

Running head: Imagined Viewer and Object Rotations

Imagined Viewer and Object Rotations Dissociated with Event-Related fMRI

Jeffrey M. Zacks, Jean M. Vettel, and Pascale Michelon

Washington University in Saint Louis

Please address communications to:

Jeffrey M. Zacks

Washington University

Psychology Department

St. Louis, MO 63130-4899

314-935-8454

jzacks@artsci.wustl.edu

IN PRESS, JOURNAL OF COGNITIVE NEUROSCIENCE

Abstract

Human spatial reasoning may depend in part on two dissociable types of mental image transformations: object-based transformations, in which an object is imagined to move in space relative to the viewer and the environment, and perspective transformations, in which the viewer imagines the scene from a different vantage point. This study measured local brain activity with event related functional MRI while participants were instructed to imagine either an array of objects rotating (an object-based transformation) or to imagine themselves rotating around the array (a perspective transformation). Object-based transformations led to selective increases in right parietal cortex and decreases in left parietal cortex, whereas perspective transformations led to selective increases in left temporal cortex. These results argue against the view that mental image transformations are performed by a unitary neural processing system, and they suggest that different overlapping systems are engaged for different image transformations.

Introduction

Components of Spatial Reasoning

Mental spatial transformations are a ubiquitous and important component of everyday reasoning. People perform them in order to read maps, use tools, arrange furniture, play chess, and drive in traffic. Consider one example in detail: Imagine you are sitting on an airport shuttle bus across from someone holding a map. You wonder whether a building marked in red on the map corresponds to the location of a cathedral you'd like to visit. To satisfy your curiosity, you could imagine the map moving until it was facing you. We refer to transformations like this as object-based transformations, because they involve the movement of an object relative to the viewer and the environment. Alternatively, you could imagine yourself sitting in the position of the person holding the map. We refer to this sort of operation as an egocentric perspective transformation (or simply perspective transformation) because one imagines one's personal point of view moving relative to the environment.

The two sorts of transformations share a number of features, and solving them likely requires the operation of a number of neural processing resources. If a person does indeed solve these problems by forming a mental image and imagining an image transformation, it is possible to describe in some detail the processes involved. First, one needs to encode the spatial situation specified by the problem (in this case, conveyed by language). Second, the locations and orientations of the objects need to be located with respect to spatial reference frames (Bryant, Tversky, & Franklin, 1992; McCloskey, 2001). Three reference frames are relevant to locating any object: an egocentric reference frame that specifies location and direction relative to the self, an intrinsic reference frame that is relative to the located object, and an environmental reference frame that is relative to the immediate surroundings¹. Third, one must form a mental image based on this computed spatial information, a process that has been described as forming a representation in a perception-like storage buffer (Farah, 1989; Kosslyn, 1994). Fourth, one must calculate what transformation of the image and reference frames should be performed to create a representation that will answer the question at hand; this process may be influenced by habit, stimulus features, and practice. Fifth, one must implement this transformation and update the mental image. Finally, one must read the answer to the problem out from the transformed image. Most of these

operations will be required for both the object-based transformation problem and the perspective transformation problem. Understanding the problem, specifying the spatial locations of the objects, and forming a mental image are likely to be similar operations in the two cases. Similarly, reading a spatial relationship out from a transformed mental image should be similar for both transformations. However, the process of updating the spatial relations and generating a new image may differ between the two transformations.

These two transformations could be implemented by one of three neurophysiological architectures. The simplest model would have object-based image updating and egocentric perspective image updating performed by one unitary processing component. On this view, all stages of processing would be performed by the same processing subsystems in both transformations. We will refer to this as the unitary model. A second possibility is that one of the two transformations requires all the processing resources of the other, plus some additional subsystems. For reasons we will detail in the following section, the most plausible version of this model has the processing resources involved in egocentric perspective transformations being a subset of those involved in object-based transformations. We will call this the hierarchical model because one processing network is a strict subset of the other. Finally, object-based and perspective transformations may each depend on unique processing resources that are not shared by the other process. We will refer to this as the double dissociation model. The hierarchical and double dissociation models are both instances of multiple systems models (Zacks, Mires, Tversky, & Hazeltine, 2002), but they differ in the configuration of the postulated components. According to all three models, solving a spatial reasoning problem by performing any mental image transformation involves a large number of shared processing resources. Thus, much of the neural network involved in solving such problems will be common across transformations. As will be seen in the following section, the extant neuropsychological and neurophysiological data provide ample support for this claim. However, the evidence to distinguish between the three models is much less clear.

It is important to note that these two classes of spatial transformation are not exclusive. In particular, imagined movements of body parts may play an important role in many spatial transformation tasks. There is evidence to support this view from behavioral experiments (Sirigu & Duhamel, 2001; Wexler, Kosslyn, & Berthoz, 1998; Wohlschläger, 2001; Wohlschläger &

Wohlschläger, 1998) and neuroimaging studies (Kosslyn, DiGirolamo, Thompson, & Alpert, 1998; Kosslyn, Thompson, Wraga, & Alpert, 2001; Parsons et al., 1995; Richter et al., 2000; Sirigu & Duhamel, 2001). A full account of the neurophysiology of mental imagery will need to integrate imagined motor movements with object-based and perspective transformations. However, here we focus on object-based and perspective transformations.

Neural Systems for Mental Spatial Transformations

Data from neuropsychology and neuroimaging consistently indicate the involvement of a large number of brain regions in solving spatial reasoning problems. Patients with focal lesions to any quadrant of the brain (anterior/posterior, right/left) are impaired on neuropsychological tests of spatial reasoning (Semmes, Weinstein, Ghent, & Teuber, 1963). Consistent with this, electrophysiological and neuroimaging data find wide and varying patterns of activity associated with performing spatial reasoning tasks (e.g., Vingerhoets et al., 2001; Wijers, Otten, Feenstra, & Mulder, 1989). A smaller number of studies have tightly controlled encoding and response requirements by comparing large-magnitude rotations to small-magnitude rotations within a single task (Carpenter, Just, Keller, Eddy, & Thulborn, 1999; Gauthier et al., 2002; Richter et al., 2000; Tagaris, Kim, Strupp, & Andersen, 1996; Zacks, Ollinger, Sheridan, & Tversky, 2002). These converge in pointing to posterior parietal, posterior temporal, lateral occipital cortex, the supplementary motor area (SMA), and sometimes the cerebellum as being particularly important for carrying out mental image transformations.

Much of the research on human spatial reasoning has focused on object-based transformations, particularly mental rotation: imagining the rotation of an external object, while retaining one's egocentric perspective. Reviews have concurred in arguing that the image updating processes necessary for performing mental rotation are localized to posterior cortex (Farah, 1989; Kosslyn, 1994; Newcombe & Ratcliff, 1989). Deficits in other areas, particularly left frontal cortex, can lead to poor performance on mental rotation tasks, but this may be due to impairments in image generation (Farah, 1989; Newcombe & Ratcliff, 1989). There is some controversy regarding the lateralization of the posterior components, but on balance it appears that the right hemisphere plays the stronger role (Corballis, 1997). Data from split-brain patients indicates a left visual field advantage for mental rotation, and data from neurologically normal participants responding to rapidly presented lateralized

stimuli shows the same pattern (Corballis & Sargent, 1989a, 1989b; Ditunno & Mann, 1990). However, this overall difference appears to be mediated by the materials used and the spatial ability of the participants (Fischer & Pellegrino, 1988; Voyer, 1995). Electrophysiological data have tended to show right-lateralized posterior responses during mental rotation tasks (Pegna et al., 1997; Yoshino, Inoue, & Suzuki, 2000), though lateralization is not always reported. The neuroimaging literature is more ambiguous: Several studies have reported right-hemisphere localized responses for mental rotation (Carpenter et al., 1999; Harris et al., 2000; Zacks, Ollinger et al., 2002; Zacks, Rypma, Gabrieli, Tversky, & Glover, 1999), but others have reported bilaterally balanced involvement (Alivisatos & Petrides, 1997; Cohen et al., 1996; Jordan, Heinze, Lutz, Kanowski, & Jancke, 2001; Kosslyn et al., 1998), and two have reported left hemisphere dominance (Just, Carpenter, Maguire, Diwadkar, & McMains, 2001; Vingerhoets et al., 2001).

