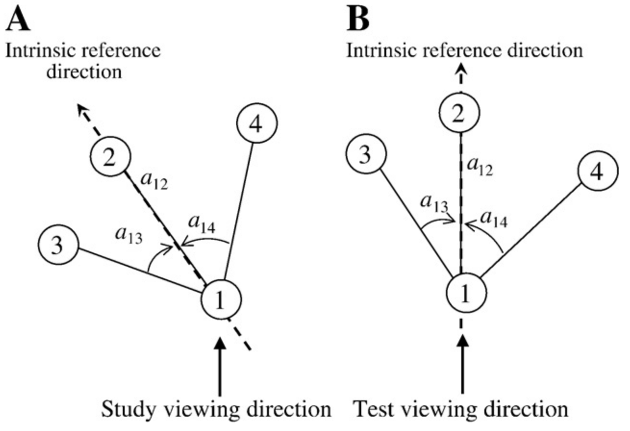

Fig. 2 – The model of spatial memory for scene recognition. Numbered circles represented objects in spatial layout. Participants' study viewing direction was illustrated by the solid arrow and the intrinsic reference direction was illustrated by the dashed arrow.

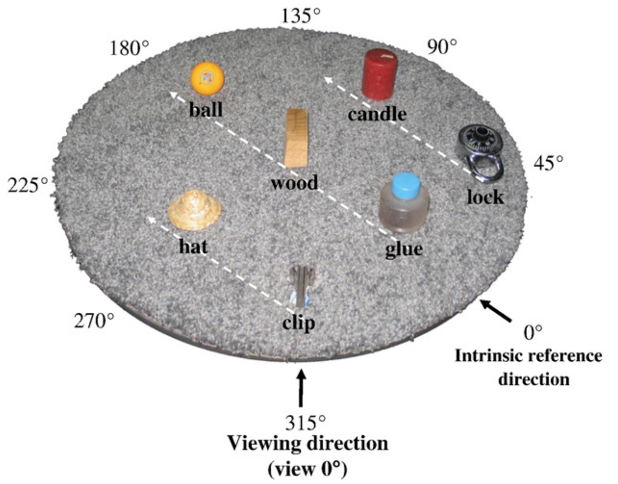

Fig. 1 – The layout of objects used in the experiment. For the intrinsic 0° group, 315° indicated the viewing position; 0° indicated the intrinsic axis. For the intrinsic 315° group, 0° indicated the viewing position; 315° indicated the intrinsic axis.

In particular, scene recognition involves at least two separate processes: (1) identifying the spatial reference directions of the test scene and (2) aligning the identified reference direction with the reference direction represented in spatial memory before confirming the interobject spatial relations in the test scene. Identifying the reference direction of the test scene is easier for the intrinsic triplets of objects than for the non-intrinsic triplets because the former explicitly indicate the spatial reference direction but the latter do not. Furthermore, when the test view is the same as the study view, the spatial reference direction in the test scene is aligned with the represented spatial reference direction of the study scene, but these two spatial reference directions are not aligned when the test view is different from the study view. Hence an extra alignment process is required for novel test views but not for the familiar test view.

The present research aimed to demonstrate a neural dissociation of the effects of RDD and View found in the study of Mou et al. (2008a). Conceptually there might be two possible neural dissociations. First there may be two separate effects of RDD and View in terms of magnitude of neural activation. In particular, magnitude of neural activation may differ in recognizing different triplets (intrinsic vs. non-intrinsic) and in recognizing different views (familiar vs. novel). The second possible dissociation may be that the RDD and View effects occur in two separate brain areas. This study examined both dissociations.

Understanding the neural mechanisms of mental rotation has been an active research area over the past decade (e.g., Carpenter et al., 1999; Cohen et al., 1996; Gauthier et al., 2002; Kosslyn et al., 1998; Milivojevic et al., 2009; Mourão-Miranda et al., 2009; Podzebenko et al., 2002; Richter et al., 1997; Richter et al., 2000; Windischberger et al., 2003). In a meta-analysis, Zacks (2008) selected 32 empirical neuroimaging articles on mental rotation of visual stimuli, such as Shepard and Metzler (1971) objects or alphanumeric characters. Results indicated several cortical areas consistently activated during mental rotation, including intraparietal sulcus and medial superior precentral cortex. These observations were taken to suggest that mental rotation depended on analog representations and motor simulation respectively (Zacks, 2008). We hypothesized that these brain areas might be the neural bases for the mental alignment between the spatial reference directions during scene recognition, i.e. the View effect.

However there were neither theoretical nor empirical reasons to hypothesize which brain areas might be the neural bases for the identification of the spatial reference direction, i.e. the RDD effect. Several brain areas have been found to be activated in spatial tasks compared to non-spatial tasks (e.g., Creem et al., 2001; Zacks et al., 2003). Creem et al. (2001) required participants to imagine themselves rotated in the middle of previously remembered four-object-array. The activation of areas such as bilateral superior parietal lobe (BA 7), middle frontal gyrus (BA 9/11), and precentral gyrus (BA 6) were reported when comparing the rotation task to the control task. Several brain areas have been reported to be relevant to long term spatial memory, such as hippocampus (Biegler et al., 2001; Burgess, 2008; Kumaran and Maguire, 2005; Maguire et al., 2000) and retrosplenial cortex (Epstein, 2008; Maguire, 2001; Sugiura et al., 2005). Hence, the candidate brain areas related to identifying spatial reference directions might be among those areas.

As discussed above, this study also aimed to explore the dissociation in terms of neural activation magnitude in scene recognition for intrinsic triplets versus non-intrinsic triplets and for familiar views versus novel views. Anterior cingulate cortex might be a good candidate area where we could find evidence of greater difficulty in recognizing non-intrinsic triplets and novel views (e.g., Zacks et al., 2003). A "default-model network" (Fox et al., 2005; Fransson, 2005; Raichle and Snyder, 2007) has been proposed to explain the task-induced deactivation (TID) in some brain areas, which exhibit stronger activations in control or rest conditions comparing to goal-directed tasks. The TID of anterior cingulate cortex was consistently found across different studies (Binder et al., 1999; Mazoyer et al., 2001; Shulman et al., 1997), and could be explained as important to accurate performance (Polli et al., 2005), reflecting externally focused attention (Gusnard and Raichle, 2001) or task difficulty (McKiernan et al., 2003). According to the reference direction model (Mou et al., 2008a), participants recognize intrinsic triplets and familiar views with greater ease than non-intrinsic triplets and novel views, respectively. These effects may be reflected in anterior cingulate cortex, with greater deactivation for non-intrinsic than intrinsic triplets and for novel views than familiar views.

