Published in final edited form as:
Q J Exp Psychol (Hove). 2016 June ; 69(6): 1210–1226. doi:10.1080/17470218.2015.1079225.

How Are Bodies Special? Effects Of Body Features On Spatial Reasoning

Alfred B. Yu[1,2] and **Jeffrey M. Zacks**[2]

[1]Army Research Laboratory, Aberdeen Proving Ground, MD

[2]Washington University in St. Louis

Abstract

Embodied views of cognition argue that cognitive processes are influenced by bodily experience. This implies that when people make spatial judgments about human bodies, they bring to bear embodied knowledge that affects spatial reasoning performance. Here, we examined the specific contribution to spatial reasoning of visual features associated with the human body. We used two different tasks to elicit distinct visuospatial transformations: *object-based transformations*, as elicited in typical mental rotation tasks, and *perspective transformations*, used in tasks in which people deliberately adopt the egocentric perspective of another person. Body features facilitated performance in both tasks. This result suggests that observers are particularly sensitive to the presence of a human head and body, and that these features allow observers to quickly recognize and encode the spatial configuration of a figure. Contrary to prior reports, this facilitation was not related to the transformation component of task performance. These results suggest that body features facilitate task components other than spatial transformation, including the encoding of stimulus orientation.

Keywords

perspective taking; spatial transformations; mental imagery

In order to engage successfully in social interactions, it is often important to consider aspects of the spatial configuration of another person's body, such as head orientation, gaze direction, and posture. For example, consider a yoga instructor who is teaching several students in a class. In some situations, it may be necessary for the instructor to look at the students and determine whether every student is adopting the same pose. In other situations, it may be necessary to determine whether a student has raised their left or right hand. The instructor also may wish to compare a student's pose to a verbal description or a stick figure drawing. Are these judgments *embodied*? That is, does performance depend on experiential knowledge of the structure and posture of one's own body? Embodied experience could contribute to spatial reasoning in two distinct ways. First, it could facilitate the

Address correspondence to: Jeffrey M. Zacks, Department of Psychology, Washington University, 1 Brookings Drive, Saint Louis, MO 63130, USA.

Supplemental Material
The Supplemental Material can be found at the address

Author Manuscript
Author Manuscript
Author Manuscript

transformation of a body's spatial configuration. Experience with the visual, kinesthetic, and vestibular consequences of one's own motion could facilitate computing transformations of a body's spatial reference frame. Second, embodied experience could facilitate non-transformation components of spatial reasoning such as the encoding of a body's spatial configuration—its orientation and pose. In this paper, we describe two experiments in which people made spatial judgments about pictures of objects. We asked whether perceiving an object as a body affected how they performed the spatial transformation components of the tasks, the non-transformation task components, or both.

Stimulus transformations and spatial reference frames

Mental spatial transformations can be characterized in terms of the spatial reference frames on which they operate (Zacks & Michelon, 2005). Spatial reference frames can be divided into three broad classes (McCloskey, 2001). An environmental reference frame is defined with respect to part of the external world, such as a classroom, a campus, or a city. For example, "the lamp is near the door" and "the bar is on the north side of town" each locate something with respect to an environmental reference frame. An egocentric reference frame is defined with respect to the vantage point of an observer. For example, "the bowl of ice cream is in front of me" locates the ice cream with respect to an egocentric reference frame. Finally, an object-based reference frame is defined with respect to an object. For example, "the steering wheel is on the left side of the car" and "the eraser is on the top of the pencil" locate things with respect to an object-centered reference frame. Using these classes of reference frames, imagined transformations of space can considered to be composed from three basic classes, each consisting of a transformation of a single class of reference frame while holding constant the other two. An object-based transformation is an imagined relocation of an object-based reference frame, such as imagining the object rotating so that it goes from upright to upside-down relative to the environmental reference frame. A perspective transformation is an imagined relocation of the observer's egocentric reference frame so that it occupies a different vantage point, such as imagining one's self viewing a room from the point of view of another person in the room. (Transformations in which an environmental reference frame is selectively updated are infrequently encountered in the real world, and mentally simulating them does not appear to play a significant role in spatial reasoning.)