A smaller body of research has examined the neurophysiology of egocentric perspective transformations. Studies of patients with focal brain lesions have used map-reading tasks and “body schema” tasks, such as pointing to body parts indicated on a diagram. Both tasks depend on mapping one’s egocentric perspective onto an external representation. In a study of focal missile wound patients, deficits on these sorts of tasks were associated with left posterior lesions (Semmes et al., 1963). A PET study by Bonda and colleagues (1996) found that a task in which participants imagined themselves in the position of a nearby experimenter led to left-lateralized posterior activity. Another recent PET study directly contrasted a condition involving a perspective transformation with one that required no such transformation (Ruby & Decety, 2001). Participants either imagined themselves performing an action on an object or imagined the experimenter performing the action. Imagining one’s self performing the action would seem to invite an egocentric perspective transformation, because one would likely imagine one’s self standing in a position appropriate to performing an action; imagining the experimenter performing the action requires no such transformation. Relative to imagining the experimenter performing the action, imagining one’s self performing the action led to increased activity in the inferior parietal lobule, posterior insula, and post-central gyrus, all in the left hemisphere. Finally, a recent fMRI study found that asking participants to imagine themselves moving led to activity in left posterior

parietal cortex (Creem, Downs, Wraga, Proffitt, & Downs, 2001). The paradigm used in this study played a particular role in motivating the present work, so we will describe it in more detail shortly.

The research reviewed to this point allows for only indirect comparison between object-based and perspective transformations because the two have rarely been considered together. Two recent neuroimaging experiments directly compared object-based and perspective transformations, using designs intended to hold other aspects of task performance constant. The first (Zacks et al., 1999) was based on the Ratcliff (1979) manikin test. Participants made judgments about the handedness of upright and inverted pictures of a human body. Responding to inverted figures led to right-dominant increases in activity in posterior parietal, temporal and occipital cortex when making judgments about inverted figures, compared to upright figures. When activity during judgments about upright figures was compared to a resting fixation control increases were observed in posterior cortex that were primarily in the left hemisphere. Both results are consistent with the pattern of deficits in Ratcliff's (1979) patients. The authors argued that the right-dominant posterior differences between inverted and upright figures reflected object-based transformations, whereas the left-dominant posterior differences between upright figures and rest reflected egocentric perspective transformations. The second study also used pictures of human bodies with an outstretched arm (Zacks, Ollinger et al., 2002). Here, participants made two different judgments about the bodies. In some blocks they indicated which arm was outstretched. In others they judged whether two bodies appearing at different orientations were identical or mirror images. Based on behavioral evidence (Zacks, Mires et al., 2002), it was hypothesized that these two tasks would selectively elicit egocentric perspective transformations and object-based transformations, respectively. Regions in right parietal, temporal, and occipital cortex and the medial cerebellum were more active for the task thought to depend on object-based transformations. However, no regions were observed showing the opposite pattern.

In sum, the existing neuropsychological and neurophysiological data argue against the unitary system model. They appear on balance to support the double dissociation model over the hierarchical model, but are not conclusive. A number of studies have suggested specialization of some cortical regions, particular in right posterior regions, for object-based transformations. The evidence for brain areas specialized for perspective transformations is weaker. In part this reflects a lack of attention to

perspective transformations, particularly in neuroimaging studies. However, the small number of studies testing for brain regions selectively active for perspective transformations have produced mixed results.

Object and Viewer Rotations

One means to attempt to decide between the hierarchical and double dissociation models comes from an experimental task originally designed to study cognitive development. In this paradigm, participants view an array of objects and are asked to make a spatial judgment about the array, such as reporting the location of a particular object or the color of an object at a particular location. Critically, participants are asked to make this judgment based not on how the array currently appears, but on how it would appear if some spatial transformation were to take place. In the original version, developed by Piaget and Inhelder (1956), children were shown a model consisting of three mountains with distinctive features (e.g., snow, a house, a cross) and asked to report how the scene would look if viewed from a different vantage point.

Huttenlocher and Presson (1973) adapted this task to study both perspective transformations and object-based transformations. They showed children an array of objects and asked them to imagine what the array would look like if either (a) they were to move around the array (viewer rotations), or (b) the array were to rotate (object rotations) and match this to one of several pictures. For both types of transformation, response time and error rate were greater for large rotations than for small rotations. For the picture-matching task, array rotations were easier than viewer rotations (Huttenlocher & Presson, 1973). However, when children were asked to report which item was at a particular location after the transformation, rather than match the transformed array to a picture, this pattern reversed (Huttenlocher & Presson, 1979). They labeled these appearance and item questions. In a later study with college-aged participants, Presson (1982) extended this paradigm, also including position questions, in which a particular object was identified and the participant was asked where it was located. As before, Presson found that the relative difficulty of object and viewer transformations varied depending on the question asked. However, he argued that in the situations that appeared to favor object rotations, participants might be translating individual elements of the array rather than rotating the array as a whole. Wraga and colleagues (2000), in a series of tightly controlled experiments, found a consistent advantage for

viewer rotations, though the magnitude of this advantage varied in a pattern consistent with the original Huttenlocher and Presson (1979) findings. They also used a catch trial design, in which most trials required only a position judgment, but occasionally an item question was asked immediately afterward. Responses for these catch trials were fast and accurate after viewer transformations, but slower and less accurate after object transformations—further evidence that people may sometimes solve object rotation problems by imagining one of the objects in the array moving, rather than rotating the array as a whole. Together, these studies indicate that participants can selectively perform viewer and object rotations as requested, but sometimes may imagine parts of an array of objects moving rather than the whole array, when the target question allows. (Piecemeal transformation of the array is of course a perfectly good object-based spatial transformation, but it is less directly comparable to a holistic viewer rotation.)

Viewer and object rotations can be performed based on depicted object arrays, as in the previously described experiments, or based on arrays learned through description. For example, Tversky and colleagues (1999) presented participants with a story that described an array of objects, and asked them to imagine object and viewer rotations in that imagined environment. They observed reliably different patterns of response time for the two transformations.

Extending the results on the neurophysiology of egocentric perspective transformations described above, a recent fMRI study used an imagined array to study viewer rotations (Creem et al., 2001). In this experiment participants learned the imagined environment by studying a picture depicting an array of four objects. During scanning, participants were asked to imagine themselves in the middle of the array, and then imagine rotating around their principal axis (a “log roll”). After each imagined transformation, they were asked to indicate where a named object would be located. Imagined viewer rotations gave rise to activity in posterior parietal cortex, lateralized to the left hemisphere.

This approach, in which participants are directly instructed to perform one of two imagined transformations, nicely complements approaches in which the transformation required is inferred from the task demands and behavioral profile (such as typical mental rotation tasks). By directly instructing participants to perform either an object or viewer rotation, the experimenter achieves a direct manipulation of the image transformation process without having to make inferences about the transformation performed based on the pattern of behavioral data. Comparing object and viewer

rotations in this paradigm has the additional advantage that identical object arrays are presented for both types of transformation. This controls the complexity of the spatial stimulus being operated on. However, it is important to keep in mind that this procedure depends on participant compliance, and it is subject to reactivity: The cognitive operation actually performed may be affected by participants' theories about how long or how difficult different sorts of mental transformations should be (Pylyshyn, 1981).

The results from behavioral studies of object and viewer rotations permit two conclusions. First, it is possible to present participants with an array of objects and ask them to imagine either an object-based or perspective transformation and then perform a spatial judgment. The resulting data indicate that people can understand and comply with these instructions, producing orderly patterns of response time and error rate. Second, the fact that different spatial judgments are easier with each of the two types of transformation indicates that they do indeed differ computationally. The single published neuroimaging study of viewer rotations (Creem et al., 2001) supports the view that there are areas specialized for viewer rotations, but does not speak to object rotations.

Goals of the Current Study

In the current study, we sought to use the object and viewer rotation tasks to test the three models of spatial transformations. The unitary model asserts that one common set of regions implements the two apparently different sorts of transformation. Based on the data reviewed previously, the junction of the parietal, temporal and occipital (PTO) cortex would be a likely principal locus of image updating under this model. The hierarchical model asserts that one set brain regions is required for both perspective and object transformations, whereas a second set of regions is selectively required for object transformations. Previous data would suggest right PTO cortex might play this role. Finally, the double dissociation model asserts that unique brain regions are involved in each of the two types of transformation. The data reviewed above suggest a predominantly left-hemisphere locus for perspective transformations and a predominantly right-hemisphere locus for object transformation².

We asked participants to perform object and viewer rotations while measuring local brain activity with functional magnetic resonance imaging (fMRI). The double dissociation model predicted that we would observe areas uniquely activated by each of the two transformations. In particular, we

hypothesized that such areas would tend to be in posterior PTO cortex, and would tend to be left lateralized for viewer rotations and right lateralized for object rotations. The hierarchical model predicted we would observe only half of this double dissociation pattern. Based on previous results (Zacks, Ollinger et al., 2002), we expected that if only one of the two transformations would give rise to unique activity, it would be the object rotations. The unitary system model, of course, predicted an absence of activity unique to either transformation.

