Spatial Transformations of Scene Stimuli: It's an Upright World

Amy L. Shelton and Jeffrey M. Zacks

At least two different types of mental spatial transformations can be used in spatial reasoning: object-based transformations—updating an object's spatial reference frame, and perspective transformations—updating the viewer's egocentric reference frame. Pictures of human bodies have been shown to flexibly engage these systems for different tasks, suggesting that the neural systems implementing these two transformations may be adapted for different spatial reasoning situations. In the present study, four experiments tested how pictures of immersive spaces—rooms—selectively engage different transformations. Response latency patterns suggested that the visual system quickly interprets pictures of scenes using two dissociable spatial transformations: object-based transformations, which re-orient the picture with respect to upright in the world, and perspective transformations, in which the viewer imagines themselves taking up a position within the depicted scene.

A. L. Shelton (✉)
The Johns Hopkins University, Baltimore, USA
e-mail: ashelton@jhu.edu

J. M. Zacks
Washington University in St. Louis, St. Louis, USA

Reasoning About Bodies in Space

Much of daily human activity involves reasoning about the changing relationships among one's body, other objects, and the world. Although several different types of mental spatial transformation may be possible, two distinct classes have been identified in the literature: object-based and perspective transformations [1]. Object-based transformations involve the mental rotation or manipulation of an object. This transformation is akin to a stationary observer "watching" a moving object in space. For example, when packing a car for a trip, one might imagine the different ways the suitcases can be turned to fit into the space. Perspective transformations are transformations of oneself in space. This is akin to an observer mentally transporting himself to a new perspective and "seeing" the world from this new view.[1] For example, when lecturing to a large class, the instructor typically faces the audience. When directing the audience to a particular portion of screen, the instructor may indicate a direction ("bottom right"). To specify the location correctly, the instructor may imagine what the screen looks like from the perspective of an audience member. Previous work has suggested that these two types of transformations are distinguishable both in terms of the behavioral profiles they produce [2] and the neural substrates that participate in them [3–6].

Behaviorally, object-based and perspective transformations have been distinguished by their temporal dynamics [2], and [7]. For example, when Wraga and colleagues asked participants to make spatial judgments about learned arrays, participants were either instructed to use array rotations (object-based transformations) or viewer rotations (perspective transformations). The results revealed that response latency increased as a function of orientation for the array rotations, whereas response latency for viewer rotations was flatter. As an alternative approach, Zacks and colleagues used different tasks to induce the transformations. Participants were asked to view images of bodies rotated in the picture plane. They were asked either to determine whether two bodies at different angular disparities had the same or different arm extended (same-different task) or to determine whether a single rotated body had its left or right arm extended (left-right task). The same-different task was hypothesized to induce object-based transformations because the judgment depends on the relationship between the two bodies irrespective of the observer. On the other hand, the left-right task was hypothesized to induce perspective transformations because left and right have clear meaning in the egocentric reference frame; by aligning oneself with the body stimulus, one can easily assess whether the extended hand is now on one's left or right. Response latency patterns supported this hypothesis. For same-different judgments response times increased monotonically with increasing stimulus orientation, replicating previous results for object-based transformations [8]. For left-right judgments response times were largely equivalent across different stimulus orientation, consistent with the pattern obtained when participants were explicitly instructed to perform perspective transformations with

---

[1] "Seeing" is not meant to restrict this to visual experience of space; the literature is agnostic as to whether the information available from an imagined perspective may be multimodal or amodal.

similar stimuli [9]. In a subsequent study, participants were given instructions that mismatched the hypothesized "natural" transformation for a given tasks; results revealed impaired performance and response latencies that resembled the other task [10]. These findings were further supported by participants' introspective reports.

Object-based transformations and perspective transformations have also been dissociated neurophysiologically. In numerous studies, object-based transformations, such as mental rotation, have been associated with the inferior parietal cortex, particularly in the right hemisphere [4, 11–22]. Although perspective transformations have received less attention in the literature, left posterior regions have been implicated in tasks that likely require perspective transformations (e.g., [23] and [24]). In a direct comparison of object-based and perspective transformations of bodies, [4], found that regions at the junction of the temporal, parietal, and occipital lobes (TPO) in the right hemisphere were selectively activated by object-based transformations [5]. Complementing this single dissociation, two studies using variants of the array and viewer rotation tasks found double dissociations between object-based and perspective transformations ([3] and [6]). In both studies, left TPO cortex was selectively activated by perspective transformations, and right parietal cortex was selectively activated by object-based transformations. (In [3], a number of other regions were selectively activated during object-based transformations as well.)

The existence of dissociations between object-based transformations and perspective transformations in their behavioral profiles and neural correlates has led to the suggestion that the brain has (at least) two systems for spatial transformations: one that supports object-based transformations and one that supports perspective transformations [1]. Zacks and Tversky, [10], proposed that the engagement of a particular system should depend not only on the task but also on the type of stimulus being manipulated. They contrasted bodies, which can move independently or serve as the source of viewpoint, with small inanimate objects, for which the independent movement or manipulation is far more common; rarely would one ask what the world looks like from an object's perspective. Consistent with the predictions, they found evidence that participants flexibly used perspective transformations or object-based transformations to make judgments about pictures of bodies, whereas participants depended heavily on object-based transformations when making judgments about manipulable inanimate objects.

Zacks and Tversky, [10], provided clear evidence for distinctions within the domain of discrete objects. However, spatial reasoning is not restricted to this class of stimuli and often entails making judgments in a multi-object environment (e.g., maneuvering a car in a parking lot). The present study was therefore designed to investigate spatial transformations of scene stimuli in the form of images of rooms. Unlike bodies or other objects, rooms are stationary, upright entities. They do not undergo movement, but can serve as the loci of potential perspectives. Thus, the tuning of the object-based transformation system would be expected to be relatively unresponsive during judgments about rooms, and the perspective transformation system would be expected to be responsive during any room judgments.

The four experiments described here were designed to test the hypothesis that rooms would selectively evoke the use of perspective transformations. We contrasted

rooms to bodies, which have consistently shown both object-based and perspective transformations depending on the spatial reasoning task. Participants made same-different and left-right judgments about pictures of rooms and pictures of bodies. If the two systems for spatial transformations are readily available for either type of stimulus, one would expect that both stimuli would yield object-based performance for the same-different task and perspective performance for the left-right task [2]. Alternatively, if rooms selectively engage the perspective transformation system, one would expect that they would tend to produce flatter slopes than those observed with bodies for the same-different task.

Experiment 1

Experiment 1 was designed to compare performance on the same-different and left-right tasks separately for bodies and rooms. All participants performed both tasks with both types of stimuli, and the patterns of performance were examined in the context of object-based and perspective transformations. Bodies provided the control condition. Based on previous studies, we expected that participants would perform object-based transformations in order to solve the same-different task and perspective transformations in order to solve the left-right task. We therefore predicted increasing response times with increasing stimulus orientation in the same-different task, and flatter response latency profiles in the left-right task, replicating previous results.