Object-based transformations are sensitive to orientation disparity. In a classic study, Shepard and Metzler (1971) presented pairs of abstract block shapes and asked participants to verify whether the shapes were the same, or whether one was a mirror-image version of the other. In this task, orientation disparity is defined as the difference in orientation between the two shapes. They found that response time increased with increasing orientation disparity; this was true both for rotation in the picture plane and in the depth plane. This pattern was taken to indicate that participants performed an object-based transformation to mentally align the two shapes. In tasks involving only one stimulus, orientation disparity can be defined relative to a canonical orientation such as upright. For example, Cooper and Shepard (1973) showed participants alphanumeric characters at various orientations and asked them to judge whether they were printed normally or mirror-reversed. In this case, orientation disparity was defined relative to the canonical upright orientation, and characters were rotated in the picture plane only. As with the block figures of Shepard and Metzler (1971), response time increased with increasing orientation disparity.

Perspective transformations are also orientation-sensitive in some situations, showing slower responses when a target is misoriented to a greater degree (Huttenlocher & Presson, 1973; Michelon & Zacks, 2006; Presson & Montello, 1994). However, this is not always the case. Parsons (1987) asked people to imagine themselves in the position of a depicted figure and to press a key when they were done. He found that the effect of orientation on response time depended on the axis of rotation of the figure. For example, when upright figures were rotated around the vertical axis response time increased monotonically with rotation. However, when front-facing figures were rotated in the picture plane there was essentially no effect of rotation on response time. Parsons offered a potential explanation of these effects in terms of embodiment: He suggested that perspective transformations are influenced by the time it would take to perform an action that would place one's viewpoint in the target orientation.

In the present studies, we took advantage of the observation that for front-facing figures rotated in the picture plane, perspective transformations produce orientation-independent performance whereas object-based transformations produce orientation-dependent performance (Yu & Zacks, 2010; Zacks, Mires, Tversky, & Hazeltine, 2002; Zacks & Tversky, 2005). This means that, for this configuration, orientation dependence can be used to diagnose whether a perspective transformation is being performed.

Different spatial transformations for different task situations

Previous research suggests that whether people use an object-based transformation or a perspective transformation to solve a spatial reasoning problem depends on the judgment they are asked to make and on the stimuli they see. Zacks and Tversky (2005) asked participants to make two sorts of judgments about rotated bodies and manipulable objects (see Figure 1). In the *same-different* task, participants viewed two pictures, one of which was upright and the other of which was rotated by varying amounts. The two objects were either identical or mirror images, and the participants were asked whether the objects were identical. In the *left-right* task, participants viewed only the rotated object and judged its handedness. They found evidence that viewers overwhelmingly used object-based transformations to perform the same-different task. For the left-right task, which transformation was used depended on the object depicted: For pictures of bodies, participants tended to perform perspective transformations; for pictures of small manipulable objects, they tended to perform object-based transformations.

Ways that spatial reasoning tasks may be embodied

One possibility is that people adopt systematically different spatial transformation strategies when dealing with human bodies and inanimate objects. In their day-to-day interactions with small objects, people often manually manipulate them or observe others doing so. As a result, they frequently observe changes that correspond to object-based transformations. For example, if one turns a door handle one observes an object-based transformation as the handle rotates. Typical interactions with people are quite different. One frequently needs to adopt another person's perspective, as when giving directions or coordinating behavior such as moving around a crowded kitchen without bumping into other cooks. Such considerations have led some to propose that spatial reasoning about human figures is *embodied*—that is, that people drawn on first-person experience of their own body when reasoning about the bodies of others. Sayeki (1997) argued that a "body analogy" can be used to reason about objects, by mapping structural components of an object onto corresponding components of a mental body schema. Sayeki found that performance in a Shepard-Metzler mental rotation task was facilitated if participants were exposed to the strategy of imagining a human head placed on top of the block figures. Facilitation also was found when the top block was labeled with an arrow pointing downwards, suggesting that the benefit of the human head was to provide a visually salient anchor for encoding and representing the orientation of the object.