A second goal was to replicate the comparison of object and viewer rotations across multiple spatial judgments. Processes that are specific to updating a spatial image should tend to be consistent across different spatial judgments. As noted previously, there is evidence that the relative difficulty of making a spatial judgment interacts with the transformation performed (Presson, 1982; Wraga, Creem, & Proffitt, 1999; Wraga et al., 2000). This means that brain activity in a single judgment that differs across transformations could reflect the transformation *per se*, or differences in how the spatial judgment is made. Moreover, different spatial judgments offer different methodological advantages and disadvantages. We first studied judgments of spatial location (Experiment 1). Participants were asked to report the location of a particular object after the transformation. This task is relatively easy to explain and for participants to perform. However, as noted previously, there is reason to believe that for object rotations participants sometimes solve this problem by imagining the individual objects in the array translating, rather than rotating the array as a whole. Although translation of individual objects is a perfectly good object-based spatial transformation, it is more difficult to compare to the imagined viewer rotations because it differs in the geometry of the transformation. For a converging spatial judgment, we chose a color judgment task (Experiment 2). In this task a particular location was cued and the participant was asked to report the color of the object that would be in that location after the transformation. This task is not prone to the individual object translation strategy. By obtaining converging evidence from two different judgments, we aimed to isolate activity due to transformation processes from those related to the judgment, and to capitalize on the advantages of each judgment task.

Overview of Experiments

In this study participants were asked to perform two sorts of transformations of an array of objects. Imagined object rotations required participants to imagine the array of objects rotating in space.

Imagined viewer rotations required participants to imagine themselves moving in a rotational path around the array. After imagining either the array or themselves moving, participants were asked to report some aspect of how the scene would appear after the imagined transformation. In both experiments, trials of both types of transformation were randomly intermingled.

The two experiments differed primarily in the type of judgment that was required after the imagined transformation. In the first experiment, participants reported the location of one of the objects by indicating whether it would be on their left or right after the transformation. In the second experiment, a particular location was specified, and the participants indicated the color of the object in that location after the transformation. (The two experiments also differed in a few procedural details, particularly in the amount of practice given, as described in Methods below.)

Both experiments employed rendered drawings of arrays of four cubes on poles set at the corners of a wooden base (see Figure 1 and Figure 2). Pacing of the trials was the same for the two experiments: Participants were required to respond within 7.32 s (three MRI acquisition frames) of presentation of the stimulus, and trials were separated by 0, 2.44, or 4.88 s of fixation (in order to provide variability in the intertrial interval, which was necessary for statistical analysis of the MRI data). Because the two experiments were so similar we analyzed the data from both experiments together, for statistical power and efficiency of presentation. However, it is important to appreciate that the two studies were conducted sequentially, and that incidental features of the stimulus design varied between the two studies (as detailed below).

Results

Task Performance

Across both object and viewer rotations in both experiments, performance showed a consistent pattern. Response time and error rate increased with degree of rotation, as shown in Figure 3. This increase was greater for object rotations than viewer rotations. Overall response time and error rate were higher for object rotations; response time and error rate also increased more steeply with orientation for object rotations. This pattern is consistent with previous behavioral results from similar paradigms using arrays of real objects (Carpenter & Proffitt, 2001; Huttenlocher & Presson, 1973, 1979; Presson, 1982; Wraga et al., 2000).

Response times were analyzed by computing each participant's mean response time for each condition and submitting these to a mixed ANOVA. There were two repeated measures: transformation (viewer or object rotation) and rotation magnitude (0, 90, 180 or 270 degrees). Judgment was a between-participants factor (location or color, corresponding to Experiments 1 and 2, respectively). Trials on which an error occurred were excluded, as were outliers (defined as any response faster than 300 ms or slower than 3 standard deviations from that participant's mean).

The response time analysis confirmed the reliable effect of rotation magnitude on response time, $F(3, 90) = 211.2, p < .001$. To test specifically that response time increased with increasing orientation, we conducted pairwise comparisons based on Tukey's W for each increase (0 to 90, 90 to 180, and 180 to 270 degrees) for all combinations of transformation and judgment. These all indicated statistically significant effects at the .05 level, except for two cases: the 180-to-270 degree increases for self-transformation conditions in both experiments. Responses were also reliably faster for viewer rotations than for object rotations, $F(1, 30) = 90.0, p < .001$, and increased more with increasing rotation magnitude for object rotations, leading to a reliable interaction between transformation and rotation magnitude, $F(3, 90) = 16.0, p < .001$.

There were also several effects on response time involving the judgment performed. We note that the experiments were not designed to examine such effects, and therefore they could reflect intrinsic differences between the two judgments, or could be due to the greater amount of practice given to the participants in Experiment 2 (color judgments). Responses were faster for color judgments than for location judgments, $F(1, 30) = 12.2, p = .001$. There was a trend that approached statistical reliability such that the effect of rotation magnitude on response time was marginally greater for location judgments, $F(3, 90) = 2.61, p = .06$. Finally, there was a three-way interaction between transformation, rotation magnitude, and judgment, $F(3, 90) = 6.29, p < .001$. There was no interaction of transformation and judgment, $F(1, 30) = 0.00$.

Given the presence of effects involving judgment, we conducted follow-up ANOVAs for each of the two experiments separately. These analyses confirmed that the response time increased with stimulus orientation for both location judgments, $F(3,45) = 131.0, p < .001$, and color judgments, $F(3,45) = 85.0, p < .001$. They also showed that responses were faster for viewer rotations than object

rotations for both location judgments, $F(1,15) = 61.4$, $p < .001$, and color judgments, $F(1,15) = 35.6$, $p < .001$. Finally, in both experiments response time increased more with orientation for object than viewer rotations, leading to a reliable interaction between transformation and rotation magnitude, $F(3,45) = 6.84$, $p < .001$ for location judgments, $F(3,45) = 16.6$, $p < .001$ for color judgments.

Error rates were analyzed with ANOVAs of the same design used for response times, with each participant's error rate in each condition as the dependent variable. The effects of transformation and rotation magnitude paralleled those for response times. Errors increased with increasing rotation magnitude, $F(3, 90) = 14.4$, $p < .001$. Error rates were higher for object rotations, $F(1, 30) = 33.4$, $p < .001$. Finally, the effect of rotation magnitude on error rate was greater for object rotations, $F(3, 90) = 12.5$, $p < .001$.

Unlike response times, error rates revealed no reliable effects involving judgment. This indicates that increased training and screening in Experiment 2 was successful in reducing errors to levels comparable to those of Experiment 1.

In short, (a) larger rotations were more difficult (slower and less accurate) than small rotations, (b) object rotations were modestly more difficult than viewer rotations, and (c) the effect of rotation on difficulty was greater for object rotations than viewer rotations. This is consistent with previous studies of imagined self and object rotations (Huttenlocher & Presson, 1973, 1979; Wraga et al., 2000).

fMRI Analyses

The local blood oxygen level dependent (BOLD) response was calculated for each combination of transformation, response time, and judgment. For each participant response time was binned across the two transformations into four categories: fast, medium-fast, medium-slow, and slow trials. Basing the analysis on response time rather than rotation magnitude allowed us to control differences in the relationship between rotation magnitude and response time across tasks (see Methods). The analysis was based on the general linear model with impulse response basis functions, adding timepoint (number of frames after trial onset) as an independent variable. Thus, the fMRI analysis had 4 independent variables and 11 interactions, only a subset of which are of interest here. We therefore begin with the effects of primary interest, those involving differences between object and viewer rotations (effects of transformation). All other reliable effects will be discussed in the following sections.

Differential Effects of Object and Viewer Rotations. In this analysis, differences in the brain response to the two transformations can appear as a main effect of transformation, indicating a reliably greater overall response in one of the two tasks. They can also appear as an interaction between transformation and timepoint, indicating a difference in the shape of the response to the two transformations, including a larger magnitude of change. One region, in the right intraparietal sulcus, had a reliably larger mean evoked BOLD signal during object rotations than viewer rotations, leading to a reliable main effect of transformation (Talairach coordinate at peak: 31, -51, 45; Brodmann's area [BA] 7/40). One region in the lateral posterior part of the left superior temporal sulcus, at the PTO junction, showed greater modulation of BOLD response over the course of a trial during viewer rotations, leading to a reliable transformation by time interaction (Talairach coordinate of peak: -59, -50, -02; BA 21/37). Finally, there was one lateral parietal region in the left hemisphere whose activity decreased following trial onset, reliably more for object rotations than for viewer rotations, resulting in a reliable transformation by time interaction (Talairach coordinate of peak: -45, -67, 24; BA 39). We also observed similar activity just below the statistical threshold in the homologous right hemisphere region. (For both the right intraparietal region and the left PTO region, thresholds had to be dropped more dramatically, into the range of the noise, before contralateral regions were present.) The locations and activity profiles of these regions are shown in Figure 4.

To further characterize this activity, we averaged each participant's estimated BOLD response for each trial type over each of the three regions, and submitted these mean responses to three region-wise ANOVAs. If the effects observed were due to activity during only one of the judgment tasks, this could lead to an interaction with the judgment variable. None of the regions showed a statistically reliable transformation by judgment interaction or transformation by judgment by time interaction (largest $F = 171$). Thus, these effects appear to be consistent across location and color judgments.