In this study participants were instructed to learn a layout of objects with respect to an intrinsic axis that was different from their viewing direction (Fig. 1, 0° and 315°). fMRI data were obtained when participants performed an "old–new" scene recognition task in which "old" stimuli were intrinsic triplets and non-intrinsic triplets from five views (Fig. 3). Areas where BOLD signal had higher activation for non-intrinsic triplets than for intrinsic triplets could reveal the neural substrates for the RDD effect. Areas where BOLD signal increased linearly with the angular distance between the testing view and the studying view could reveal the neural basis of the View effect. As we hypothesized above, intraparietal sulcus and medial superior precentral cortex, which were consistently activated during mental rotation processes, might be relevant to the View effect. The candidate areas for the processes associated with selecting reference directions might be any areas activated in a spatial task, including superior parietal lobe, intraparietal sulcus, medial superior precentral cortex, hippocampus, and retrosplenial cortex. The anterior cingulate cortex, with larger deactivation for a more difficult task, might have higher deactivation for non-intrinsic triplets than for intrinsic triplets and for novel views than for familiar views.

2. Results

2.1. *Behavioral results*

Only responses to target configurations were analyzed. Both percent correct responses and response latency were submitted to separate repeated measures analyses of variance (ANOVAs), with test views (distance 0° to 180° in 45° steps with respect to the learning viewpoint) and type of triplet (intrinsic or non-intrinsic) as within-subject factors. Mean percent correct responses and response latency are plotted in Fig. 4A and B as a function of test view and triplet type.

As shown in the figures, the major findings were as follows. First, participants were more accurate and quicker in recognizing intrinsic triplets than non-intrinsic triplets. Second, percent correct responses decreased and response latency increased as the distance between the test view and the study view increased, regardless of the triplets. For percent correct responses, the effect of triplet type was significant, $F(1, 15) = 14.343$, $p < .005$, MSE = .029, $\eta_p^2 = .489$. The overall effect of the distance to the learning view was significant, $F(4, 60) = 5.372$, $p < .001$, MSE = .014, $\eta_p^2 = .264$. A planned linear contrast (2 1 0 −1 −2) showed that percent correct responses decreased linearly with the distance to the learning view, $t(60) = 3.45$, $p < .005$. The interaction between the triplet type and the distance to the learning view was not significant, $F(4, 60) = 1.322$, $p > .1$, MSE = .015, $\eta_p^2 = .081$. For response latency, the effect of triplet type was significant, $F(1, 15) = 17.127$, $p < .001$, MSE = .711, $\eta_p^2 = .533$. The overall effect of the distance to the learning view was significant, $F(4, 60) = 10.270$, $p < .001$, MSE = .781, $\eta_p^2 = .406$. A planned linear contrast (−2 −1 0 1 2) showed that response latency increased linearly with the

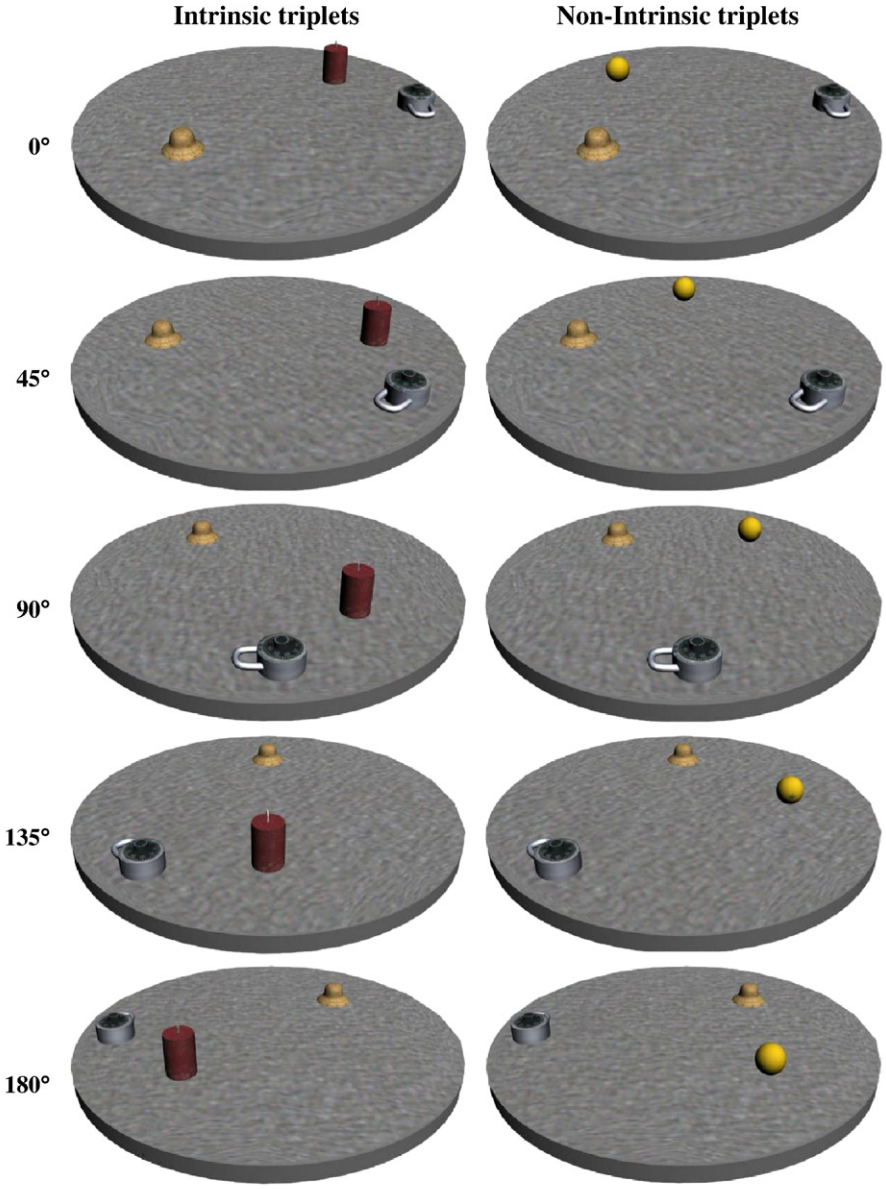

Fig. 3 – Experiment design for the group of intrinsic 0° group. Intrinsic triplets were defined as among which two objects were presented along the intrinsic direction but neither two objects were presented along the learning view. And non-intrinsic triplets were defined as among which two objects were presented along the learning view but neither two objects were presented on intrinsic direction. Both types of triplets were tested from five test views defined with respect to the learning view.

distance to the learning view, $t(60)=6.12$, $p<.001$.The interaction between the triplet type and the distance to the learning view was not significant, $F(4, 60)=1.625$, $p>.1$, MSE=.749, $\eta_p^2=.098$.