Amorim, Isableu, and Jarraya (2006) extended Sayeki's findings by manipulating *body-likeness*. They arranged arranging non-human objects in poses that were globally analogous to a human body in similar poses. They contrasted spatial transformation performance for stimuli with human bodies, lamps with movable joints, and abstract objects built from Shepard-Metzler cubes. These researchers drew a distinction between *spatial embodiment* and *motoric embodiment*. In spatial embodiment, knowledge of the spatial configuration of the human body facilitates performance by providing an analogous structure for representing non-human shapes. In motoric embodiment, observation of another's body in a particular pose leads to motor resonance, that is, activation of the neural and mental representations that would have been active if one's own body were adopting that pose. Amorim et al. proposed that both spatial embodiment and motoric embodiment play a role in shape-matching performance. Spatial embodiment allows the observer to construct a representation of another person's axes (e.g. head-feet, left-right, front-back) by using an analogy to one's own body. Motoric embodiment enables the observer to understand the posture of the other person's body via a covert simulation of their posture. In their view, the facilitation associated with body features may occur via both spatial embodiment and motoric embodiment, and indeed they found that this facilitation depended on both the degree to which the orientation of the stimulus matched that of the observer and whether the figure's pose was kinematically possible. Consistent with this proposal, they found that rotated human and human-like figures were less affected by orientation disparity than non-human figures. They interpreted this as evidence for embodied spatial transformations, meaning facilitation of spatial transformation performance by knowledge of one's own body.

Two open questions

The previous results suggest that reasoning about the human body is "special," but leave in doubt at least two important claims about the way it is special. First, does embodiment make the transformation component itself more efficient, or does the faster performance result from non-transformation processes such as stimulus encoding or responding? Second, do people reason differently about bodies because bodies have a familiar visual appearance that is linked to one's experience of one's own body, or because bodies give unique cues to object orientation that allow one to efficiently establish relevant frames of reference. Many objects other than bodies have only zero or one salient axes of asymmetry—think of apples, balls, trees, dining tables, or tissue boxes. The bodies of humans and other animals are relatively unusual in that they have clear up-down and front-back axes. In the experiments of Sayeki (1997) and Amorim and colleagues (2006), the stimuli described as more embodied also had stronger orientation cues. To what degree does the fact that a picture depicts an animate embodiable thing affect spatial reasoning above and beyond the benefits provide by such orientation cues?

We previously reported that the degree to which a picture shows an animate creature affects spatial reasoning performance (Yu & Zacks, 2010). We used the same-different and left-right tasks described previously. Animacy was manipulated by using stimuli depicting human bodies, animals, and inanimate objects. We assessed whether participants used perspective transformations or object-based transformations by testing whether their performance was orientation dependent. We found that for left-right judgments, people were more likely to use perspective transformations when the stimuli depicted animate beings than inanimate objects. (For same-different judgments, people were overwhelmingly likely to use object-based transformations across all stimulus types.) In addition, responses to animate stimuli were faster than responses to inanimate objects. The results suggested that animacy can influence *which* spatial transformation is preferred in some spatial reasoning tasks, and also that animacy can affect how efficient task performance is overall. However, they do not answer the question whether judgments about bodies make any particular spatial transformation more efficient.

In sum, the claims that bodies are special or that spatial reasoning is embodied can be made more precise by specifying the cognitive mechanisms that are affected by representations of the body. In these studies, we focused on the effects of body-related features on transformation-specific and non-transformation specific components of simple spatial reasoning tasks. We directly manipulated the degree to which our stimulus figures resembled a human body, and performed analyses designed specifically to assess effects on transformation-specific and non-transformation components of spatial reasoning.