In short, we observed a double dissociation such that a region in right parietal cortex increased more during object-based transformations and a region in left PTO cortex increased more during viewer rotations. Also, a lateral left parietal region decreased reliably more in activation during object rotations.

Overall Evoked Responses: Increases and Decreases. The analysis used here identifies regions that overall have a consistent evoked BOLD response as well as those that show a reliable main effect of timepoint. One would expect most of the brain regions involved in performing the task to show such responses, leading to a wide network of activated regions. This was the case, as shown in Figure 5. Most regions showed increases in activity relative to baseline, but a smaller network was observed that decreased in activity. This included dorsomedial frontal cortex (BA 8/9/10), the anterior cingulate gyrus (BA 32), lateral posterior parietal cortex (BA 19/39/40), and the fundus of the frontal operculum. The pattern of activity, including decreases, was highly symmetric. The left parietal area identified by the transformation by time interaction was also evident here, as was its right hemisphere homolog.

Effects of Response Time. Areas whose activity was modulated by response time gave rise to a main effect of response time or a response time by timepoint interaction. A large network of regions showed such responses, as shown in Figure 6. This network included both regions that increased in activity and regions that decreased. These responses were consistent in pattern: As response time increased, the evoked BOLD response became larger in magnitude and its mass shifted later in time relative to trial onset (see Figure 6, bottom). This is consistent with the view that these regions were characterized by neural activity whose duration was proportional to the length of time between stimulus onset and response on each trial. (One exception was a region in left motor cortex, which appeared to show a pure shift of temporal offset without a systematic increase in magnitude. This is consistent with a transient increase in activity at the time of responding.)

We also conducted a focused analysis to test for effects of response time in the regions whose activity differed for object and viewer rotations (see above). Estimated responses were averaged over voxels in each region and submitted to ANOVAs for each region. All three regions had reliable response time by time interactions, minimum $F(24, 720) = 1.84, p < .01$. The left parietal region that decreased in activity did not have a main effect of response time, $F(3, 90) = .57$; the two other regions did, minimum $F(3, 90) = 6.7, p < .001$. Thus, all three regions whose activity was modulated by which transformation was performed also were affected by response latency.

Differences Between Location and Color Judgments. In this analysis areas whose response differed between location and color judgments showed a main effect of judgment or a judgment by time

interaction. Differences between the two judgments were not the focus of the present investigation, and the design was not optimized to detect them. First, the two experiments differed in procedural aspects (particularly amount of practice) as well as in the judgment required. Second, the procedure we adopted to control differences in overall response time between object and viewer rotations does not control the between-groups comparison between location and color judgments. Therefore, we will simply summarize these data briefly. A number of regions showed an interaction between judgment and timepoint, and a smaller number showed main effects of judgment. Prominent amongst these were pre- and post-central regions, and the cerebellum. These were mainly on the left in cortex and the right in the cerebellum. Right hemisphere regions were largely more active for location judgments, whereas left hemisphere regions were generally more active for color judgments. This is likely due to the fact that 50% of responses in the location task were made with each hand, whereas all responses in the color task were made with the right hand. Consistent with this, there was a network of regions similar to those showing judgment by time interactions (but smaller in magnitude) whose activity showed greater modulation by response time for color judgments than location judgments, leading to reliable three-way interactions between response time, judgment, and timepoint. There was also one region in the fundus of the left precentral sulcus, on the anterior bank, with a reliable response time by judgment interaction; it was modulated more by response time for location judgments than for color judgments. Finally, there was one region in the right posterior cingulate gyrus with a reliable three-way interaction between transformation, response time, and judgment. The response in this region was a small decrease; examination of the response did not suggest a clear interpretation of the interaction.

None of the other interactions in the ANOVA gave rise to statistically reliable regions of activation.

Discussion

A Double Dissociation Between Object-Based Transformations and Egocentric Perspective Transformations

Across two experiments using different spatial judgments, we observed a double dissociation of increases in BOLD activity between object-based transformations and egocentric perspective transformations. In the right intraparietal cortex, BOLD activity was greater when participants were

asked to perform an object-based transformation, compared to when they performed a perspective transformation. Conversely, in the PTO cortex we observed a larger increase in BOLD signal during perspective transformations.

The observation of right parietal activity in the condition associated with object-based transformations is consistent with the majority of previous neuropsychological and neurophysiological results from mental rotation tasks (see Neural Systems for Mental Spatial Transformations, above).

The finding that left PTO cortex increased more for perspective transformations is broadly consistent with previous reports of left posterior activity in conditions associated with imagined movements of one's body or perspective (see same section above). However, the activity observed in the present case was located inferior to those regions, mostly in the superior temporal sulcus. This region likely overlapped the MT complex, an area in the human thought to be homologous to the monkey areas MT and MST, which responds selectively to visual motion (e.g., Huk, Dougherty, & Heeger, 2002; Tootell et al., 1995), and is activated imagined motion (Goebel, Khorram-Sefat, Muckli, Hacker, & Singer, 1998) and static scenes that imply visual motion (Kourtzi & Kanwisher, 2000). The left PTO region also appeared to include an adjacent area in the posterior superior temporal sulcus which has been shown to be activated specifically by biological motion (Bonda, Petrides, Ostry, & Evans, 1996; Grèzes et al., 2001; Grossman et al., 2000; Howard et al., 1996; Servos, Osu, Santi, & Kawato, 2002; Vaina, Solomon, Chowdhury, Sinha, & Belliveau, 2001). One possibility is that this activity reflects top-down activation of general motion processing and biological motion processing regions associated with the movement of one's body.

The relatively inferior location of the left PTO increases for perspective transformations may also reflect in part averaging with the near and robust parietal deactivations, such that the more superior voxels were grouped with the superior parietal region that showed a deactivation pattern. Inspection of sub-regions of this parietal area suggested this might be the case, though the statistical methods used here don't provide means to test this hypothesis rigorously.

Widespread Modulation by Task Parameters Other Than Transformation

We also observed widespread modulation of cortical activity by other experimental factors. A large number of regions showing both increases and decreases from baseline were modulated by

response latency, showing larger deflections during trials on which the participant took longer to respond. We have previously reported such widespread response time effects in a different spatial judgment task (Zacks, Ollinger et al., 2002). The demonstration of large effects of response time on cortical activity in spatial reasoning paradigms leads to two methodological conclusions. First, in order to evaluate task differences in evoked brain activity it is critical to control effects of response time. Here, this was done statistically by binning data based on response time. Second, although the finding that activity in a given brain area is affected by response time in a spatial transformation task is consistent with the conclusion that this area is directly responsible for implementing the transformation, it is by no means sufficient evidence. Modulation by response time may be evident in areas responsible for decision-making or response production, due to increased conflict on slow trials (Botvinick, Braver, Barch, Carter, & Cohen, 2001). It may also be found in early visual areas, due to increased attention to the stimulus. In the present study we saw clear evidence that motor cortex was modulated by response time, which was likely a simple reflection of the shift in the motor response relative to trial onset.

All three regions whose activity differed for the two tasks were also affected by response time. The joint finding of modulation by spatial transformation and response time dependence constitutes stronger evidence for specialized involvement in spatial transformations than response time dependence alone.

Decreases in Activity Affected by Spatial Transformation and Response Time

The evoked responses in this study included a number of regions that decreased in activity. Consistent with previous reports of regions that decrease in activity across a wide range of cognitive tasks (Shulman, Fiez, Corbetta, & Buckner, 1997), these included midline frontal and posterior structures and lateral parietal cortex. It has been suggested that these regions, which have a high basal level of activity, implement ongoing monitoring processes that are transiently suspended during execution of a demanding cognitive task (Gusnard & Raichle, 2001; Raichle et al., 2001). Unlike previous results for spatial transformations (Zacks, Ollinger et al., 2002), decreases were modulated by task parameters (response time, judgment, and transformation). One possibility is that this discrepancy reflects the fact that trials in the present study took approximately four times as long to complete, which might provide time for differential suppression of these monitoring processes.

In the present experiments the lateral parietal decreases were more extreme for object rotations than viewer rotations. (This was statistically reliable only in the left hemisphere.) Differences in BOLD decreases between the two transformations was not predicted by any of the three models reviewed in the Introduction. Gusnard and Raichle (2001) have suggested that these regions are involved in monitoring for unexpected visual targets, and decreases in this area may reflect suspension of this monitoring. However, such activity is typically lateralized to the right hemisphere, and it is not immediately clear why such processing would be selectively suspended during object rotations. One possibility is that during object rotations participants focus their attention on the central spatial field where the transformation is taking place, at the expense of the spatial periphery. Viewer rotations, which affect the whole attentional field, may not cause such a narrowing of spatial attention.