For rooms, we hypothesized that participants would be less likely to perform object-based transformations, even in the same-different task. This led to the specific prediction that the relationship between stimulus orientation and response time during the same-different task would be weaker for rooms than for bodies. That is, we expected that the same-different task would show a flatter response latency curve for rooms than for bodies, and this curve should be similar to the response latency curve observed for the left-right task (for both rooms and bodies).

Method

Participants

Sixty-five participants (33 male) from the Stanford University community volunteered in return for experimental credits in Psychology courses. All participants reported normal or corrected-to normal vision and hearing.

Materials

Body stimuli were line drawing images of human bodies with one arm extended in two different poses. Images were created at 12 different picture plane rotations

ranging from 0° (upright) to 330° in 30° increments. Room stimuli were created by first creating two different rooms in a desktop virtual reality program (Virtus Walkthough Pro, Virtus Corporation, Cary, NC). Virtual snapshots were then taken of each room with a plant placed on either the right or left side of a door in the center of the back wall of the room. Given that rooms are inherently scenes, the images had to be cropped. To prevent the image boundaries from providing a salient reference frame specifying orientation, the images were cropped in a circular window as if looking through a large porthole. Rooms were then rotated to create images from the same 12 orientations as for the bodies. For both types of stimuli, angular disparity in the left-right task corresponded to this angular disparity from upright. Half of the trials were right-handed (right arm extended or plant to the right), and half were left-handed. For the same-different tasks, the stimuli at the 12 different orientations were paired to create 12 different angular disparities ranging from 0° to 330°, in 30° increments (e.g., the 30° and 90° images might be paired to create a 60° test trial). Half the trials had the arm or plant on the same side, and half had it on different sides.

Procedure

All participants performed 112 trials in each of the four combinations of task (same/different and left/right) and stimulus (rooms and bodies), in a counterbalanced order on a Macintosh computer running PsyScope software [25]. Prior to testing, participants received instructions for each task in written form. In the same-different task they were told to press the left button for "same" and the right button for "different." In the left-right task they were told to press the left and right button for "left" and "right" responses, respectively. For bodies, left and right were defined by the arm of the figure, whereas in rooms participants had to determine where the plant would be upon entering the door. Participants were then given 10 practice trials that were identical to the actual trials just prior to each task. For each trial, a cue appeared ("Hit any button to go on"). A fixation cross appeared for stimuli for 1500 s followed by the test stimuli presented either in pairs (for same-different) or alone (for left-right) (see Fig. 1). If the response was correct, the computer indicated so with a pleasant tone and the trial ended. If the response was incorrect, the computer buzzed and the stimuli remained on the screen until the correct response was entered. Both the response latency (to the first response) and the accuracy were recorded.

Results

Three participants (2 male) were removed before analyses due to incomplete data or error rates exceeding 30 % in any task block or 15 % overall. For the remaining 62 participants, error rates were low in both judgment tasks (4.3 % for left-right, 5.5 % for same-different). The small task difference in errors was statistically significant, $F(1, 61)=4.68$, $p=0.03$, but did not interact with stimulus set, $F(1, 61)=0.002$,
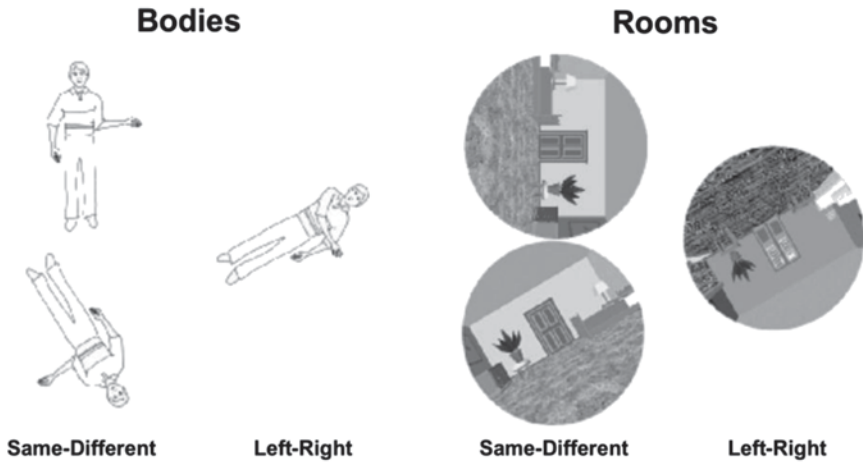
Bodies   ## Rooms

Same-Different   **Left-Right**   **Same-Different**   **Left-Right**

Fig. 1 Examples of the same-different and left-right tasks with pictures of bodies and rooms in Experiment 1. Room stimuli were presented in color during the actual experiment (Answers from left to right: same, right, different, left)

$p=0.96$. The main effect of stimulus set on error rate did not approach statistical significance, $F(1, 61)=0.0021$, $p=0.96$.

All response time analyses were performed on correct responses only. Prior to analysis, outlying response times were trimmed by excluding observations 3 standard deviations from a participant's mean for a given combination of stimulus set and judgment task. This led to the elimination of 1.9% of correct responses.

Two analyses were performed, following the approach of Zacks and Tversky, [10]. First, each participant's mean response times were calculated as a function of stimulus set, task, and orientation. These were then subjected to an analysis of variance (ANOVA) with stimulus type, task, and orientation as within-subject factors. As can be seen in Fig. 2, response times for same-different judgments about bodies increased substantially with increasing stimulus orientation, but response times for left-right judgments about bodies did not. For judgments about rooms, this difference was attenuated and both tasks showed smaller increases in response time with increasing stimulus orientation. This led to a statistically significant three-way interaction between stimulus orientation, task, and stimulus set, $F(6, 366)=18.2$, $p<0.001$. It also led to significant main effects of orientation, $F(6, 366)=85.1$, $p<0.001$, and of task, $F(1, 61)=54.8$, $p<0.001$, and to a significant two-way interaction between task and orientation, $F(6, 366)=6.345$, $p<0.001$. The two-way interaction between task and stimulus set approached but did not reach statistical significance, $F(1, 61)=3.71$, $p=0.059$. The main effect of stimulus set was not significant, $F(1, 61)=1.15$, $p=0.29$; nor was the interaction between stimulus set and orientation, $F(6, 366)=0.73$, $p=0.63$.