Experiment 1

Experiment 1 tested three hypotheses. First, do body features facilitate spatial reasoning performance? In particular, does the head alone facilitate? Second, if body features affect performance, do they do so for the transformation-specific or non-transformation components of the task? Finally, are benefits of body cues specific to tasks involving perspective transformations or to object-based transformations, or do they manifest in tasks eliciting both transformations?

To assess the effect of body features on spatial reasoning performance, we created three different figure types: intact human *bodies*, block objects similar to those used by Shepard and Metzler (1971), but with a human head placed at the top (*block-head* figures), and the same objects with a teapot placed at the top (*block-teapot* figures). As shown in Figure 2, all three stimulus types have clear up-down and front-back axes. The body stimuli have the most familiar body features, and the block-teapot stimuli have the fewest.

The presence of a head and face is a powerful cue to the structure of the human body. We included the block-head figures because they provide this cue while preserving much of the geometric structure of the abstract block figures used by Shepard and Metzler (1971). The head provides familiar visual features that could facilitate rapid visual identification of an object's top and front through the viewer's history of associations between visual features and experience of the body senses: kinesthetic, proprioceptive, and vestibular features. The head also provides intrinsic geometric cues to define a top and front. Most important, it defines a top in the object-centered reference frame. Rock (1973, pg. 23) pointed out that familiar shapes such as bodies have intrinsically-defined orientations even when they are misoriented with respect to the environment and observer. The block-teapot stimuli control lack the associations with body sense features but have equivalent geometric cues: Both heads and teapots include a clear intrinsic top (the hair on the head or the lid of the teapot) and front (the face on the head or the spout of the teapot), and both also display horizontal symmetry about their vertical axes. Therefore, the primary difference between the stimulus sets was the relative "embodiability" of the former. Note that embodiability is likely confounded with familiarity. This likely mirrors human experience: There are probably not any highly-familiar inanimate objects with a similar size and top-bottom, front-back, and left-right organization as the human head.

Body features could have three potential effects on performance. First, adding visual cues provided by bodies could facilitate the perception of objects' orientation. This should reduce overall response time in both the same-different and left-right tasks. Second, adding body cues could make object-based transformations or perspective transformations more efficient, as suggested by Amorim, Isableu, and Jarraya (2006). This should reduce the slope of the function relating orientation to response time under conditions where there is a positive slope. Recall that positive slopes are ubiquitous for the object-based transformations, but are not found for perspective transformations of frontal pictures rotated around the depth axis. Third, adding body cues could encourage more frequent use of perspective transformations, as observed previously for the left-right task (Yu & Zacks, 2010; Zacks and Tversky, 2005). This would be expected to reduce the correlation between orientation and response time (and also the slope).

Method

Participants—Thirty-two participants from the Washington University Psychology Department's participant pool took part for course credit. Data from one participant were discarded due to a high error rate (48%). We analyzed data from 31 participants (8 female, mean age 20.2 years).

Apparatus—Stimuli were presented using PsyScope X (http://psy.ck.sissa.it) on a 19-inch CRT monitor at a distance of approximately 45 cm from the participant's eyes. The resolution of the display was set to 1920 by 1440 pixels. Participants responded by pressing buttons on a USB button box (ioLab Systems).

Stimuli—The stimulus set consisted of color images of constructed in Poser 7 (Smith Micro, Aliso Viejo, CA). A human male figure was posed in 12 different ways by varying the position of his shoulders, elbows, waist, knees, and ankles. For each of the 12 poses, a corresponding block figure was constructed to approximate the figure's pose in 3-dimensional space. Each block figure consisted of a marker (either a human head or a teapot) placed at the top of 10 adjacent and identical cubes spaced evenly within a regular grid such that the faces of adjacent cubes were completely touching. The widest figure was 4 cubes wide, the tallest figure was 6 cubes tall, and the deepest figure was 3 cubes deep. Figures were marked on one side (left or right) by superimposing a multi-colored dot onto the picture. These dots were superimposed onto identical locations across corresponding human body, block-head, and block-teapot figures. Each individual figure spanned approximately 8 degrees of visual angle and was presented on a light grey background. Rotated versions of each unique stimulus were created by rotating the original picture in the picture plane in 30 degree increments. Horizontally mirrored versions of each rotated stimulus were then created by flipping each image along the vertical axis. Figure 2 shows a body, a block-head figure, and a block-teapot figure in the same pose.