Converging Evidence From Two Tasks

The patterns of brain activity reported here indicated widespread differences between location and color judgments. These differences may have arisen from differences in the readout and comparison processes required for the spatial judgments, from the fact that location judgments were made with both hands whereas color judgments were made only with the right hand, or from the greater practice given to participants in the second experiment.

One aspect of the difference between the location and color judgments deserves particular attention. Location judgments are relatively easy for participants to learn, and were found by Presson (1982) and Wraga et al. (2000) to produce the most similar response time patterns between viewer and object rotations. However, these authors also noted that participants may “cheat” during object rotations with such problems, moving one object rather than rotating the array holistically. Questions such as the color judgment used here do not afford this strategy. However, they were much more difficult for participants to learn, requiring more practice and the exclusion of noncompliant or poorly performing participants.

In the face of these differences between the tasks, the data indicated focal differences between viewer and object rotations that were consistent across spatial judgments. We note that none of these regions showed reliable interactions involving transformation and spatial judgment in the well-powered region-wise analysis. The fact that these effects were robust in the face of possible differences in

strategy and in the face of differences in learning strengthens the case they make for the neural dissociability of the two types of spatial transformation.

An Emerging Model of Mental Spatial Transformations

The results reported here, together with previous findings, clearly rule out the unitary model of spatial transformation processing. The current results would appear to militate for the double dissociation model over the hierarchical model. The finding that brain regions in right posterior cortex are activated more during object-based transformations than during perspective transformations appears to be a robust result, having now been replicated in two other neuroimaging studies using different paradigms (Zacks, Ollinger et al., 2002; Zacks et al., 1999). The present finding of left posterior activity selective for perspective transformations completes the double dissociation. However, this result has been reported only once previously (Zacks et al., 1999) and in that case the left posterior activity was substantially more dorsal. Replication and extension of this result would clearly be desirable.

It is important to consider the differences observed here between object-based and perspective transformations in relation to the overall evoked response in those regions (Figure 4). The left posterior temporal region (marked A in the figure) was minimally active during object rotations while showing modest activity during perspective rotations. However, the right parietal region (B) was clearly active during both types of transformation, though it was more active for object rotations. Conversely, the left parietal region (C) decreased during both transformations, but did so to a greater degree for viewer rotations. If the computational units postulated by the hierarchical and double dissociation models are indeed anatomically localized, the patterns observed in B and C require further explanation. One possibility is that individual neurons in these regions change their firing rate in both types of transformation, but are more affected by one transformation than the other. A second possibility is that individual neurons in these regions tend to be active only for one of the two transformations, but that a given cortical area consists of a mixed population of the two types of cell. Such a distribution would be consistent with the finding that in a given region in posterior cortex, different cells may code spatial relations in terms of different reference frames (Colby, 1998). Thus two cortical areas may be compared in a graded fashion based on their relative proportions of cells implementing a particular reference frame. Finally, the appearance of graded, rather than all-or-none, association with one of the two

transformations may be an artifact of the limits of the spatial resolving power of fMRI. An appearance of graded activity in both transformations may result from the blurring of activity in a region activated during both transformations with one activated during only one. The ability to detect graded activations such as were observed, rather than reducing them to binary differences, is a valuable feature of event related fMRI designs with a parametric component. It will be important in future research to tease apart whether differences between BOLD activity during object-based and perspective transformations reflect differences in firing rate, in relative proportion of cell types, or spatial blurring.

Conclusions

In sum, the finding here of a double dissociation between object-based and perspective transformations suggests that mental spatial transformations are not performed by a single unitary image updating computational mechanism. In particular, it argues for a model that contains unique processing units responsible for computing object-based and egocentric perspective image updating.

Methods

Participants

Participants were recruited from the Washington University community by advertising, and were paid \$25 per hour for their time. Participants were screened for neurological disorders and contraindications for MRI scanning, and they were right-handed as assessed by the Edinburgh handedness inventory (Oldfield, 1971). Sixteen participants were tested in each experiment (age range 18-30, 14 female).

Stimulus Materials and Task Procedure

In both experiments, each trial consisted of the presentation of a picture with a brief instruction indicating which transformation to perform (object or viewer rotation). The picture remained on the screen for 7.32 s (three MRI acquisition frames). On 1/2 of the trials the next stimulus was presented immediately. On 1/4 of the trials the next stimulus was preceded by a one frame (2.44 s) interval in which a crosshair was shown in the middle of the screen and participants were asked to maintain fixation on the crosshair. On 1/4 of the trials the crosshair was presented for two frames (4.88 s). This distribution of intertrial intervals is approximately optimal for estimating evoked BOLD responses to the experimental trials using the general linear model (Ollinger, Corbetta, & Shulman, 2001).

Stimuli were presented by an Apple Power Macintosh computer (Cupertino, CA) with PsyScope experimental presentation software (Cohen, MacWhinney, Flatt, & Provost, 1993). An LCD projector was used to project the images onto a screen behind the scanner, where they were viewed through a mirror attached to the scanner head coil. This provided an image subtending approximately 16 degrees of horizontal vertical angle, and 12 vertical degrees. Responses were recorded with a custom made fiber-optic button box.

After an initial briefing and provision of informed consent, each participant was made comfortable on the scanner bed and fitted with a thermoplastic mask to reduce head motion. They were then given the button box and mirror and moved into the scanner bore. The scanning session began with approximately 30 minutes of structural image acquisition, during which the participants received training on the tasks to be performed. Instructions and sixteen practice trials were presented by the computer during the structural scans. (Participants also were trained on a different spatial reasoning task, involving judgments about pictures of human bodies. That task was performed during the first BOLD run, followed by the two BOLD runs of present interest. Those data will be reported elsewhere.) Following scanning participants were debriefed and released.

Experiment 1: Location Judgments

The two experiments differed primarily in the judgment required of participants. In the first experiment, an object was cued and the participant was asked to report the location of the object after an imagined transformation. Participants viewed an array of four blocks mounted on wooden posts at the corner of a square wooden board. The four blocks differed in color (red, blue, green, and yellow). On each trial a cue below the picture indicated (a) whether a viewer or object rotation should be performed, (b) which direction to rotate (clockwise or counterclockwise), and (c) how many degrees to rotate (0, 90, 180, 270). The participant was asked to report whether a particular block would be on their left or right after the transformation. An example of a trial, and a detailed explanation of the stimulus and cue, are given in Figure 1.

Participants responded by pressing one of two buttons on a button box. Before the beginning of the scanning session, the box was placed horizontally on the participant's lap, with their left index finger on the left button and their right index finger on the right button.

The effectiveness of this paradigm depends on participants performing the imagined transformation stipulated on each trial. Therefore, the task instructions emphasized to participants that it was important to perform each imagined transformation. Participants were instructed to prioritize performing the task as instructed and responding accurately over speed of responding. Data from two participants had to be replaced, due to equipment failure in one case and participant movement in the other.

Experiment 2: Color Judgments

In the second experiment, a location was cued on each trial and participants were asked to perform an imagined viewer or object rotation and then report the color of the object at the cued location. The object array used the same layout as Experiment 1: four blocks on posts on a square wooden board. However, the blocks were drawn in only two colors, red and green, and the board contained two of each color. As in Experiment 1, the cue on each trial indicated (a) whether a viewer or object rotation should be performed, (b) which direction to rotate, and (c) how many degrees to rotate. In addition, a fourth component of the cue indicated at which location the object's color should be reported. An example of a trial, and a detailed explanation of the stimulus and cue, are given in Figure 2.

Participants responded by pressing one of two buttons on a button box. Before the beginning of the scanning session, the box was placed on the participant's lap, oriented such that they could press both buttons comfortably with their right index finger. Half of the participants were trained to respond by pressing the upper button for "red," and half were trained to respond by pressing the upper button for "green."

Pilot testing indicated that the color judgment task required more initial practice to perform comfortably, and also that participants were more likely to report using "short cuts" in this task. To address these concerns, we added a practice session before the scanning session. Each participant came to the laboratory one to seven days before the scanning session and completed the behavioral paradigm exactly as it would later be conducted in the scanner. They received instructions, performed 16 practice trials, and then performed two blocks of 64 trials each of the task. We debriefed participants after this practice session and examined their performance. Eight participants reported employing a non-imagery strategy to solve the problems on a more than a few trials. For example, participants sometimes reported

simply mapping left onto right and vice versa when the rotation was 180 degrees. These participants were replaced. An additional eight participants failed to achieve accurate performance during the practice session (better than 85% correct for both viewer and array rotations). These participants also were replaced. Finally, data from three participants were unusable due to technical difficulties or movement during scanning. These also were replaced. In short, we employed highly selective procedures to ensure that participants were performing the desired mental spatial transformation. This may limit the generalizability of the results; the participants who satisfied our criteria were likely above average in spatial ability (and also in motivation to comply with experimental instructions). However, we can be moderately confident that the behavioral and neuroimaging data collected reflect the imagery processes of interest. Supporting this conclusion, we observed no instance of a participant achieving the requisite accuracy and reporting compliant performance during the training session who then failed either criterion during the testing session.