2.2. *Functional brain imaging data*

Linear contrasts of event-related fMRI response to different test views were performed to identify the brain areas sensitive to the View effect. Distances 0°, 45°, 90° , 135°, and 180° were weighted as −2 −1 0 1 2 to test the linear increase pattern, and weighted as 2 1 0 −1 −2 to test the linear decrease pattern. There was greater activation as the distance of the test view to the learning view increased in bilateral middle frontal gyrus (left BA 6, right, BA 9/46/13, see Fig. 5A, B and Table 1) and bilateral intraparietal sulcus (left, BA 40, right, BA 40/7/39, see Fig. 5C, D and Table 1). In general, these results were similar to those reported relative to mental rotation (Zacks, 2008). There was no greater activation as the distance of the test view to the learning view decreased.

To detect brain areas related to RDD effect, we compared event-related fMRI signals in responding to different triplets (intrinsic vs. non-intrinsic). The only area that differentiated triplet types was an area that showed a decrease in activity relative to the no-task fixation interval. There was greater deactivation for non-intrinsic triplets in anterior cingulate cortex (BA 32/10; see Fig. 6 and Table 1). The deactivation of

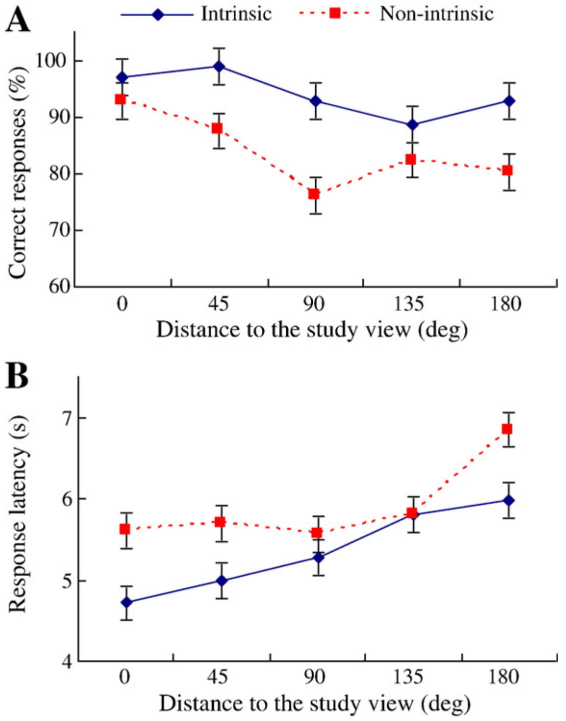

Fig. 4 – Correct responses and response latency in scene recognition as function of distance to the learning view and triplets. (Error bars were confidence intervals corresponding to ±1 standard error of the mean as estimated from the ANOVA.)

this area was also reported in tasks of mental transformations of spatial layouts compared to non-spatial tasks (Creem et al., 2001; Zacks et al., 2003). Furthermore, a whole-brain GLM analysis of the interaction between triplets and viewpoints effects revealed no differential activation in any brain area.

Region of Interest (ROI) analyses were implemented in the above significant areas, i.e. bilateral middle frontal gyrus, bilateral intraparietal sulcus, and anterior cingulate cortex, in the MarsBar Toolbox (Brett et al. 2002). The percentage signal change for the different conditions averaged across all voxels within the activation areas were extracted and used as the dependent measure for second-level ANOVA analyses. In bilateral middle frontal gyrus and bilateral intraparietal sulcus, the overall effects of the distance to the learning view were significant, $Fs(1, 15) > 6.04$, $ps < .001$. The activation increased with the distance between the test view and the learning view. However the significant effect of triplet type (RDD) was found in right intraparietal sulcus, $F(1, 15) = 5.307$, $p < .05$, MSE = .047. The activation was greater for the non-intrinsic triplets than for the intrinsic triplets (see Fig. 5, bottom). No other effect was significant. In the anterior cingulate cortex, the effect of triplet type was significant, $F(1, 15) = 22.675$, $p < .001$, MSE = .017. The deactivation was greater for the non-intrinsic triplets than for the intrinsic triplets. Neither the overall effect of the distance to the learning view nor the interaction between the triplet type and the distance to the learning view was significant. However, the planned linear contrast (2 1 0 −1 −2) indicated that the percentage signal change decreased linearly when the distance to the learning view increased, $t(60) = 2.20$, $p < .05$. In other words, the degree of deactivation increased with the distance to the learning view.

3. Discussion

The behavioral data replicated our previous findings (Mou et al., 2008a). Participants were quicker and more accurate in recognizing intrinsic triplets, regardless of test view, and in recognizing familiar views, regardless of triplet type. The fMRI data revealed that the bilateral middle frontal gyrus and left intraparietal sulcus were related to the superior performance for familiar views regardless of the types of triplets. The activation of right intraparietal sulcus, and the deactivation in anterior cingulate cortex, which was sensitive to task difficulty, were related to the superior performance for intrinsic triplets and for familiar views. There was no evidence that these two effects interacted. These results showed a clear dissociation between the effects of RDD and View in terms of magnitude of neural activation.

This study showed a less clear dissociation between these two effects in terms of the brain locations. Although the bilateral middle frontal gyrus and left intraparietal sulcus were specifically sensitive to the View effect, the right intraparietal sulcus and the anterior cingulate cortex were sensitive to both the View and the RDD effect. The right intraparietal sulcus and the anterior cingulate cortex might be involved in both identifying the spatial reference directions of the test scene (RDD effect) and aligning it with the represented reference direction in spatial memory (View effect).