To more precisely characterize the relationship between stimulus orientation and response time we computed, for each participant, the Pearson correlation between orientation and response time for each combination of stimulus set and judgment
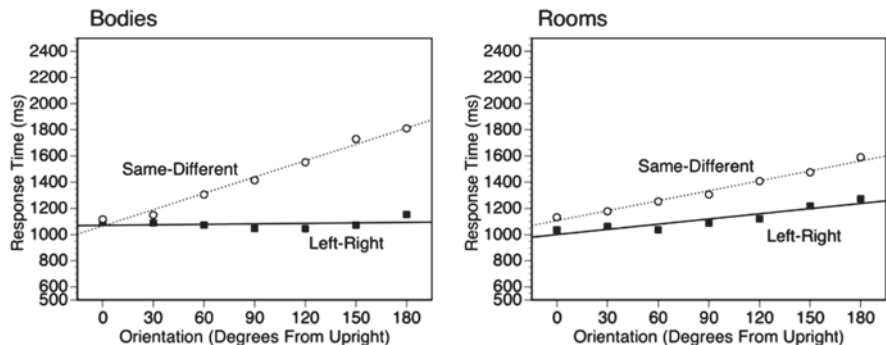
Fig. 2 Response time as a function of stimulus orientation for each combination of judgment (same-different or left-right) and stimulus set (bodies or rooms) in Experiment 1. Each point is the mean across participants of the mean within-participant trimmed response time. The lines are least-squared regression fits
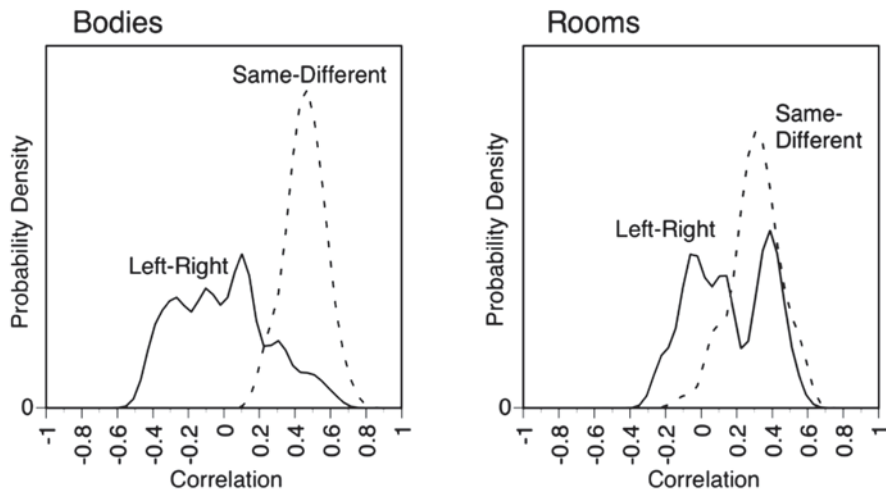

Fig. 3 Distributions of correlations between stimulus orientation and response time, as a function of the judgment made (same-different or left-right) and the stimulus set (bodies or rooms). Data are from Experiment 1. (For this figure and Figs. 5, 8 and 10, density functions were calculated by kernel estimation with a Gaussian kernel of bandwidth 0.05.)

task. The correlation gives a straightforward characterization of the strength of the linear relationship between stimulus orientation and response time. To the extent that response times increase with increasing orientation, this correlation will be positive (see [10]). The distribution of correlations for each condition is plotted in Fig. 3. As can be seen in the figure, for judgments about bodies, a clear task difference was observed: Correlations were robustly positive for same-different judgments, but centered on zero for left-right judgments. For judgments about rooms, correlations tended to be somewhat positive for both same-different and left-right

judgments. This led to a significant main effect of task, $F(1, 61)=102.3, p<0.001$, and a significant interaction between stimulus set and task, $F(1, 61)=80.2, p<0.001$. The main effect of stimulus set was not significant, $F(1, 61)=0.11, p=0.74$. Follow-up t-tests revealed that the difference between the left-right and same-different tasks was significant for both the body stimulus set, $t(61)=13.2, p<0.001$, and for the room stimulus set, $t(61)=4.22, p<0.001$. Correlations were significantly positive for all conditions except for left-right judgments about bodies. [For that condition, $t(61)=-0.16, p=0.87$. For the other conditions, the smallest $t(61)$ was 5.7, $p<0.001$.]

To summarize, when making judgments about bodies, a strong difference was observed between same-different and left-right judgments: response time increased with increasing stimulus orientation for same-different judgments, but not for left-right judgments. However, for judgments about rooms, this task difference was attenuated; response time increased modestly for both left-right and same-different judgments. This led to a significant three-way interaction in the analysis of mean response times, and a significant two-way interaction in the analysis of correlations.

Discussion

The results of Experiment 1 were consistent with the hypothesis that room stimuli selectively engage perspective transformations: The relationship between stimulus orientation and response time in the same-different task was weaker for rooms than for bodies (flatter curve). However, there was an unanticipated pattern to these data. Although the stimulus orientation and response time relationship was attenuated when making same-different judgments about rooms, it was not fully orientation-independent; instead, response times increased significantly with orientation. Even more surprisingly, response times also increased significantly with orientation when making left-right judgments about rooms. Taken together, these results suggest that rooms differed from bodies in their engagement of object-based and perspective transformations, but they also raised questions about how rooms might show an attenuated increase in response time with stimulus orientation in both tasks.

One possibility is that participants performed perspective transformations for both the left-right and same-different tasks with rooms, but used trajectories that did not lead to the orientation-independent performance found for perspective transformations of front-facing bodies rotated in the picture plane [9]. The spatial framework of the room may have constrained participants' imagined perspective transformations, for example, if they imagined themselves deviating from the simplest path to avoid imagining themselves intersecting the objects near the door. On this interpretation, the data would provide support for the hypothesis that when participants thought of the room stimuli as immersive spaces, this produced a bias to solve the spatial reasoning problems using perspective transformations.

However, these data could also be explained by proposing that participants performed *object-based* transformations in both the left-right and same-different tasks

with rooms. Perhaps presenting room stimuli as pictures induced participants to first resolve the discrepancy between the picture they were presented and the gravitational upright that they were experiencing by treating the picture of the room as an object unto itself and mentally rotating it to upright. Although experiences in which rooms rotate are presumably quite rare, experiences in which the reference frame of a room is misaligned with the gravitational upright are also atypical. In most cases where we see a room from an odd viewing angle, it is due to our own misorientation relative to the gravitational upright.

A third alternative is that the increased latency as a function of orientation reflects a natural tendency to upright a scene stimulus. Not only are actual rooms usually experienced in alignment with gravity, *pictures* of rooms are generally viewed such that the room depicted is aligned with the gravitational upright or the egocentric front of the viewer. (For example, paintings of rooms in museums generally are hung with the depicted floor and ceiling aligned with the actual floor and ceiling, and pictures of rooms in books are generally printed with the floor toward the bottom of the page and the ceiling toward the top.) As such, seeing a rotated scene stimulus may cause the participant to rapidly engage in an object-based transformation of the stimulus to reorient the depiction to upright, regardless of the task. By this explanation, there may be a bias to use perspective transformations for the spatial reasoning, but response latencies may be slowed down by the need to upright the image as well. In this sense, the object-based rotation of the depiction to upright is essentially interference.