Design and procedure—Experiment 1 consisted of a total of six blocks, each consisting of 96 trials of either the same-different task or the left-right task. Task alternated by block, and stimulus type switched every two blocks. Both figure type order and task order were counterbalanced across participants.

At the beginning of each trial, a fixation cross was presented for 500 ms, and then replaced with a pseudo-randomly selected stimulus item. In the left-right task, participants judged whether the left or right side (relative to the figure's egocentric perspective) of a centrally-presented stimulus figure was marked with a dot. Participants responded by pressing the left button on the button box if the mark was on the figure's left, and the right button if the mark was on the right. In the same-different task, participants viewed two similar figures: an upright reference figure presented in the top half of the screen, and a rotated comparison figure presented in the bottom half. Participants judged whether the comparison figure was the same as the reference figure, or whether it was a mirror-flipped version. Each stimulus item remained on the screen until the response. Feedback was given at the end of the trial, in the form of an auditorily-presented beep for correct trials and a buzzer for incorrect trials. At the end of all six blocks, participants completed a strategy questionnaire with separate questions for each of the blocks (the questionnaires can be found in the online supplement).

Results

RT analyses—We trimmed RT data by excluding RTs from error trials, trials with responses faster than 300 ms, and trials with responses slower than the mean plus 3 standard deviations for each subject. This procedure excluded a total of 6.1% of all trials. Per-subject mean RTs were calculated for each combination of task, stimulus type, and orientation by taking the mean of the trimmed RTs. We collapsed data from trials with clockwise and counterclockwise rotations from upright, resulting in seven possible orientation values (0, 30, 60, 90, 120, 150, and 180 degrees). Figure 3 depicts the mean RTs for all experimental conditions and supplementary Figure S1 depicts the mean error rate for the same conditions. Trimmed RTs were submitted to a 3-way (task * stimulus type * orientation) ANOVA. All main effects were significant (task: $F(1, 30) = 52.1$, $p < .001$; stimulus type: $F(2, 60) = 7.1$,

$p < .01$; orientation: $F(6, 180) = 56.5$, $p < .001$). The task by orientation interaction was significant ($F(6, 180) = 29.0$, $p < .001$), showing that the cost associated with orientation disparity differed between the two tasks. The task by stimulus type interaction ($F(2, 60) = 3.2$, $p < .05$) was also significant and appeared to be driven by a greater overall benefit of seeing human body stimuli in the same-different task, relative to the left-right task where mean RTs for human bodies and block-teapot figures were more similar. Follow-up *t* tests showed that the difference in mean RT between body and block-teapot figures was significant in the same-different task ($t(30) = 3.0$, $p < .01$) but only marginally so in the left-right task ($t(30) = 2.0$, $p = .056$). The stimulus type by orientation interaction was not significant ($F(12, 360) = 0.7$, $p = .707$). Importantly, the three-way interaction between task, stimulus type and orientation was also not significant ($F(12, 360) = 1.0$, $p = .472$), nor was the stimulus type by orientation interaction significant in follow-up ANOVAs for each of the individual tasks (left-right task: $F(12, 360) = 0.9$, $p = .539$; same-different task: $F(12, 360) = 0.8$, $p = .601$). The consistent lack of interactions involving stimulus type and orientation suggests that performance differences in processing the three stimulus types are independent of orientation; in other words, the benefit associated with viewing the human body relative to the block figures applies equally to all orientations. We examined potential carryover effects by adding stimulus type order and task order variables to the previously-mentioned ANOVAs. The only significant effect was an stimulus type by stimulus type order interaction, indicating that participants responded more quickly in later blocks ($F(10, 50) = 7.9$, $p < .001$).