Consistent with previous studies of mental rotation (Zacks, 2008), activation in bilateral middle frontal gyrus and bilateral intraparietal sulcus increased with the angular difference between the test view and the learning view, which suggested that the alignment of the reference directions was implemented by mental transformations. In the present study, we did not explicitly ask participants to imagine object layout rotation or rotation of themselves about the layout, therefore participants might take either object-based or perspective transformation strategy (Creem et al., 2001; Wraga et al., 2005; Zacks et al., 2003; Zacks and Michelon, 2005).

The parahippocampal place area (PPA) was reported to be involved in perceiving scenes from different viewpoints (Epstein et al., 2003; Epstein et al., 2005). In the present study, there was no evidence that the parahippocampal cortex was activated in the contrast of different test views. There are several possible explanations of this null effect: First, the stimuli in the present study were different from those used in previous studies on the PPA, most of which were photographs of natural environments. Epstein and Kanwisher (1998) suggested that the PPA responded less strongly to arrays of multiple objects than to natural scenes. Second, the test scenes in the current study were only part of the studied scene, and participants had to compare every test scene with a representation of study scene in long-term memory. Because each test scene was different from the remembered scene, the PPA could be equally active in all trials. Third, the design or data collected in this experiment might not have been powerful enough to detect a significant effect in this area.

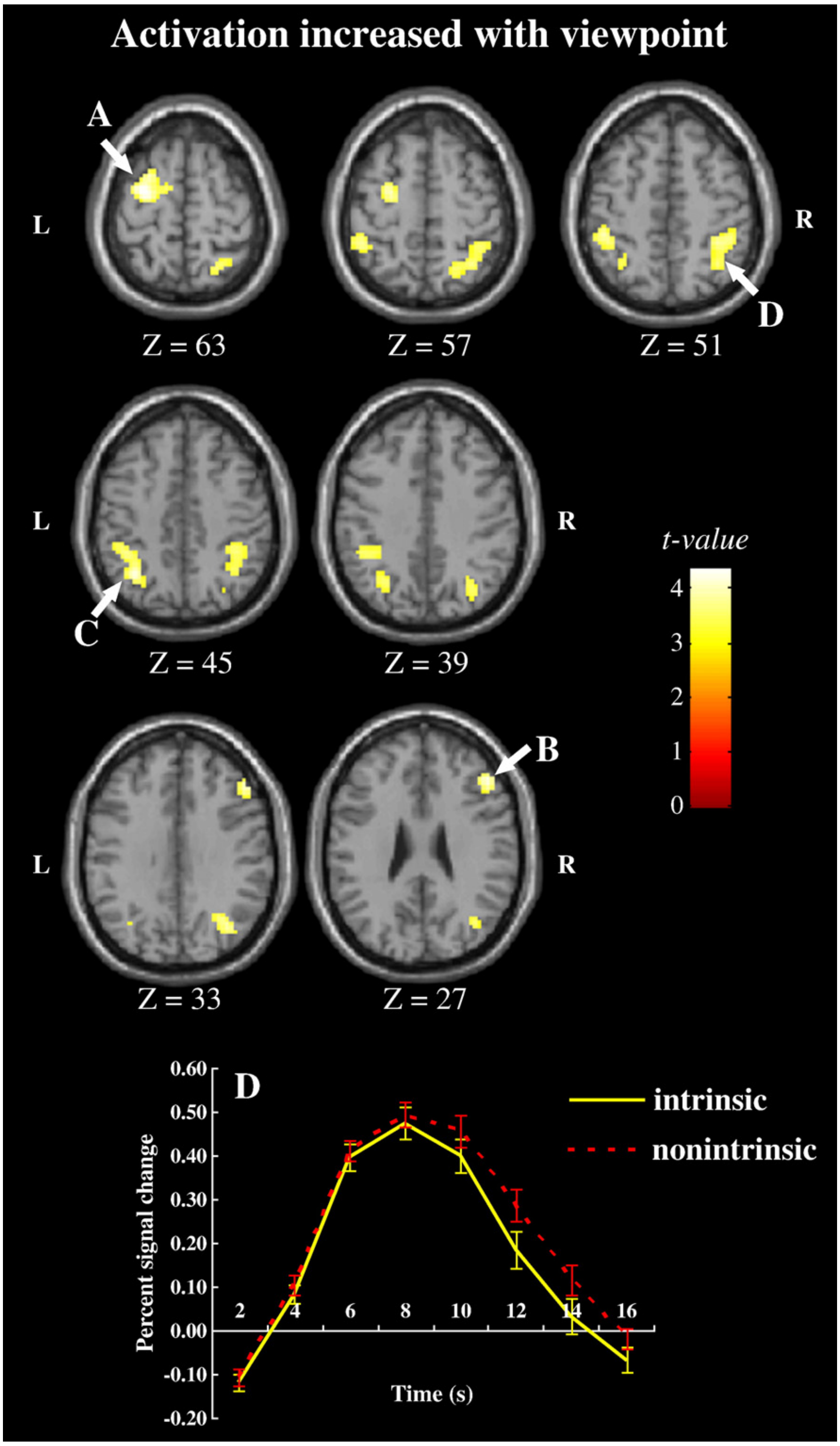

Fig. 5 – Brain areas increasingly activated with distance to the learning view. (distance 0°, 45°, 90° , 135°, 180° were weighted as −2 −1 0 1 2 to test the linear increase pattern). A, left middle frontal gyrus (Brodmann's area 6; MNI peak coordinate = −27, −6, 63; z = 4.23). B, right middle frontal gyrus (Brodmann's area 9/46/13; MNI peak coordinate = 48, 30, 30; z = 4.03). C, left intraparietal sulcus (Brodmann's area 40; MNI peak coordinate = −33, −57, 45; z = 3.99). D, right intraparietal sulcus (Brodmann's area 40/7/39; MNI peak coordinate = 36, −42, 51; z = 3.72). Bottom panel, the time course of the activations in area D in the intrinsic (solid line) and non-intrinsic (dashed line) conditions. Error bars reflect ± 1 SEM.

Studies have shown that the right intraparietal sulcus is involved in a line bisection task (Fink et al., 2000), figure axis orientation task (Faillenot et al., 1999), and object orientation task (Altmann et al., 2005). According to the reference direction model, identifying the spatial reference of a scene is analogous to determining the “top” of a figure (e.g., Rock,

Table 1 – Peak MNI coordinates (mm) showing significant differences ($p < 0.001$ at the voxel level, $p < 0.05$ corrected at the cluster level).