In Experiment 2 we attempted to distinguish these three interpretations by directly instructing participants to perform perspective transformations with both body stimuli and room stimuli.

Experiment 2

To directly characterize the relationship between stimulus orientation and response time for perspective transformations with the room stimuli, we explicitly instructed participants to perform perspective transformations with those stimuli. Following the manipulations used by Parsons (1987), we asked participants either to perform the left-right task or to imagine a perspective transformation for both rooms and bodies. We predicted that participants performing the left-right task would show the same pattern of performance found in Experiment 1, with bodies showing a flat slope and rooms showing a slight increasing relationship. If participants who were asked to imagine performing perspective transformations showed the same pattern, this would support the hypothesis that the participants in Experiment 1 had tended to use perspective transformations when performing the left-right and same-different tasks with rooms. However, if participants who were asked to imagine performing perspective transformations showed orientation-independent performance for both bodies and rooms, this would suggest the participants in Experiment 1 had tended to use *object-based* transformations when performing the left-right and same-different tasks with rooms.

Method

Participants

Thirty-two participants (16 male) from the undergraduate population at Washington University volunteered in return for $ 10 or partial fulfillment of a course requirement.

Materials

The materials were the same room and body stimuli used in Experiment 1.

Procedure

Participants were randomly divided into two groups and asked to perform two different tasks. In the *left-right* group, participants performed the left-right task described in Experiment 1. In the *imagine* group, participants were asked simply to "imagine [themselves] standing in the door of the room," and "form a vivid mental picture of [themselves] lined up with the door as shown on the screen." They were instructed to press a button when they had formed the image. Participants in each group performed 112 trials with each type of stimuli. The room and body blocks were counterbalanced across participants. Stimuli were presented in the same manner as in Experiment 1, except that there was no correct or incorrect response in the imagine task, and so no feedback was provided.

Results and Discussion

For the group that performed the left-right task, error rates were comparable to those in Experiment 1. They were low (4.6 % for bodies, 3.9 % for rooms) and did not differ significantly across stimulus sets, $t(15)=0.91$, $p=0.38$. Response time data from error trials were excluded, and the response time data were trimmed as described in Experiment 1, which resulted in the elimination of 1.5 % of correct responses.

The response time data were analyzed using the same approach as for Experiment 1. First, each participant's mean response times were calculated as a function of group, stimulus set, and orientation, and these were submitted to an ANOVA with group as a between-participants factor and stimulus set and orientation as repeated measures. As Fig. 4 shows, response time was relatively independent of orientation for both types of judgment about bodies, but increased somewhat with increasing orientation for both types of judgment about rooms. This led to a significant main effect of orientation, $F(6, 180)=3.12$, $p=0.006$, and a significant orientation-by-stimulus set interaction, $F(6, 180)=8.25$, $p<0.001$. Performance of the imagine task
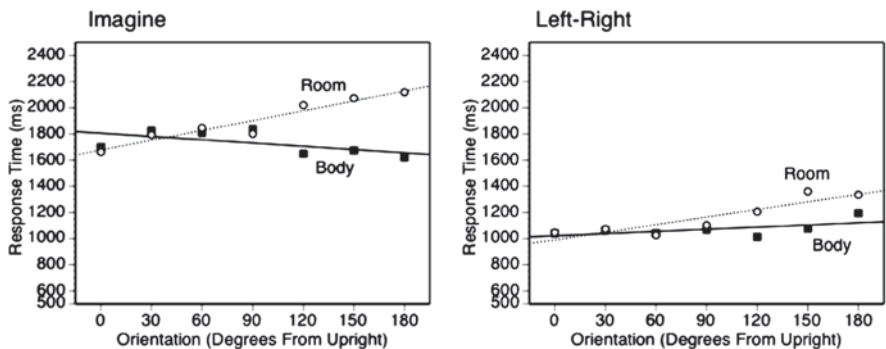
Fig. 4 Response time as a function of task (imagining oneself in the position indicated by the picture, or making a left-right judgment about the picture) and stimulus set (bodies or rooms) in Experiment 2. Each point is the mean across participants of the mean within-participant trimmed response time. The lines are least-squared regression fits
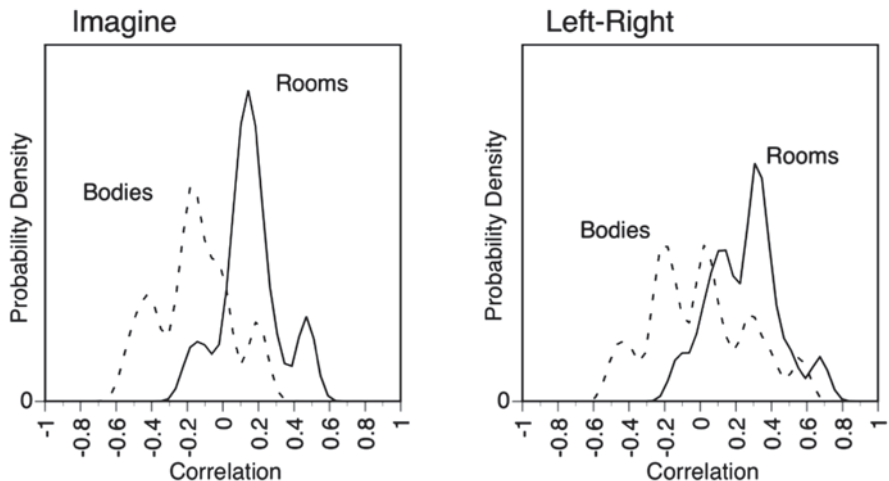

Fig. 5 Distributions of correlations between stimulus orientation and response time, as a function of task (imagining oneself in the position indicated by the picture, or making a left-right judgment about the picture) and stimulus set (bodies or rooms) in Experiment 2

was overall slower than performance of the left-right task, leading to a significant main effect of group, $F(1, 30)=4.93$, $p=0.034$. None of the other main effects or interactions approached statistical significance (largest $F=1.70$). To follow up the significant two-way interaction, we conducted separate ANOVAs for each of the two groups. These showed that the orientation-by-stimulus set interaction was significant for both the left-right group [$F(6, 90)=4.23$, $p<0.001$] and the imagine group [$F(6, 90)=5.15$, $p<0.001$]. Separate ANOVAs for each of the four combinations of group and stimulus set showed that the effect of orientation was statistically significant for both the left-right and imagine tasks with rooms [left-right: $F(6,$

90)=6.79, $p<0.001$; imagine: $F(6, 90)=3.30, p=0.006$]. There was no significant effect or orientation for either task when performed with pictures of bodies [left-right: $F(6, 90)=0.99, p=0.44$; imagine: $F(6, 90)=1.74, p=0.12$]