Response time component analyses—We examined three different components of each participant's RT-orientation profile: the slope, the intercept, and the correlation between RT and orientation. We fitted a first-order linear regression model to predict RT based on orientation disparity for each unique combination of type, task, and participant. Each linear model yielded both a slope and an intercept. The slope represents the millisecond increment in RT associated with a one-degree increment in orientation, whereas the intercept represents the model's estimate of the RT at zero degrees of orientation disparity. The intercept is a good estimator of the components of task performance other than the mental spatial transformation, including stimulus encoding and responding. (We will set aside responding in discussing intercept effects, because there is not a good theoretical rationale for hypothesizing that body features affect the decision or response stage in these tasks.) Steeper slopes indicate that performance (a) depends on orientation, and (b) is inefficient. Flatter slopes show that performance is either orientation-dependent and efficient, or not orientation-dependent. The correlation between RT and orientation indexes the orientation-dependence of a participant's performance, separate from efficiency. We used separate ANOVAs to examine the influence of task and stimulus type on each of these three DVs. Results for each of the three summary statistics are plotted in Figure 4.

Intercepts of the RT-orientation function were lowest for bodies and highest for block-teapots, resulting in a main effect of stimulus type ($F(2, 60) = 9.0$, $p < .001$). In the left-right task, bodies differed from block-head figures ($t(30) = 2.1$, $p < .05$) and block-teapot figures ($t(30) = 3.0$, $p < .01$), whereas the block-head to block-teapot comparison was not significant ($t(30) = 1.7$, $p = .092$). In the same-different task, the body to block-teapot comparison was significant ($t(30) = 2.6$, $p < .05$), whereas the body to block-head comparison ($t(30) = 1.9$, $p = .070$) and the block-head to block-teapot comparison ($t(30) = 2.0$, $p = .059$) were marginally significant. The numerical trend was in the expected direction; that is, increasingly body-like stimuli were associated with lower intercepts across both tasks. Intercepts were lower for the left-right task than the same-different task, resulting in a main effect of task ($F(1, 30) = 28.3$, $p < .001$). The interaction between task and stimulus type was not significant ($F(2, 60) = 0.9$, $p = .398$).

The slope of the RT-orientation function was higher for the same-different task than the left-right task, leading to a main effect of task ($F(1, 30) = 51.2$, $p < .001$). There was no main effect of stimulus type on slope ($F(2, 60) = 2.0$, $p = .149$), although there was a significant interaction between task and stimulus type ($F(2, 60) = 3.8$, $p < .05$). Responses to human bodies were associated with the greatest slopes in the left-right task, but the smallest slopes in the same-different task. Follow-up *t* tests within each task showed that slopes generally did not differ between stimulus types, with the exception of one comparison in the left-right task: the slopes for bodies were greater than those for block-teapot figures in the left-right task ($t(30) = 2.7$, $p < .05$).

Correlations between orientation and response time were positive for the same-different task and close to zero for the left-right task, replicating previous results with bodies (Yu & Zacks, 2010; Zacks & Tversky, 2005). This led to a significant main effect of task ($F(1, 30) = 47.7$, $p < .001$). Correlations were highest for bodies and lowest for block-teapots, leading to a main effect of stimulus type ($F(2, 60) = 5.0$, $p < .01$). Additionally, the interaction was significant ($F(2, 60) = 3.3$, $p = .042$). In the left-right task, both the body to block-head comparison ($t(30) = 2.2$, $p < .05$) and the body to block-teapot comparison ($t(30) = 2.7$, $p < .05$) were significant, but the block-head to block-teapot comparison was not ($t(30) = 0.9$, $p = .395$). In the same-different task, only the block-head to block-teapot comparison was significant ($t(30) = 2.4$, $p < .05$). The body to block-head comparison was not significant ($t(30) = 1.5$, $p = .137$), nor was the body to block-teapot comparison ($t(30) = 0.6$, $p = .544$).