Region	Brodmann's areas	x	y	z	Z score	Cluster extent (mm^3)
View linear increase						
L middle frontal gyrus	6	−27	−6	63	4.23	4887
	6	−24	3	69	3.85	
	6	−12	−3	63	3.45	
R middle frontal gyrus	9	48	30	30	4.03	2997
	46	42	27	18	3.66	
	13	39	18	9	3.14	
L intraparietal sulcus	40	−33	−57	45	3.99	6075
	40	−45	−36	51	3.77	
	40	−39	−42	45	3.61	
R intraparietal sulcus	39	36	−69	33	3.73	2187
R intraparietal sulcus	40	36	−42	51	3.72	6318
	7	24	−60	60	3.65	
	40	45	−39	51	3.59	
Intrinsic > non-intrinsic						
Anterior cingulate cortex	32	9	48	−3	3.84	2511
	32	−6	48	0	3.69	
Medial frontal gyrus	10	−12	51	6	3.68	

1973). And the dissociation of View and RDD effect is also observed when participants recognized a shape (Li et al., 2009). Therefore, we speculated that the right intraparietal sulcus might be related to identifying the spatial reference directions of the test scene.

In the present study, there was no evidence that the brain areas relevant to long-term spatial memory, such as hippocampus (Biegler et al., 2001; Burgess, 2008; Kumaran and Maguire, 2005; Maguire et al., 2000) and retrosplenial cortex (Epstein, 2008; Maguire, 2001; Sugiura et al., 2005), differed in activation for intrinsic and non-intrinsic triplets. The absence of such differences may not be surprising as long-term memory should be used in recognizing both kinds of triplets.

Anterior cingulate cortex was related to the superior performance in recognizing the intrinsic triplets as well as recognizing the familiar view. Deactivation has been reported in both spatial tasks and non-spatial tasks. For instance, deactivation of anterior cingulate cortex has been reported in spatial tasks similar to those used in the present experiment when compared to non-spatial tasks (Creem et al., 2001; Zacks et al., 2003). Moreover, the deactivation of this area was also found when participants were forced to use distal room cues to navigate to a hidden platform in a virtual environment (Shipman and Astur, 2008). Anterior cingulate cortex was also involved in task-induced deactivation in many other non-spatial tasks (e.g., McKiernan et al., 2003). Hence the deactivation of anterior cingulate cortex, which was related to the superior performance in recognizing the intrinsic triplets as well as recognizing the familiar view, might reflect the general difficulty of a task. According to the reference direction model, recognizing intrinsic triplets should be easier than recognizing non-intrinsic triplets, because more effort is required to reconstruct the spatial reference direction in the latter case. Similarly, recognizing familiar views should be easier than recognizing novel views because more effort is needed for novel views to align the identified reference direction with the reference direction represented in spatial memory.

In summary, the present research identified two groups of brain areas specifically related to the viewpoint dependency and the spatial reference direction dependency of scene recognition. The bilateral intraparietal sulcus and bilateral middle frontal gyrus, which had been reported to be involved

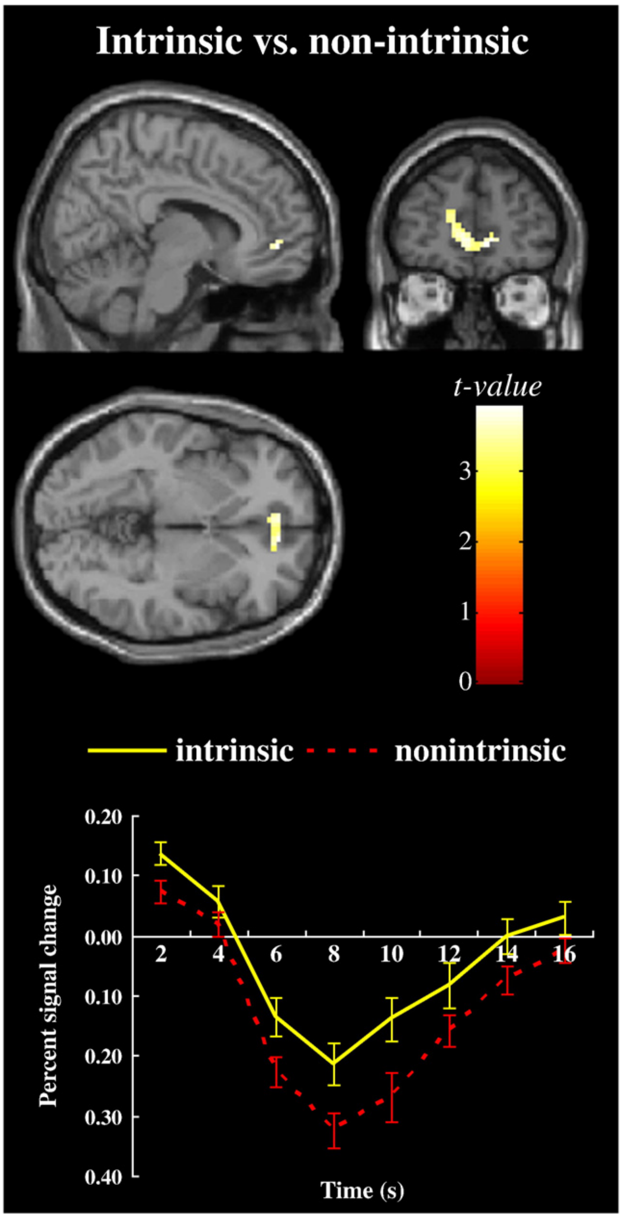

Fig. 6 – Brain areas less deactivated when participants recognizing intrinsic triplets than non-intrinsic triplets. Anterior cingulate cortex (Brodmann's area 32/10; MNI peak coordinate = 9, 48, −3; z = 3.84). Bottom panel, the time course of the deactivations in the intrinsic (solid line) and non-intrinsic (dashed line) conditions. Error bars reflect ± 1 SEM.

in mental rotation processing, were related to the View effect. The right intraparietal sulcus, which was involved in line bisection task and figure/object orientation task, was also related to the RDD effect. The larger deactivation in the anterior cingulate cortex was observed in recognizing non-intrinsic triplets and in recognizing novel views. These findings are in accordance with the claims of the reference direction model of spatial memory that there are at least two different components to recognizing spatial configurations: (a) identifying spatial reference directions in the test scene and (b) aligning the identified reference directions with the represented reference directions of the studied scene in spatial memory. Identifying spatial reference directions is easier for intrinsic triplets than for non-intrinsic triplets and aligning reference directions is easier for familiar views than for novel views.