To further characterize the relationship between orientation and response time across the experimental conditions, we calculated the correlation between orientation and response time for each participant, for each of the two stimulus sets. As can be seen in Fig. 5, correlations for both groups were higher when making judgments about pictures of rooms than when making judgments about pictures of bodies. This led to a significant main effect of stimulus set, $F(1, 30)=41.5, p<0.001$. Correlations also were slightly higher for the group that performed the left-right task, leading to a marginally significant main effect of group, $F(1, 30)=3.86, p=0.06$. The group-by-stimulus set interaction did not approach statistical significance, indicating that the two groups showed similar stimulus set effects, $F(1, 30)=0.46$, $p=0.51$. T-tests confirmed that for both groups, the difference in correlations between the two stimulus sets was significant [left-right: $t(15)=4.14, p<0.001$; imagine: $t(15)=4.96, p<0.001$]. Correlations between orientation and response time for judgments about room pictures were significantly positive, $t(31)=6.03, p<0.001$. For judgments about bodies, the correlations were slightly negative, and this difference approached statistical significance, $t(31)=-1.82, p=0.08$.

In sum, response time patterns when participants were explicitly asked to imagine themselves in a particular position were similar to response time patterns when participants were asked to make left-right judgments about the same position. For judgments about bodies, response time was essentially independent of stimulus orientation both when participants made left-right judgments, and when they were explicitly instructed to imagine themselves in the position of the body, replicating previous findings [9]. For judgments about rooms, response times increased with increasing stimulus orientation, both for left-right judgments and for imagined movements. These nearly identical patterns replicate those observed for the left-right task in Experiment 1, ruling out the possibility that participants were using strictly object-based transformations on the left-right task.

However, these data do not explain why these perspective transformations for rooms should be more sensitive to orientation than perspective transformations of bodies. More specifically, why should the time to imagine one's self in the door of a room should differ from the time to imagine one's self in that same position when the to-be-assumed position is cued by a picture of a body standing alone? The interference explanation introduced previously may account for this oddity. That is, when shown a depiction of a room in an atypical orientation, participants may perform an object-based transformation to upright the stimulus *in addition* to the transformations that are required for appropriately completing that task. By this explanation, the representation of the space depicted by the picture evokes a tendency to perform a perspective transformation, but the representation of the picture as a picture evokes a tendency to upright the picture using an object-based transformation. If this is correct, then the surface properties of the room pictures should be necessary and sufficient to evoke the object-based uprighting transformation.

Fig. 6 Example of the combined room/body pictures used in Experiment 3. Stimuli were presented in color in the actual experiment

Experiment 3 provided a rigorous test of the interference hypothesis using exactly the same stimuli to depict rooms and bodies. In this experiment, participants made left-right or same-different judgments about pictures that included both a body and a room, but were instructed to attend either to the spatial reference frame of the body, or of the room.

Experiment 3

Experiment 3 replicated Experiment 1, except that the stimuli were identical in the rooms and bodies conditions. We created new stimuli that included a body standing in a room (Fig. 6), and then manipulated the instructions to direct attention either to the room or to the body. These instructions did not tell the participant what type of transformation to use, but rather indicated what aspect of the stimulus (the room or the body) was relevant to the task.

By holding the physical stimuli constant, Experiment 3 allowed us to directly test what gave rise to the differences observed for rooms versus bodies. First, if the difference in response latency patterns between rooms and bodies on the same-different task resulted from the preferential engagement of perspective transformations when reasoning about rooms compared to bodies, then the response latencies should again show a more pronounced monotonic relationship to orientation for bodies than for rooms. Second, holding the stimulus constant allowed us to test how the stimulus differences may have affected the patterns of performance, particularly on the left-right task. In Experiment 2, when participants were asked to perform perspective transformations with rooms, small but significant increases in response time with increasing stimulus orientation were observed. In Experiments 1 and 2, response time increased slightly but significantly with increasing orientation for

left-right judgments. This result differed from the pattern observed for left-right judgments about bodies in the same spatial configuration, in Experiment 1 and previous research [2] and [9]. We hypothesized that room stimuli might invoke some automatic transformation to upright, irrespective of the reference frame for making the judgment. Based on this hypothesis, we predicted that in the left-right task in Experiment 3, we would observe small but significant increases in response time with increasing stimulus orientation for *both* the body and room conditions.

Method

Participants

Sixty-four participants (32 male) from the Johns Hopkins Community volunteered in return for extra credit in Psychology and Cognitive Science courses or for monetary compensation.

Materials

Using Poser 3.0 software (Curious Labs, Santa Cruz, CA), rendered images of rooms (2 different rooms) and bodies (2 different poses) were created. In the images, a lamp was placed either to the left or right of the doorway and the body had either the left or right arm extended (see Fig. 6). The two rooms and two poses were combined such that room, pose, left or right lamp, and left or right arm were completely counterbalanced. As in the previous experiments, the images were cropped in a circular aperture, and images were taken at 12 different orientations ranging from 0° (upright) to 330° in 30° increments. These images were combined to create the different angular disparities for the same-different task.

Procedures

All participants performed the same-different and left-right tasks with both rooms and bodies, completing 112 trials of each combination. The trials were blocked hierarchically, first by attentional instruction and then by task. Participants were assigned to groups according to the complete counterbalancing of instruction and task within instruction. For the attend-rooms instructions, participants were asked to determine whether the lamp was on the same side of the door in two images in the same-different task and asked to determine whether the lamp was on the right or left of the door when entering the room in the left-right task. For the attend-bodies instructions, participants were asked to determine whether the two figures had the same arm extended in the same-different task and asked to determine which of the figure's arm was extended in the left-right task. The correspondence between the
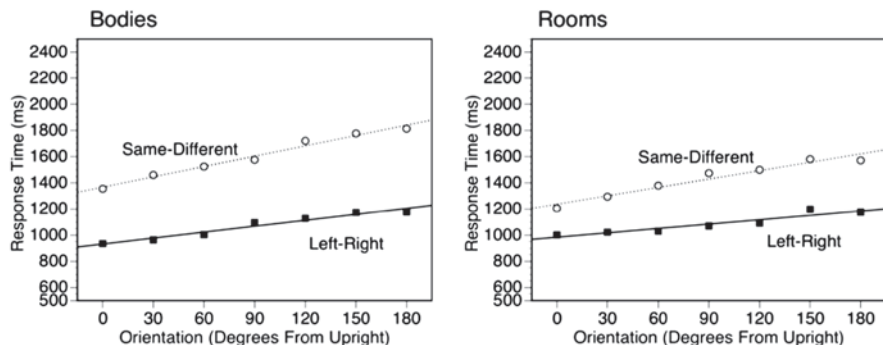
Fig. 7 Response time as a function of stimulus orientation for each combination of judgment (same-different or left-right) and object about which the judgment was made (bodies or rooms) in Experiment 3. Each point is the mean across participants of the mean within-participant trimmed response time. The lines are least-squared regression fits

location of the lamp and the extended arm was counterbalanced, such that attending to the wrong stimulus would produce chance performance. Trial procedures were identical to those used in Experiment 1.