Magnetic Resonance Imaging

Imaging was performed on a 1.5 T Vision scanner (Siemens, Erlangen, Germany) at the Research Imaging Center of the Mallinckrodt Institute of Radiology at Washington University. Structural images were acquired using a sagittal 3D magnetization prepared rapid acquisition gradient recalled echo (MP-RAGE) T1-weighted sequence, with 1 mm³ isotropic voxels. Functional imaging was performed using an asymmetric spin-echo echo-planar pulse sequence with a flip angle of 90° and a time to echo of 37 ms, optimized for blood oxygen level dependent (BOLD) contrast (T2*) (Conturo et al., 1996; Ogawa, Lee, Kay, & Tank, 1990). Eighteen axial slices were acquired with a thickness of 7 mm and in-plane resolution of 3.75 mm. The time to recall (TR) for each slice was 135.2 ms, resulting in a total acquisition time of 2.44 s for each functional image. T2-weighted structural images were acquired in the planes of the functional images, with an in-plane resolution of .938 mm to facilitate alignment of the functional data to a standard stereotactic space.

Each functional run took 595 s (244 image acquisitions), and included 64 trials. The first 4 images were acquired before beginning the task, to allow transient signals to diminish. Each stimulus was presented at the beginning of an MR acquisition frame, and remained on the screen for 3 frames (7.32 s). Participants were instructed and trained to respond within this interval if possible. Each trial

was followed by a variable intertrial interval, as described in the Stimulus Materials and Task Procedure section above. Each participant completed two runs of each of the spatial reasoning tasks, for a total of 128 trials.

Image Analysis

Functional data were preprocessed prior to statistical analysis using methods standard for our laboratory (Zacks et al., 2001; Zacks, Ollinger et al., 2002). First, individual images for each scan were collated into a single four dimensional array. Second, timing offsets among slices were compensated for using sinc interpolation. Third, systematic odd vs. even intensity differences due to contiguous interleaved slice acquisition were removed using suitably chosen scale factors. Fourth, head motion was corrected using a six parameter rigid body realignment with 3D cubic spline interpolation. Finally, the MP-RAGE image and functional data were aligned to an atlas constructed by the methods of Lancaster et al. (2000) to conform to the coordinate scheme of Talairach and Tournoux (1988).

BOLD data were analyzed using an event-related fMRI procedure based on the general linear model. Each voxel's response was modeled as a function of three independent variables: timepoint, transformation, and response time. Timepoint represented the time of BOLD acquisition relative to the trial onset, and had 9 levels (covering a 22.96 s window following the onset of each trial). By modeling each timepoint (i.e., using delta functions as basis functions in the model), this approach avoids assumptions about the shape of the BOLD response to a trial (Josephs, Turner, & Friston, 1997). This was important in the present context, because the trials were relatively long in duration and the timecourse of neural activity within a trial could not be specified a priori. The transformation variable had two levels: object rotation or viewer rotation. The response time variable was calculated by binning response time into quartiles for each participant across tasks, leading to four levels: fast, medium-fast, medium-slow, and slow. This follows our previous procedure for a similar task (Zacks, Ollinger et al., 2002). In this procedure it is important to bin response times using the data combined across the two tasks, in order to control task differences in response time. This allows for comparisons across tasks that are not biased by differences in participants' speed of responding across the two tasks. Because degree of rotation was a manipulated variable in these experiments, it would be natural to analyze the data as a function of transformation and rotation. However, in both experiments the two tasks had small but

reliable differences in the effects of rotation on response time (see Figure 3). This means that an analysis in terms of orientation could be confounded by response time differences.

Data were analyzed using a two-stage procedure, in which estimates of the hemodynamic response to all trial types were estimated for each participant, and these estimates were submitted to a mixed analysis of variance (ANOVA). This is conceptually analogous to the typical procedure for analyzing behavioral data in designs with a large number of trials: A measure of the “typical” response to each trial type is calculated for each participant and then submitted to an ANOVA with participants as a random variable. In the first stage, a linear model was fit to each participant’s BOLD data with timepoint, transformation, and response time as independent variables. Trials on which an error was made were not included. These models also included covariates to model scan-to-scan baseline shifts and linear and nonlinear low-frequency signal drift. In fitting the model, the data were smoothed with a gaussian filter with a full width at half maximum of 6 mm. In the second stage, these estimates were submitted to a mixed ANOVA with timepoint, transformation and response time as repeated measures. The ANOVA also included a between-participants variable, judgment, with two levels: location (for Experiment 1 participants) and color (for Experiment 2 participants). F statistics from the ANOVA were transformed to Z statistics, and regions that showed a statistically reliable main effect or interaction in the ANOVA were selected according to the following criteria: a contiguous cluster of 45 contiguous voxels (1.215 cm^3) with a Z statistic greater than 3. This has been shown to correspond to a mapwise false positive rate of $p = .05$ (McAvoy, Shulman, Corbetta, Buckner, & Ollinger, under review). Analyses were computed using in-house software (Ollinger, Shulman, & Corbetta, 2001).

For visualization, thresholded functional activation data were projected onto maps of human cortical and cerebellar surfaces (Van Essen, in press) using the CARET software package (Van Essen et al., 2001).

References

- Alivisatos, B., & Petrides, M. (1997). Functional activation of the human brain during mental rotation. Neuropsychologia, 35(2), 111-118.
- Bonda, E., Frey, S., & Petrides, M. (1996). Evidence for a dorso-medial parietal system involved in mental transformations of the body. Journal of Neurophysiology, 76(3), 2042-2048.
- Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. Journal Of Neuroscience.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. Psychological Review, 108(3), 624-652.
- Bryant, D. J., Tversky, B., & Franklin, N. (1992). Internal and external spatial frameworks for representing described scenes. Journal of Memory & Language, 31(1), 74-98.
- Carpenter, M., & Proffitt, D. R. (2001). Comparing viewer and array mental rotations in different planes. Memory and Cognition, 29(3), 441-448.
- 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. Journal of Cognitive Neuroscience, 11(1), 9-24.
- Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. Behavior Research Methods, Instruments & Computers, 25(2), 257-271.
- 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(Pt 1), 89-100.
- Colby, C. L. (1998). Action-oriented spatial reference frames in cortex. Neuron, 20(1), 15-24.
- Conturo, T. E., McKinstry, R. C., Akbudak, E., Snyder, A. Z., Yang, T. Z., & Raichle, M. E. (1996). Sensitivity optimization and experimental design in functional magnetic resonance imaging. Society for Neuroscience Abstracts, 22, 7.
- Corballis, M. C. (1997). Mental rotation and the right hemisphere. Brain & Language, 57(1), 100-121.
- Corballis, M. C., & Sergent, J. (1989a). Hemispheric specialization for mental rotation. Cortex, 25(1), 15-25.
- Corballis, M. C., & Sergent, J. (1989b). Mental rotation in a commissurotomed subject. Neuropsychologia, 27(5), 585-597.
- Creem, S. H., Downs, T. H., Wraga, M., Proffitt, D. R., & Downs, J. H., III. (2001). An fMRI study of imagined self-rotation. Cognitive, Affective & Behavioral Neuroscience, 1(3), 239-249.
- Ditunno, P. L., & Mann, V. A. (1990). Right hemisphere specialization for mental rotation in normals and brain damaged subjects. Cortex, 26(2), 177-188.
- Farah, M. J. (1989). The neuropsychology of mental imagery. In F. Boller & J. Grafman (Eds.), Handbook of Neuropsychology (Vol. 2, pp. 395-413). Amsterdam: Elsevier.
- Fischer, S. C., & Pellegrino, J. W. (1988). Hemisphere differences for components of mental rotation. Brain & Cognition, 7(1), 1-15.
- 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(1), 161-171.
- Goebel, R., Khorram-Sefat, D., Muckli, L., Hacker, H., & Singer, W. (1998). The constructive nature of vision: direct evidence from functional magnetic resonance imaging