4. Experimental procedures

4.1. *Participants*

Sixteen volunteers (eight females; mean age 22 years, range 18–27 years) participated for monetary compensation. All participants were right-handed and had normal or corrected to normal vision, and reported no history of psychiatric or neurological disorders, and no current use of any psychoactive medications. Ethical approval for the study was obtained from the Institutional Review Board of Beijing MRI Center for Brain Research. Informed written consent was obtained in accordance with procedures approved by the Institutional Review Board of Beijing MRI Center for Brain Research.

4.2. *Materials and design*

The layout consisted of a configuration of seven common objects sized around 5 cm (see Fig. 1). The objects were placed on a circular table covered by a grey mat (50 cm in diameter, 45 cm above the floor). The distance between the adjacent objects, in the directions aligned with the 0°–180° and 270°–90° axes of the array were 14.14 cm. The table was placed in a corner of the rectangular learning room, with the 315°–135° axis of the array aligned with the longer walls of the room. There was a stool (seated 46 cm high) located at the position of 315° and another identical one located at the position of 0°, with the backside of the stool 100 cm from the middle of the table.

When participants were tested in the scanner, virtual objects instead of real objects were presented. All of the virtual objects and table were created by Autodesk 3ds Max (http://www.autodesk.com/3dsmax). They were virtual analogs of the real objects in the learning room and were presented with the exact scale. The program written with ImageTclAR (Owen et al., 2003) was used to display the stimuli, which were synchronized with MRI data acquisition. The virtual layouts were presented from the perspective as if participants were sitting on the stool in the learning room. Stimuli were displayed on a rear-projection screen mounted in the bore of the scanner, which the participants could view through an angled mirror fixed to the head coil.

Each test trial was constructed from three virtual objects displayed on the virtual table according to the study configuration or according to the mirror reflection of the study configuration with respect to the viewpoint. Thus, half of the trials were target triplets, and the other half of the trials were mirror reflections. Participants were required to respond whether the configuration of the three objects was consistent with the study configuration regardless of the viewpoints. There were 12 triplets used in the study. Six triplets along the axis of 315° were created as following: hat–ball–candle, hat–ball–lock, glue–candle–clip, glue–candle–hat, clip–wood–lock, and clip–wood–candle. Six triplets along the axis of 0° were created as following: wood–ball–candle, wood–ball–lock, lock–candle–ball, lock–candle–hat, clip–hat–candle, and clip–hat–glue. Triplets along the axis of 315° did not contain any two objects lying along the axis of 0°, and triplets along the axis of 0° did not contain any two objects lying along the axis of 315°. Therefore the triplets along the axis of 315° and 0° were exclusively defined.

One of the independent variables was learning direction. Half of the participants viewed the layout from the direction of 315° and were instructed to learn the layout along the direction of 0° (intrinsic 0° group); the other half of the participants viewed the layout from the direction of 0° and were instructed to learn the layout along the direction of 315° (intrinsic 315° group). Participants were randomly assigned to the intrinsic 0° group and the intrinsic 315° group with the restriction that each group contained an equal number of men and women.

Another independent variable was the type of triplet, either intrinsic or non-intrinsic. Intrinsic triplets were defined as those in which two objects were located along the intrinsic direction but no two objects were located along the learning view. Non-intrinsic triplets were defined as those in which two objects were located along the learning view but no two objects were located along the intrinsic direction. For example, the triplets along the axis of 0° (e.g. lock–candle–hat) were intrinsic triplets to the intrinsic 0° group but were non-intrinsic triplets to the intrinsic 315° group.

The last independent variable was the angular distance of the test view to the learning view. For the intrinsic 0° group, the distance of the test view relative to the learning view of 315° were defined clockwise whereas for the intrinsic 315° group, the distance of the test view relative to the learning view of 0° were defined counter-clockwise so that the test view parallel to the intrinsic direction was defined as 45° from the learning view in both groups. Five distances of the test view to the learning view were selected: 0° (learning view), 45°, 90°, 135°, and 180°. Fig. 3 demonstrates intrinsic and non-intrinsic triplets presented from five test views for the intrinsic 0° group.

4.3. *Procedure*

4.3.1. *Learning phase*

Before entering the learning room, each participant was instructed to learn the locations of the objects for a scene recognition test and given one configuration of four objects as a practice so that participants were familiar with the procedure. The participant was then blindfolded and led to be seated in the stool at the viewing position (315° in Fig. 1 for the intrinsic 0°

group and 0° in Fig. 1 for the intrinsic 315° group) in the learning room. The blindfold was removed and participants in the intrinsic 0° group were asked to learn the locations of the objects according to the columns in the 0°–180° direction, as indicated by the experimenter (e.g., "Please try to remember the location of each object as accurately as you can. During memorization, please learn the location of objects in columns, for example, Clip–Hat, Glue–Wood–Ball, Lock–Candle.); participants in the intrinsic 315° group were asked to learn the locations of the objects according to the columns in the 315°–135° direction, as indicated by the experimenter (e.g., Hat–Ball, Clip–Wood, Glue–Candle, Lock, etc.). Participants viewed the display for 30 s before being asked to name and point to the objects with eyes closed. All participants named and pointed to objects in an order consistent with the intrinsic axis. The accuracy of spatial memory was emphasized by giving feedback and corrections if the experimenter observed obvious pointing errors made by participants. Verbal strategies, such as memorized lists of object names, were not encouraged. This learning and pointing procedure was repeated five times for each participant.

4.3.2. *Testing phase*

After learning the spatial layout, participants were led to the MRI scanner. Scanning sessions consisted of one T2-weighted spin-echo anatomical scan followed by four experimental scans, and then followed by a T1-weighted anatomical scan. Each experimental scan was 11 min and 32 s long. Each scan began with a 16-s presentation of a red fixation cross in the middle of black screen. This was followed by 30 16-s-long trials. In each trial a test triplet was presented on the virtual table, and participants were instructed to make a judgment as accurately as they could before the testing triplet disappeared (right button for target configurations and left button for mirror reflection configurations). Each test trial was followed by a fixation cross for 6 s. Each experimental scan ended with a 16-s fixation period. The testing trials were randomized across four experimental scans.