Results and Discussion

Error rates were low (4.9%) and did not differ significantly across conditions. [There was a marginally significant task-by-stimulus set interaction, such that error rates were slightly lower in the left-right task with bodies than the other three conditions, but this did not reach statistical significance, $F(1.64)=3.63$, $p=0.06$. Neither main effect was significant: For the effect of task, $F(1, 63)=1.04$, $p=0.31$; for the effect of stimulus set, $F(1, 63)=1.96$, $p=0.17$.]

   Response time data were trimmed and analyzed as described for Experiment 1. First, mean response times were calculated for each participant for each combination of task, instructions, and orientation, and these mean response times were submitted to a repeated measures ANOVA. As can be seen in Fig. 7, when participants attended to the bodies there was a large difference between the same-different and left-right tasks, such that response times increased more with increasing orientation during the same-different task. When participants attended to the space of the rooms, this difference was attenuated. This pattern led to a three-way interaction between task, instructions, and orientation, $F(6, 378)=3.46$, $p=0.002$, and replicated the pattern observed in Experiment 1. However, response times increased with increasing orientation for all four conditions, including a small but significant increase for left-right judgments when attending to the body [overall $F(6, 378)=71.6$, $p<0.001$; smallest individual-condition $F(6, 378)=10.2$, $p<0.001$]. Overall, responses were slower in the same-different task, $F(1, 63)=85.8$, $p<0.001$, and slower when attending to the bodies than when attending to the space of the rooms, $F(1, 63)=7.17$,
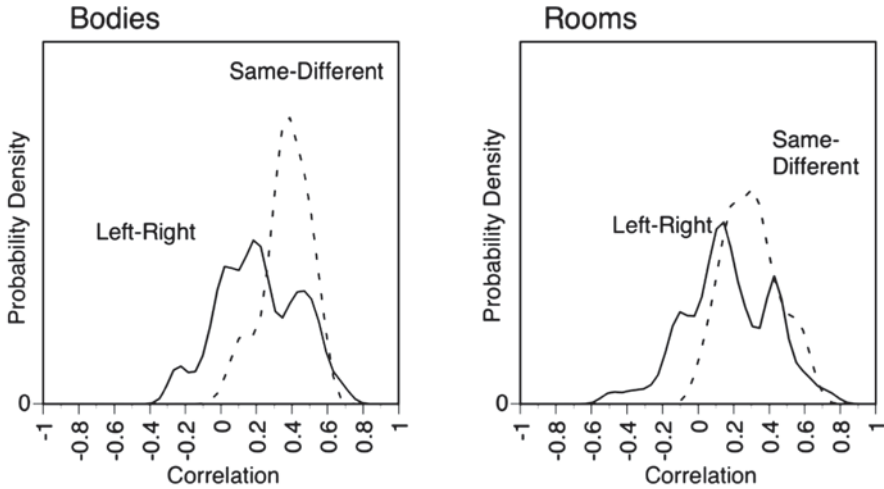
Fig. 8 Distributions of correlations between stimulus orientation and response time, as a function of the judgment made (same-different or left-right) and the object of the judgment (bodies or rooms) in Experiment 3

$p=0.009$. All three two-way interactions were also significant, smallest $F=9.68$, $p<0.001$.

Analyses of the correlations between stimulus orientation and response time largely converged with the ANOVAs on response time. Correlations were significantly higher for same-different judgments than for left-right judgments, $F(1, 63)=26.5$, $p<0.001$, and were higher when participants attended the bodies than when they attended to the rooms, $F(1, 63)=16.7$, $p<0.001$ (Fig. 8). For all four conditions, the correlations between stimulus orientation and response time were significantly positive, smallest $t(63)=5.46$, $p<0.001$. However, the correlation analyses failed to provide additional evidence that the relationship between stimulus orientation and response time depended on the interaction of task and instructions; this was not statistically significant, $F(1, 63)=0.69$, $p=0.41$

In short, the results replicated the main finding of Experiment 1: The relationship between stimulus orientation and response time depended both on the judgment participants were asked to make, and on the target of that judgment.

The same-different task revealed the same pattern of stronger orientation dependence for bodies than for rooms, arguing against the possibility that this difference was due to stimulus differences alone in Experiment 1. This predicted difference supports the hypothesis that participants used more perspective transformations and fewer object-based transformations when making same-different judgments about rooms compared to bodies. The same-different judgments for rooms as well as the left-right judgments for both stimuli have patterns nearly identical to that observed in Experiment 2, when participants were directly instructed to imagine making a perspective transformation and cued with a picture of a room. This pattern further

supports the preferential use of perspective transformations for bodies in the left-right task and rooms more generally.

The weak but significant increase in response latency as a function of stimulus orientation in both versions of the left-right task suggest an influence of the room stimulus, irrespective of the focus of the transformation. This finding is consistent with the hypothesis that pictures of rooms at atypical orientations tend to evoke object-based transformations to mentally upright the pictures, in addition to perspective transformations that may be performed to accomplish the left-right judgment.

The strong influence of the room in the bodies condition supports the claim that the uprighting is occurring in a task-irrelevant manner. However, it is notable that the rotation of these stimuli was locked such that the body and room rotated together. If pictures of rooms at atypical orientations evoke object-based transformations to upright them, and if people also tend to perform perspective transformations to make left-right judgments about a potential viewpoint from a body within a room, then manipulations of the room's orientation and the body's orientation should have separable effects: in this paradigm, response times should increase with increasing rotation of the room, but not with increasing rotation of the body. Experiment 4 provided a stronger test of the task-independent uprighting account by asking participants to make left-right judgments about bodies only and rotated the room or the body independently.

Experiment 4

If effect of orientation on response time with room stimuli across both tasks reflects a task-irrelevant tendency to upright a room stimulus, then this effect should occur even if the room rotation is independent of the body that is being judged. To test this, participants in Experiment 4 were asked to make left-right judgments about bodies only while we varied the relationship between the rooms and bodies separately. In the *body-rotate* condition, the room was maintained in the upright position in the background and the body was rotated. In the *room-rotate* condition, the body remained in the upright position and the room was rotated in the background (Fig. 9).