Discussion

The data provided clear support for the hypothesis that body features facilitate spatial reasoning performance. They suggest that the presence of a head alone aids performance, though the stimuli with heads alone were not always significantly better than stimuli with teapots rather than heads. The data failed to support the hypothesis that body features aid the transformation-specific components of task performance, as seen in the analysis of RT slopes and correlations. Finally, the benefits of body features appeared to accrue to both the left-right and same-different tasks, though the statistical evidence was stronger for the same-different task.

These results stand in contrast to those reported by Sayeki (1997) and Amorim, Isableu, and Jarraya (2006). The paradigm in those experiments was very similar to our same-different task. We did not replicate their primary finding that body features alter the cost associated with orientation in the same-different task. In the RT analyses, orientation did not interact with stimulus type for either task. In the slope and correlation analyses, effects of stimulus type were modest, and when they occurred stimuli with body features showed *more* rather than less of a rotation cost. Instead, we found that body features reduced RT intercepts, which we interpret as differences in non-orientation-related mechanisms such as stimulus encoding. Importantly, our results do not support Amorim et al.'s hypothesis that participants use orientation-sensitive motoric embodiment processes to make same-different judgments about rotated figures.

Why were stimuli with more body features slightly more orientation-dependent? One possibility is that viewers are *less* likely to engage in motor embodiment with body stimuli. We find this counterintuitive; moreover, the fact that RT-orientation slopes were not highest for body stimuli in the same-different task argues against this interpretation. We believe a more plausible explanation is that trial-to-trial performance was more consistent with the body stimuli and more variable with the block stimuli, leading to higher and lower correlations, respectively. This could arise if the block stimuli produce more frequent errors in the assignment of axes during encoding.

Experiment 2

In Experiment 2 we attempted to replicate the finding that body features improved the non-transformation components of task performance, using a paradigm that more closely replicated that of Amorim, Isableu, and Jarraya (2006). We also tested the hypothesis that the geometric cues provided by the head are by themselves facilitatory. All of the figures in Experiment 1 included a distinctive feature at the intrinsic top of the object. This may have facilitated assigning an intrinsic up-down axis and the polarity of that axis (which end is "up"), which could speed orientation coding (Gregory & McCloskey, 2010). To test whether this was the case, we added a fourth figure type that did not have the geometric cues provided by the teapots in the block-teapot objects. These *block-only* figures had neither a head nor a teapot to individuate the object's top and front. To the extent that body features improve performance by facilitating the assignment of axes during encoding, RT intercepts should be slower for these stimuli than for block-teapot stimuli.

In Experiment 1, each figure was marked on one side with a colored dot. This was necessary for the left-right task, and we held this constant for the same-different task. However, this may have affected performance in the same-different task, because the dot provided a consistent target for judging laterality; in theory, participants could have performed the task by attending only to the head (or teapot) and the dot, ignoring the object's shape. In Experiment 1 a small number of subjects (4 out of 30) reported that they used this strategy. To address the possibility that the markers affected performance in the same-different task, in Experiment 2 we used eliminated the marker dots.

Following Amorim, Isableu, and Jarraya (2006), we included only the same-different task in Experiment 2.

Method

The overall procedure was similar to that used in Experiment 1. The major changes were the elimination of the left-right task, the removal of the dot markers, and the addition of block-only figures.

Apparatus—The display used in Experiment 1 was no longer available and a different display was used in Experiment 2. It was set to a lower resolution (1280 × 1024 pixels) but the size of the stimuli were adjusted to span the same visual angle as in Experiment 1.