4.4. *fMRI acquisition and analysis*

Scanning was carried out on a 3-T Siemens Trio system using a standard head coil at Beijing MRI Center for Brain Research. T2-weighted spin-echo anatomical images were acquired in 26 contiguous 3.5 mm axial slices with 1.5 mm interslice gaps [TE = 92 ms; TR = 3200 ms]. Functional images were acquired in the same slices using T2*-sensitive gradient echo spiral pulse sequence (TE = 30 ms; TR = 2000 ms; flip angle = 90°; FOV = 220 mm; 64 × 64 acquisition matrix; 26 axial slices, slice thickness = 3.5 mm, interslice gap = 1.5 mm, in-plane resolution = 3.4 × 3.4 mm,). Three-dimensional (3-D) high-resolution T1-weighted FLASH anatomical images were acquired in 144 contiguous 1.3-mm slides with 50% interslice gaps [TE = 3.37 ms; flip angle = 7°; FOV = 256 mm]. Functional images were preprocessed and analyzed using Statistical Parametric Mapping (SPM5) (Welcome Department of Cognitive Neurology, London). Images were corrected for differences in slice timing and motion artifacts and then normalized to the Montreal Neurological Institute template. Images were spatially smoothed with a Gaussian filter of 8 mm full-width half maximum (FWHM).

The canonical hemodynamic response function (HRF) was used to estimate the event-related hemodynamic responses (HDRs) for each participant. Because judgment latencies were relatively long (average RT = 6.61 s), BOLD responses were treated as epoch related. Our interest was in the 16-s period when subjects were viewing the triplets and performing the tasks, therefore responses to each trial were modeled using a 16-s duration boxcar function, aligned with the onset of each trial and convolved with the canonical HRF. Target trials were modeled separately by intrinsic and non-intrinsic triplets at each viewpoint, and mirror distracter trials were collapsed and modeled. The target-trial parameter estimates were included in the second, group level analysis. Statistical analyses were performed in the context of the General Linear Model. Whole-brain statistical maps were thresholded at $p < 0.001$ (uncorrected) and considered significant at $p < 0.05$ corrected for multiple comparisons at the cluster level based on the theory of Gaussian fields (Friston et al., 1996).

Acknowledgments

We wish to thank Adriane Seiffert for her insightful suggestions during the course of this research. We also thank Xiaoou Li for her help with assistance in data collecting. We are grateful to the Beijing MRI Center for Brain Research for using the imaging facility.

REFERENCES

Altmann, C.F., Grodd, W., Kourtzi, Z., Bülthoff, H.H., Karnath, H.-O., 2005. Similar cortical correlates underlie visual object identification and orientation judgment. Neuropsychologia 43, 2101–2108.

Biegler, R., McGregor, A., Krebs, J.R., Healy, S.D., 2001. A larger hippocampus is associated with longer-lasting spatial memory. Proc. Natl. Acad. Sci. U. S. A. 98, 6941–6944.

Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S.F., Rao, S.M., Cox, R.W., 1999. Conceptual processing during the conscious resting state: a functional MRI study. J. Cognit. Neurosci. 11, 80–93.

Brett, M., Anton, J.-L., Valabregue, R., Poline, J.-B., 2002. Region of interest analysis using an SPM toolbox. Neuroimage 16, 497 (abstract).

Burgess, N., 2008. Spatial cognition and the brain. Ann. N. Y. Acad. Sci. 1124, 77–97.

Carpenter, P.A., Just, M.A., Keller, T.A., Eddy, W., Thulborn, K., 1999. Graded functional activation in the visuospatial system with the amount of task demand. J. Cognit. Neurosci. 11, 9–24.

Cohen, M.S., Kosslyn, S.M., Breiter, H.C., DiGirolamo, G.J., Thompson, W.L., Anderson, A.K., Bookheimer, S.Y., Rosen, B.R., Belliveau, J.W., 1996. Changes in cortical activity during mental rotation. A mapping study using functional MRI. Brain 119, 89–100.

Creem, S.H., Downs, T.H., Wraga, M., Harrington, G.S., Proffitt, D.R., Downs III, J.H., 2001. An fMRI study of imagined self-rotation. Cogn. Affect. Behav. Neurosci. 1, 239–249.

Diwadkar, V.A., McNamara, T.P., 1997. Viewpoint dependence in scene recognition. Psychol. Sci. 8, 302–307.

Epstein, R.A., 2008. Parahippocampal and retrosplenial contributions to human spatial navigation. Trends Cogn. Sci. 12, 388–396.

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

Epstein, R., Graham, K.S., Downing, P.E., 2003. Viewpoint-specific scene representations in human parahippocampal cortex. Neuron 37, 865–876.

Epstein, R.A., Higgins, J.S., Thompson-Schill, S.L., 2005. Learning places from views: variation in scene processing as a function of experience and navigational ability. J. Cognit. Neurosci. 17, 73–83.

Faillenot, I., Decety, J., Jeannerod, M., 1999. Human brain activity related to the perception of spatial features of objects. Neuroimage 10, 114–124.

Fink, G.R., Marshall, J.C., Weiss, P.H., Shah, N.J., Toni, I., Halligan, P.W., Zilles, K., 2000. 'Where' depends on 'what': a differential functional anatomy for position discrimination in one- versus two-dimensions. Neuropsychologia 38, 1741–1748.

Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Essen, D.C.V., Raichle, M.E., 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc. Natl. Acad. Sci. U. S. A. 102, 9673–9678.

Fransson, P., 2005. Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. Hum. Brain Mapp. 26, 15–29.

Friston, K.J., Holmes, A., Poline, J.-B., Price, C.J., Frith, C.D., 1996. Detecting activations in PET and fMRI: levels of inference and power. Neuroimage 40, 223–235.

Gauthier, I., Hayward, W.G., Tarr, M.J., Anderson, A.W., Skudlarski, P., Gore, J.C., 2002. BOLD activity during mental rotation and viewpoint-dependent object recognition. Neuron 34, 161–171.

Gusnard, D.A., Raichle, M.E., 2001. Searching for a baseline: functional imaging and the resting human brain. Nat. Rev. Neurosci. 2, 685–694.