- studies of apparent motion and motion imagery. The European journal of neuroscience, 10(5), 1563-1573.
- Grezes, J., Folupt, P., Berenthal, B., Delon-Martin, C., Segebarth, C., & Decety, J. (2001). Does perception of biological motion rely on specific brain regions? NeuroImage, 13(5), 775-785.
- Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain areas involved in perception of biological motion. Journal of Cognitive Neuroscience, 12(5), 711-720.
- Gusnard, D. A., & Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. Nature Reviews Neuroscience, 2(10), 685-694.
- Harris, I. M., Egan, G. F., Sonkkila, C., Tochon-Danguy, H. J., Paxinos, G., & Watson, J. D. G. (2000). Selective right parietal lobe activation during mental rotation: A parametric PET study. Brain, 123(1), 65-73.
- Howard, R. J., Brammer, M., Wright, I., Woodruff, P. W., Bullmore, E. T., & Zeki, S. (1996). A direct demonstration of functional specialization within motion-related visual and auditory cortex of the human brain. Current biology : CB, 6(8), 1015-1019.
- Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. Journal Of Neuroscience, 22(16), 7195-7205.
- Huttenlocher, J., & Presson, C. C. (1973). Mental rotation and the perspective problem. Cognitive Psychology, 4, 277-299.
- Huttenlocher, J., & Presson, C. C. (1979). The coding and transformation of spatial information. Cognitive Psychology, 11(3), 375-394.
- Jordan, K., Heinze, H. J., Lutz, K., Kanowski, M., & Jancke, L. (2001). Cortical activations during the mental rotation of different visual objects. Neuroimage, 13(1), 143-152.
- Josephs, O., Turner, R., & Friston, K. (1997). Event related fMRI. Human Brain Mapping, 5(4), 243-248.
- Just, M. A., Carpenter, P. A., Maguire, M., Diwadkar, V., & McMains, S. (2001). Mental rotation of objects retrieved from memory: a functional MRI study of spatial processing. Journal of Experimental Psychology: General, 130(3), 493-504.
- Kosslyn, S. M. (1994). Image and brain: the resolution of the imagery debate. Cambridge, Mass.: MIT Press.
- 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(2), 151-161.
- Kosslyn, S. M., Thompson, W. L., Wraga, M., & Alpert, N. M. (2001). Imagining rotation by endogenous versus exogenous forces: distinct neural mechanisms. NeuroReport, 12(11), 2519-2525.
- Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. Journal of Cognitive Neuroscience, 12(1), 48-55.
- Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., Kochunov, P. V., Nickerson, D., Mikiten, S. A., & Fox, P. T. (2000). Automated Talairach atlas labels for functional brain mapping. Human Brain Mapping, 10(3), 120-131.
- McAvoy, M. P., Shulman, G. L., Corbetta, M., Buckner, R. L., & Ollinger, J. M. (under review). A sphericity correction for a repeated measures analysis of variance in fMRI.
- McCloskey, M. (2001). Spatial representation in mind and brain. In B. Rapp (Ed.), The handbook of cognitive neuropsychology: What deficits reveal about the human mind (pp. 101-132). Philadelphia: Psychology Press.
- Newcombe, F., & Ratcliff, G. (1989). Disorders of visuospatial analysis. In F. Boller & J. Grafman (Eds.), Handbook of Neuropsychology (Vol. 2, pp. 333-356). Amsterdam: Elsevier.
- Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proceedings of the National Academy of Science of the United States, 87(24), 9868-9872.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia, 9(1), 97-113.

Ollinger, J. M., Corbetta, M., & Shulman, G. L. (2001). Separating Processes within a Trial in Event-Related Functional MRI II: Analysis. NeuroImage, 13(1), 218-229.

Ollinger, J. M., Shulman, G. L., & Corbetta, M. (2001). Separating Processes within a Trial in Event-Related Functional MRI I: The Method. NeuroImage, 13(1), 210-217.

Parsons, L. M., Fox, P. T., Downs, J. H., Glass, T., Hirsch, T. B., Martin, C. C., Jerabek, P. A., & Lancaster, J. L. (1995). Use of implicit motor imagery for visual shape discrimination as revealed by PET. Nature (London), 375(6526), 54-58.

Pegna, A. J., Khateb, A., Spinelli, L., Seeck, M., Landis, T., & Michel, C. M. (1997). Unraveling the cerebral dynamics of mental imagery. Human Brain Mapping, 5(6), 410-421.

Piaget, J., & Inhelder, B. (1956). The child's conception of space. London: Routledge & K. Paul.

Presson, C. C. (1982). Strategies in spatial reasoning. Journal of Experimental Psychology: Learning, Memory, & Cognition, 8(3), 243-251.

Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. Psychological Review, 88(1), 16-45.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. Proceedings of the National Academy of Science of the United States, 98(2), 676-682.

Ratcliff, G. (1979). Spatial thought, mental rotation and the right cerebral hemisphere. Neuropsychologia, 17, 49-54.

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. Journal of Cognitive Neuroscience, 12(2), 310-320.

Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: a PET investigation of agency. Nature Neuroscience(5), 546-550.

Semmes, J., Weinstein, S., Ghent, L., & Teuber, H.-L. (1963). Correlates of impaired orientation in personal and extrapersonal space. Brain, 86, 747-772.

Servos, P., Osu, R., Santi, A., & Kawato, M. (2002). The neural substrates of biological motion perception: an fMRI study. Cerebral Cortex, 12(7), 772-782.

Shulman, G. L., Fiez, J. A., Corbetta, M., & Buckner, R. L. (1997). Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. Journal of Cognitive Neuroscience, 9(5), 648-663.

Sirigu, A., & Duhamel, J. R. (2001). Motor and visual imagery as two complementary but neurally dissociable mental processes. Journal of Cognitive Neuroscience, 13(7), 910-919.

Tagaris, G. A., Kim, S.-G., Strupp, J. P., & Andersen, P. (1996). Quantitative relations between parietal activation and performance in mental rotation. Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience, 7(3), 773-776.

Talairach, J., & Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system, an approach to cerebral imaging. Stuttgart: G. Thieme.

Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., Rosen, B. R., & Belliveau, J. W. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. Journal Of Neuroscience, 15(4), 3215-3230.

Tversky, B., Kim, J., & Cohen, A. (1999). Mental models of spatial relations and transformations from language. In G. Rickheit & C. Habel (Eds.), Mental Models in discourse processing and reasoning. (pp. 239-258). Amsterdam, Netherlands: North-Holland / Elsevier Science Publishers.

Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P., & Belliveau, J. W. (2001). Functional neuroanatomy of biological motion perception in humans. Proceedings of the National Academy of Science of the United States, 98(20), 11656-11661.

Van Essen, D. C. (in press). Organization of visual areas in Macaque and human cerebral cortex. In L. Chalupa & S. Werner (Eds.), The Visual Neurosciences. Cambridge, MA: MIT Press.

Van Essen, D. C., Drury, H. A., Dickson, J., Harwell, J., Hanlon, D., & Anderson, C. H. (2001). An integrated software suite for surface-based analyses of cerebral cortex. Journal of the American Medical Informatics Association, 8(5), 443-459.

Vingerhoets, G., Santens, P., Van Laere, K., Lahorte, P., Dierckx, R. A., & De Reuck, J. (2001). Regional brain activity during different paradigms of mental rotation in healthy volunteers: a positron emission tomography study. Neuroimage, 13(2), 381-391.

Voyer, D. (1995). Effect of practice on laterality in a mental rotation task. Brain and Cognition, 29(3), 326-335.

Wexler, M., Kosslyn, S. M., & Berthoz, A. (1998). Motor processes in mental rotation. Cognition, 68(1), 77-94.

Wijers, A. A., Otten, L. J., Feenstra, S., & Mulder, G. (1989). Brain potentials during selective attention, memory search, and mental rotation. Psychophysiology, 26(4), 452-467.

Wohlschlager, A. (2001). Mental object rotation and the planning of hand movements. Perception and Psychophysics, 63(4), 709-718.

Wohlschlager, A., & Wohlschlager, A. (1998). Mental and manual rotation. Journal of Experimental Psychology: Human Perception & Performance, 24(2), 397-412.

Wraga, M., Creem, S. H., & Proffitt, D. R. (1999). The influence of spatial reference frames on imagined object- and viewer rotations. Acta Psychologica, 102(2-3), 247-264.

Wraga, M., Creem, S. H., & Proffitt, D. R. (2000). Updating scenes after object- and viewer-rotations. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26(1), 151-168.

Yoshino, A., Inoue, M., & Suzuki, A. (2000). A topographic electrophysiologic study of mental rotation. Cognitive Brain Research, 9(2), 121-124.

Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., Buckner, R. L., & Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. Nature Neuroscience, 4(6), 651-655.

Zacks, J. M., Mires, J., Tversky, B., & Hazeltine, E. (2002). Mental spatial transformations of objects and perspective. Spatial Cognition & Computation, 2(4), 315-322.

Zacks, J. M., Ollinger, J. M., Sheridan, M. A., & Tversky, B. (2002). A parametric study of mental spatial transformations of bodies. NeuroImage, 16, 857-872.

Zacks, J. M., Rypma, B., Gabrieli, J., Tversky, B., & Glover, G. (1999). Imagined transformations of bodies: an fMRI study. Neuropsychologia, 37(9), 1029-1040.

Author Note

This research was supported by a grant from the McDonnell Center for Higher Brain Function at Washington University. The authors would like to thank Margaret A. Sheridan for extensive assistance in data collection, and to thank Mark McAvoy, Barbara Tversky, Abraham Z. Snyder, and the Dynamic Cognition Laboratory for scientific advice and technical assistance.