If the task-independent uprighting hypothesis is correct, then performance should be orientation independent when the room is upright and the body is rotating, just as in the conditions where the body is presented alone, whereas performance should be orientation dependent when the room is rotating even though the body about which the judgment is made remains in the upright position. By contrast, if the participants can ignore the task-irrelevant room rotation, then both conditions should produce patterns identical to the bodies alone in Experiment 1. Finally, our stimuli could introduce a third type of discrepancies by having the bodies and rooms in inconsistent orientations. If the "uprighting" tendency is sensitive to any type of incongruence, then both conditions might show orientation dependence as the participant attempts to reconcile the angular disparity between the room and body.
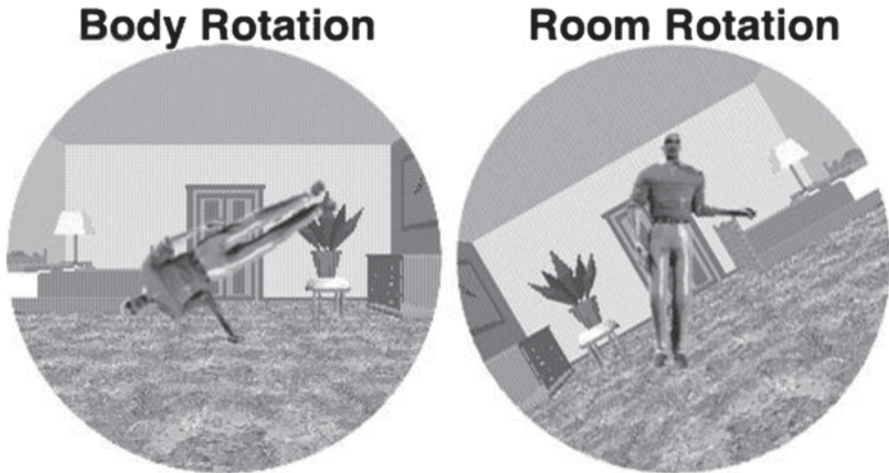
Fig. 9 Examples of the stimuli used in Experiment 4 showing a body rotating against a stable upright room and a stable upright body against a rotating room

Method

Participants

Twenty-four participants (12 male) from the Johns Hopkins Community volunteered in return for monetary compensation.

Materials

Using the same room images from Experiment 1 and bodies created as in Experiment 3, images were created that had the bodies in front of the doors of rooms as in Fig. 6. We used four base images that counterbalanced whether the extended hand of the body was on the same side as the plant in the room, even though participants were never asked about the plant (or any other feature of the room). From these four base images, two sets of stimuli were created. For the body rotation conditions, the room remained upright and the body in front of the door was rotated in the 12 different orientations ranging from 0° (upright) to 330° in 30° increments. For the room rotation condition, the body remained upright and the room in the background was rotated in the same 12 orientations. The 0° images for the two conditions were identical, so trials were randomly designated as belonging to one condition or the other to maintain independence of the two conditions.
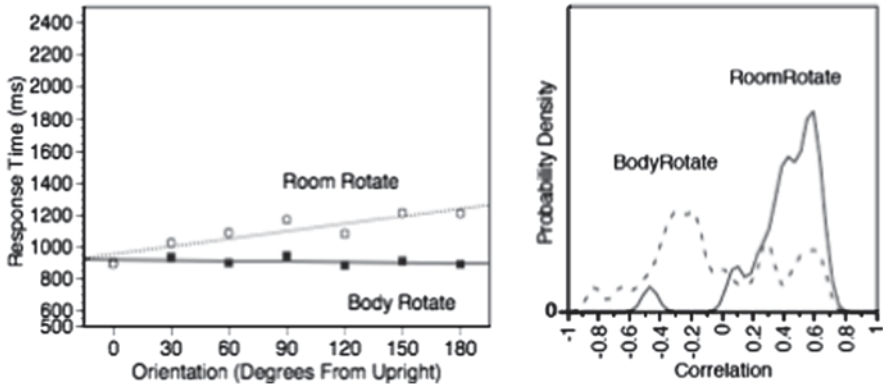
Fig. 10 Data from Experiment 4: Top panel shows response latency as a function of the orientation of either the body (closed squares) or the room (open circles) for the left-right task. Bottom panel shows the distribution of correlations between stimulus orientation and response time as function of which part of the stimulus was rotating (body or room)

Procedures

All participants performed left-right task on the bodies only using both sets of stimuli. Stimuli from the body rotation and room rotation conditions were presented in random order, and conditions were not explicitly revealed to the participants. Trial procedures were identical to the left-right task used in Experiment 1.

Results and Discussion

Error rates were low—3.5 and 2.4% for the body and room rotation, respectively. Response time data were trimmed and analyzed as described for Experiment 1. First, mean response times were calculated for each participant for each combination of condition and orientation, and these mean response times were submitted to a repeated measures ANOVA. As shown in the left panel of Fig. 10, there was a pronounced condition-by-orientation interaction, $F(6, 138)=6.87, p<0.001$, with response latency showing a stronger linear relationship with the room rotations than with the body rotations, $F(1, 23)=21.9, p<0.001$. Overall, responses were slower in the room rotation condition, $F(1, 23)=34.7, p<0.001$, and showed orientation dependence, $F(6, 138)=4.91, p=0.003$ [Linear contrast, $F(1, 138)=12.2, p=0.002$]. However, these effects were likely due to the interaction. The correlation between orientation and response time was greater for the room rotations than the body rotations, $t(23)=4.30, p<0.001$. Moreover, the average correlations were -0.04 and 0.41 for the body and room rotations, respectively, supporting the observation that the room rotations showed a substantial influence of orientation (Fig. 10, right panel).

These results support the hypothesis that rotated scene stimuli—even when they are task irrelevant—invoke some degree of automatic transformation to upright the world. In the body rotation condition, participants appeared to use perspective transformations; neither the rotation of the body relative to upright nor the discrepancy between the irrelevant room stimulus and the body affected response times. However, in the room rotation condition, the to-be-judged body stimulus was always upright with respect to the participant (and the computer screen, the testing room, etc.), but response times were affected by the rotation of the irrelevant room stimulus in the background, supporting the hypothesis that some task-irrelevant transformation is occurring in response to the presence of a rotated scene stimulus.

General Discussion

The four experiments reported here tested the degree to which scene stimuli (rooms) preferentially engaged perspective transformations more than object-based transformations. Previous research [10] has suggested that people tend to perform object-based transformations when making judgments about pictures of small, manipulable objects. The present results argue that people tend to perform perspective transformations when making judgments about pictures of scenes. This pattern is consistent with people's everyday experience of objects and places: Objects often move around us or are moved by us, and it is important to predict the consequences of those movements. Places, however, are generally stable. For places it is important to predict the consequences of occupying one location or another within the space. Bodies occupy a unique intermediate role: We experience them both as objects that can move around, when we watch other people, and as cues to potential locations of perspective, when we ourselves move around in the world. Consistent with this dual role, in these experiments and in previous studies, [2] and [10], when cued with a body, participants appeared to be able to flexibly perform either an object-based transformation or a perspective transformation, depending on the spatial judgment that needed to be made.