Klatzky, R.L., 1998. Allocentric and egocentric spatial representations: definitions, distinctions, and interconnections. In: Freksa, C., Habel, C., Wender, K.F. (Eds.), Spatial Cognition: An Interdisciplinary Approach to Representing and Processing Spatial Knowledge LNAI 1404. Vol. Springer-Verlag, Berlin, pp. 1–17.

Kosslyn, S.M., Digirolamo, G.J., Thompson, W.L., Alpert, N.M., 1998. Mental rotation of objects versus hands: neural mechanisms revealed by positron emission tomography. Psychophysiology 35, 151–161.

Kumaran, D., Maguire, E.A., 2005. The human hippocampus: cognitive maps or relational memory? J. Neurosci. 25, 7254–7259.

Li, X., Mou, W., McNamara, T.P., 2009. Intrinsic orientation and study viewpoint in recognizing spatial structure of a shape. Psychon. Bull. Rev. 16, 518–523.

Maguire, E.A., 2001. The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings. Scand. J. Psychol. 42, 225–238.

Maguire, E.A., Gadian, D.G., Johnsrude, I.S., Good, C.D., Ashburner, J., Frackowiak, R.S.J., Frith, C.D., 2000. Navigation-related structural change in the hippocampi of taxi drivers. Proc. Natl. Acad. Sci. U. S. A. 97, 4398–4403.

Mazoyer, B., Zago, L., Mellet, E., Bricogne, S., Etard, O., Houdé, O., Crivello, F., Joliot, M., Petit, L., Tzourio-Mazoyer, N., 2001. Cortical networks for working memory and executive functions sustain the conscious resting state in man. Brain Res. Bull. 54, 287–298.

McKiernan, K.A., Kaufman, J.N., Kucera-Thompson, J., Binder, J.R., 2003. A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. J. Cognit. Neurosci. 15, 394–408.

Milivojevic, B., Hamm, J.P., Corballis, M.C., 2009. Functional neuroanatomy of mental rotation. J. Cognit. Neurosci. 21, 945–959.

Mou, W., McNamara, T.P., 2002. Intrinsic frames of reference in spatial memory. J. Exper. Psychol. Learn. Mem. Cogn. 28, 162–170.

Mou, W., McNamara, T.P., Valiquette, C.M., Rump, B., 2004. Allocentric and egocentric updating of spatial memories. J. Exper. Psychol., Learn., Mem. Cogn. 30, 142–157.

Mou, W., Fan, Y., McNamara, T.P., Owen, C.B., 2008a. Intrinsic frames of reference and egocentric viewpoints in scene recognition. Cognition 106, 750–769.

Mou, W., Xiao, C., McNamara, T.P., 2008b. Reference directions and reference objects in spatial memory of a briefly viewed layout. Cognition 108, 136–154.

Mourão-Miranda, J., Ecker, C., Sato, J.R., Brammer, M., 2009. Dynamic changes in the mental rotation network revealed by pattern recognition analysis of fMRI data. J. Cognit. Neurosci. 21, 890–904.

Owen, C., Tang, A., Xiao, F., 2003, Octoberr. ImageTclAR: a blended script and compiled code development system for Augmented Reality. STARS2003, The International Workshop on Software Technology for Augmented Reality Systems. Tokyo, Japan.

Podzebenko, K., Egan, G.F., Watson, J.D.G., 2002. Widespread dorsal stream activation during a parametric mental rotation task, revealed with functional magnetic resonance imaging. Neuroimage 15, 547–558.

Polli, F.E., Barton, J.J.S., Cain, M.S., Thakkar, K.N., Rauch, S.L., Manoach, D.S., 2005. Rostral and dorsal anterior cingulate cortex make dissociable contributions during antisaccade error commission. Proc. Natl. Acad. Sci. U. S. A. 102, 15700–15705.

Raichle, M.E., Snyder, A.Z., 2007. A default mode of brain function: a brief history of an evolving idea. Neuroimage 37, 1083–1090.

Richter, W., Ugurbil, K., Georgopoulos, A., Kim, S.-G., 1997. Time-resolved fMRI of mental rotation. NeuroReport 8, 3697–3702.

Richter, W., Somorjai, R., Summers, R., Jarmasz, M., Menon, R.S., Gati, J.S., Georgopoulos, A.P., Tegeler, C., Ugurbil, K., Kim, S.-G., 2000. Motor area activity during mental rotation studied by time-resolved single-trial fMRI. J. Cognit. Neurosci. 12, 310–320.

Rock, I., 1973. Orientation and Form. Academic Press, New York.

Shelton, A.L., McNamara, T.P., 2004. Spatial memory and perspective taking. Mem. Cogn. 32, 416–426.

Shepard, R.N., Metzler, J., 1971. Mental rotation of three-dimentional objects. Science 171, 701–703.

Shipman, S.L., Astur, R.S., 2008. Factors affecting the hippocampal BOLD response during spatial memory. Behav. Brain Res. 187, 433–441.

Shulman, G.L., Fiez, J.A., Corbetta, M., Buckner, R.L., Miezin, F.M., Raichle, M.E., Petersen, S.E., 1997. Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. J. Cognit. Neurosci. 9, 648–663.

Simons, D.J., Wang, R.F., 1998. Perceiving real-world viewpoint changes. Psychol. Sci. 9, 315–320.

Sugiura, M., Shah, N.J., Zilles, K., Fink, G.R., 2005. Cortical representations of personally familiar objects and places: functional organization of the human posterior cingulate cortex. J. Cognit. Neurosci. 17, 183–198.

Windischberger, C., Lamm, C., Bauer, H., Moser, E., 2003. Human motor cortex activity during mental rotation. Neuroimage 20, 225–232.

Wraga, M., Shephard, J.M., Church, J.A., Inati, S., Kosslyn, S.M., 2005. Imagined rotations of self versus objects: an fMRI study. Neuropsychologia 43, 1351–1361.

Zacks, J.M., 2008. Neuroimaging studies of mental rotation: a meta-analysis and review. J. Cognit. Neurosci. 20, 1–19.

Zacks, J.M., Michelon, P., 2005. Transformations of visuospatial images. Behav. Cogn. Neurosci. Rev. 4, 96–118.

Zacks, J.M., Vettel, J.M., Michelon, P., 2003. Imagined viewer and object rotations dissociated with event-related fMRI. J. Cognit. Neurosci. 15, 1002–1018.