Footnotes

¹ It is important to note that both object-based and egocentric perspective transformations depend on both an intrinsic (object-based) and an egocentric reference frame, as well as an environmental reference frame. It is tempting to assume that object-based transformations depend preferentially on an object-based reference frame because that is the reference frame that is moving, and perspective transformations depend preferentially on an egocentric reference frame by the same logic. However, we know of no empirical evidence that this is the case. Moreover, in each case the relationship between all three reference frames must be updated.

² We note that the extant literature also suggests that patterns of hemispheric laterality for these tasks are complex, relative rather than absolute, and modulated by task difficulty. Simple dichotomous classifications are therefore not recommended.

Figure Captions

Figure 1. An example of a trial in Experiment 1 (location judgments). On each trial, an array of four colored blocks mounted on wooden poles on a wooden platform was presented as shown here. The arrangement of the blocks and orientation of the board was fixed throughout the experiment. One pole was colored black to indicate that it held the target object. A three-part cue was presented below the picture. The first element in the cue was either the word ARRAY, indicating that participants should perform an object rotation, or SELF, indicating that participants should perform a viewer rotation. The second element was a three-dimensional rendered arrow pointing either clockwise or counter-clockwise. The third element was a number indicating how large a rotation to imagine: 0, 90, 180 or 270 degrees. All parameters varied randomly from trial to trial: the target block, the type of transformation cued, the direction of the transformation, and the degree of rotation. All possible combinations were tested over the course of a block.

Figure 2. An example of a trial in Experiment 2 (color judgments). On each trial, an array of four red and green blocks mounted on wooden poles on a wooden platform was presented as shown here. The array was constructed with two blocks of each color placed next to each other on the board. The board could appear at any of the four possible orientations. A four-part cue was presented below the picture. The first element in the cue was either the word ARRAY, indicating that participants should perform an object rotation, or SELF, indicating that participants should perform a viewer rotation. The second element was a three-dimensional rendered arrow pointing either clockwise or counter-clockwise. The third element was a number indicating how large a rotation to imagine: 0, 90, 180 or 270 degrees. The fourth element was either the word LEFT or RIGHT, indicating the participant should report the color of the block in front to the left or right after the rotation. All parameters varied randomly from trial to trial: the orientation of the board, the type of transformation cued, the direction of the transformation, the

degree of rotation, and the location tested. All possible combinations were tested over the course of the two scanning sessions.

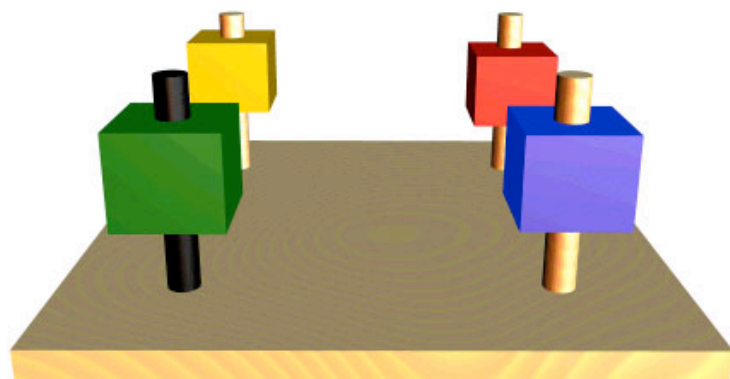
Figure 3. Response time and error rate as a function of transformation (viewer or array rotation), judgment (location or color), and degree of rotation. Panel A plots the mean of each participant's mean response time in each condition. Panel B plots the mean error rate for each condition. For both graphs, error bars represent standard errors of the mean calculated across participants.

Figure 4. Regions whose activity was differentially affected by object and viewer rotations. The top panel shows the strength of the transformation-related effect superimposed on inflated representations of the human cerebral hemispheres. Increases from baseline are plotted in red, and decreases in blue. Values are Z statistics derived from the ANOVA F statistics, thresholded to correct for multiple comparisons as described in the Methods. Regions A and B were identified based on a statistically reliable Task by timepoint interactions, whereas region C had a reliable main effect of Task. Plotted below are estimated mean evoked responses for each combination of transformation (object or viewer) and judgment (location or color). The region marked as A was located in the superior temporal sulcus at the parietal-temporal-occipital junction (Talairach coordinate of peak: -59, -50, -02; BA 21/37), and increased more for viewer rotations than object rotations. Region B was located higher in the posterior bank of the superior temporal sulcus at the intersection of the temporal and parietal lobes (Talairach coordinate of peak: -45, -67, 24; BA 39), and decreased in activity on each trial, more for object rotations than viewer rotations. Region C was located in the intraparietal sulcus (Talairach coordinate of peak: 31, -51, 45; BA 7/40); its level of activity was greater for object rotations than for viewer rotations.

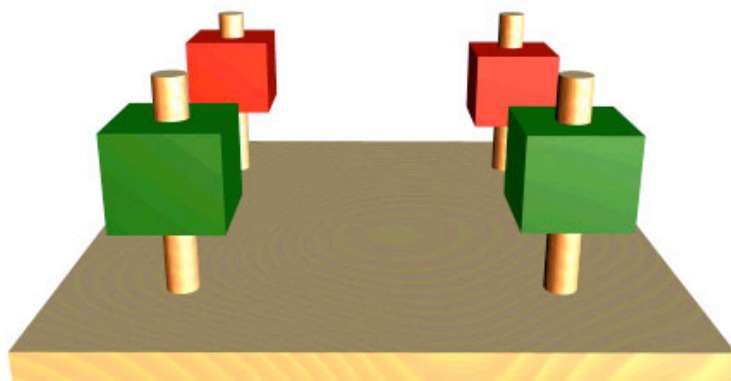
Figure 5. Overall evoked increases and decreases in BOLD signal. The figure shows the strength of the overall evoked response superimposed on inflated representations of the human cerebral

hemispheres and cerebellum. Increases from baseline are plotted in red, and decreases in blue. Values are Z statistics derived from the ANOVA F statistics for the main effect of timepoint, thresholded to correct for multiple comparisons as described in the Methods.

Figure 6. Regions whose activity was affected by response time. The top panel shows the strength of BOLD response's modulation by response time superimposed on inflated representations of the human cerebral hemispheres and cerebellum. Increases from baseline are plotted in red, and decreases in blue. Values are Z statistics derived from the ANOVA F statistics for the interaction of response time and timepoint, thresholded to correct for multiple comparisons as described in the Methods. (Inspection of the statistical map for the main effect of response time revealed a similar pattern, though reduced in magnitude.) Plotted below are estimated mean evoked responses for each level of response time in one typical region (marked A in the cortical map). This region was located in the temporo-occipital sulcus (Talairach coordinate at peak: 31, -79, 21; BA 19).

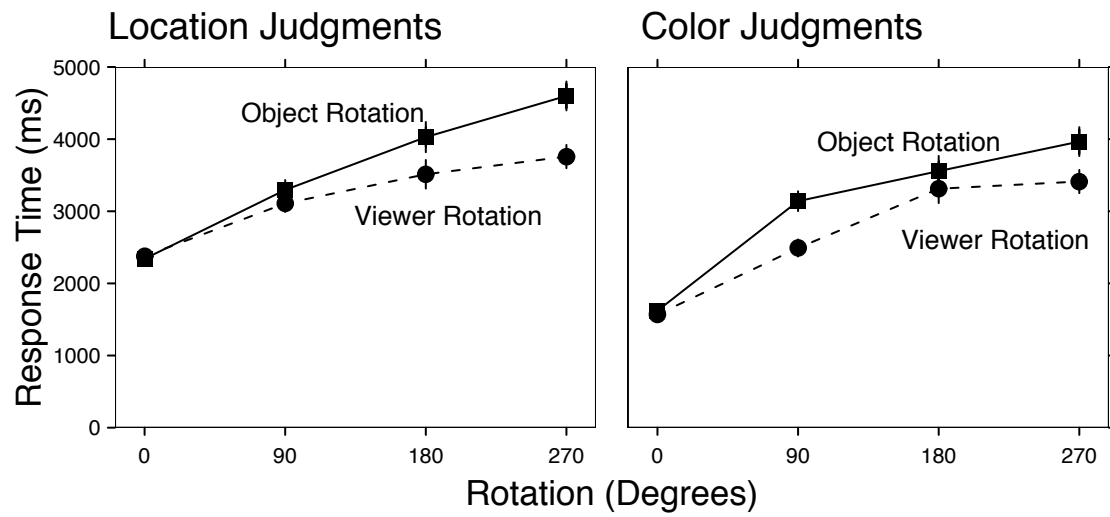


Self ↻ 270



Self  **180** **Left**

A.



B.