Experiments 1 and 3 provided evidence that spatial judgment response times depend on both the spatial judgment one is making and the thing about which that judgment is made. For the same-different task, there was a relationship between stimulus orientation and response, consistent with the performance of object-based transformations. However, this relationship was stronger for bodies than for rooms, consistent with the hypothesis that participants would be less inclined to use object-based transformations when reasoning about the room stimuli. A substantially weaker relationship was observed for both types of stimuli in the left-right task, supporting the use of perspective transformations, as expected.

In addition to the robust difference between rooms and bodies, there was a small but consistent effect of orientation on response latency for room stimuli in both tasks such that the response latency patterns for room stimuli were neither strongly linear (as expected for object-based transformations) nor orientation-independent.

Instead, for both the same-different and left-right tasks, we observed an attenuated trend for increased response latency as a function of angular disparity.

When presented with a picture of a room at an orientation that conflicts with other salient reference frames, participants may initially perform an object-based transformation of the picture to bring it into alignment with those other reference frames, independent of the spatial judgment task. Unlike pictures of bodies, pictures of rooms include salient straight lines and 90° intersections, establishing the planes of the walls. These features are strong cues to the reference frame of the picture. When room pictures are rotated, that reference frame conflicts with the reference frames defined by the participant's eye position, the computer screen, the room in which the experiment takes place, and gravity. The fact that response times for pictures of rooms increased less with orientation than response times for pictures of bodies, and were less affected by task manipulations than were response times to body pictures, argues for the view that participants tended to solve problems involving pictures of rooms by performing a perspective transformation to place themselves in the position depicted by the room. The results of Experiment 4 suggest that this uprighting need not be relevant and may not be requisite for the actual judgment but is an interference occurring in a more automatic fashion any time a rotated scene is presented. Recent studies provide additional evidence for the uprighting hypothesis by identifying the reference frame(s) used to define upright for scenes [26].

Together, these data provide clear evidence that performance in spatial reasoning tasks depends both on the type of spatial judgment required and on the stimulus about which the judgment is made. In particular, participants showed evidence of a tendency to use perspective transformations when reasoning about room stimuli, even for same-different judgments, which strongly evoke object-based transformations when made about pictures of bodies [2] and [10]. This interaction of task and stimulus provides compelling support for the view that multiple spatial transformation systems are tuned to be responsive to the requirements of different spatial reasoning situations. The adaptive deployment of these computational tools may form building blocks for complex skills such as navigation, long-term spatial memory, and abstract reasoning.

References

1. Zacks JM, Michelon P (2005) Transformations of visuospatial images. Behav Cogn Neurosci Rev 4:96–118
2. Zacks J, Mires J, Tversky B, Hazeltine E (2002) Mental spatial transformations of objects and perspectives. Spat Cogn Comput 2:315–332
3. Wraga MJ, Shephard JM, Church JA, Inati S, Kosslyn SM (2005) Imagined rotations of self versus objects: an fMRI study. Neuropsychologia 43:1351–1361
4. Zacks J, Rypma B, Gabrieli JDE, Tversky B, Glover GH (1999) Imagined transformations of bodies: an fMRI investigation. Neuropsychologia 37:1029–1040
5. Zacks JM, Ollinger JM, Sheridan MA, Tversky B (2002) A parametric study of mental transformations of bodies. NeuroImage 16:857–872

6. Zacks JM, Vettel JM, Michelon P (2003) Imagined viewer and object rotations dissociated with event-related fMRI. J Cogn Neurosci 15:1002–1018

7. Wraga MJ, Creem SH, Proffitt DR (2000) Updating scenes after object- and viewer-rotations. J Exp Psychol Learn Mem Cogn 26:151–168

8. Shepard RN, Metzler J (1971) Mental rotation of three-dimensional objects. Science 171:701–703

9. Parsons LM (1987) Imagined spatial transformation of one's body. J Exp Psychol Gen 116:172–191

10. Zacks JM, Tversky B (2005) Multiple systems for spatial imagery: transformations of objects and bodies. Spat Cogn Comput 5:271–306

11. Alivisatos B, Petrides M (1997) Functional activation of the human brain during mental rotation. Neuropsychologia 35:111–118

12. Barnes J, Howard RJ, Senior C, Brammer M, Bullmore ET, Simmons A et al (2000) Cortical activity during rotational and linear transformations. Neuropsychologia 38:1148–1156

13. Bonda E, Petrides M, Frey S, Evans A (1995) Neural correlates of mental transformations of the body-in-space. Proc Natl Acad Sci U S A 92:11180–11184

14. Carpenter PA, Just MA, Keller TA, Eddy W, Thulborn K (1999) Graded functional activation in the visuospatial system with the amount of task demand. J Cogn Neurosci 11:9–24

15. Cohen MS, Kosslyn SM, Breiter HC, DiGirolamo DJ, Thompson WL, Anderson AK et al (1996) Changes in cortical activity during mental rotation: a mapping study using functional MRI. Brain 119:89–100

16. Gauthier I, Hayward WG, Tarr MJ, Anderson AW, Skudlarski P, Gore JC (2002) BOLD Activity during mental rotation and viewpoint-dependent object recognition. Neuron 34:161–171

17. Harris IM, Egan GF, Sonkkila C, Tochon-Danguy HJ, Paxinos G, Watson JDG (2000) Selective right parietal lobe activation during mental rotation: a parametric PET study. Brain 123:65–73

18. Richter W, Somorjai R, Summers R, Jarmasz M, Menon RS, Gati JS et al (2000) Motor area activity during mental rotation studied by time-resolved single-trial fMRI. J Cogn Neurosci 12:310–320

19. Shelton AL, Pippitt HA (2006) Motion in the mind's eye: comparing mental and visual rotation. Cogn Affect Behav Neurosci 6:323–332

20. Tagaris GA (1998) Functional magnetic resonance imaging of mental rotation and memory scanning: a multidimensional scaling analysis of brain activation patterns. Brain Res Rev 26:106–112.

21. Vanrie J, Béatse E, Wagemans J, Sunaert S, Van Hecke P (2002) Mental rotation versus invariant features in object perception from different viewpoints: an fMRI study. Neuropsychologia 40:917–930

22. Zacks JM (2008) Neuroimaging studies of mental rotation: A meta-analysis and review. J Cogn Neurosci 20:1–19

23. Aguirre GK, D'Esposito M (1997) Environmental knowledge is subserved by separable dorsal/ventral neural areas. J Cogn Neurosci 17:2512–2518

24. Bonda E, Petrides M, Ostry D, Evans A (1996) Specific involvement of human parietal systems and the amygdala in the perception of biological motion. J Cogn Neurosci 16:3737–3744

25. Cohen JD, MacWhinney B, Flatt M, Provost J (1993) PsyScope: a new graphic interface environment for designing psychological experiments. Behav Res Methods, Instrum Comput 25

26. Shelton A, Lau Y, Zacks J, Yoon BC (2008) The opportunistic use of reference frames for rotating scene stimuli. J Vis 8:739