Participants—Thirty-five participants from the Washington University Psychology Department's participant pool took part for course credit. Data from five participants were discarded for having error rates above 15% (ranging from 18% to 26%). We analyzed data from 30 participants (23 female, mean age 19.6 years).

Stimuli—Figures from Exp. 1 were used with the following changes. The colored dot was removed because the paradigm with only the same-different task did not require stimuli to be clearly marked on one side or the other. The clothes worn by the human figure were changed to white in order to remove the color gradient that could provide information about the figure’s orientation (previously, the figure’s shirt was orange and its pants were blue). Small changes were made to the figure’s pose in order to better correspond with the blocks. Finally, a fourth, block-only stimulus type was introduced by removing the head or teapot from the cube figures. Rotated versions of these block-only figures had no obvious cues to their uprightness, aside from minor lighting and perspective cues that were not explicitly pointed out to participants.

Task—The same-different task was performed in the same way as the same-different task blocks in Experiment 1. Participants completed four blocks of the same-different task, with block order counterbalanced across participants.

Results

Individual trials were trimmed as described for Experiment 1 (8.0% of trials were excluded). Figure 5 depicts the mean RTs for all experimental conditions and supplementary Figure S2 depicts mean error rates. Trimmed RTs were submitted to n ANOVA with stimulus type and orientation as within-subjects variables. Response time increased with stimulus orientation, resulting in a significant main effect of orientation ($F(6, 174) = 92.2$, $p < .001$). Responses were fastest for bodies, slower for block-head figures, and slower yet for block-teapot figures, replicating the pattern observed in Experiment 1. Responses were slower still for block-only figures. These differences resulted in a main effect of stimulus type ($F(3, 87) = 22.2$, $p < .001$). Follow-up t-tests showed that all pairwise differences were significant (smallest ($t(29) = 2.34$, $p = .03$) except for the comparison between the body and block-heads condition ($t(29) = 1.33$, $p = .19$). As in Experiment 1, the interaction between orientation and stimulus type was not significant ($F(18, 522) = 1.3$, $p = .168$). Once again, we checked for carryover effects by adding stimulus type order to the previous ANOVA. In similar fashion as Experiment 1, the only significant effect was an stimulus type by stimulus type order interaction, indicating that participants responded more quickly in later blocks ($F(3, 9) = 5.1$, $p < .001$).[1]

Intercepts, slopes and correlation coefficients were calculated for each participant using the same method as described for Experiment 1. These three dependent measures were then each examined using an ANOVA with stimulus type as a within-subjects variable. Slopes did not differ across stimulus types ($F(3, 87) = 1.0$, $p = .395$), whereas intercepts ($F(3, 87) = 21.1$, $p < .001$) and correlation coefficients ($F(3, 87) = 4.5$, $p < .01$) did. Follow-up *t* tests on intercepts revealed that only the body to block-head comparison failed to reach significance ($t(29) = 1.7$, $p = .105$; all other $p < .05$). Follow-up *t*-tests on correlation coefficients showed significant differences for the body to block-only comparison ($t(29) = 3.1$, $p < .01$) and the block-head to block-only comparison ($t(29) = 3.4$, $p < .01$). Results for each of the three summary statistics are plotted in Figure 6.

Discussion

The results of Experiment 2 replicated and extended those of Experiment 1: Body features reduced the intercept of the function relating orientation to response time but not the slope or the correlation between orientation and response time. The data also supported the hypothesis that geometric features are sufficient to facilitate performance but that body-specific features add additional value: RTs for block-teapot stimuli were faster than for block-only stimuli, but block-head stimuli were faster still.

These data point to a very consistent effect of body features on spatial reasoning: overall performance is better with increased similarity to the human body, but this effect is independent of orientation. This suggests that body features facilitate the non-transformation components of task performance. As in Experiment 1, the differences between bodies and block-head stimuli were small (though in Experiment 1 the intercept for body stimuli was slightly but significantly faster than for block-heads.) This suggests that much of the benefit of reasoning about a body can result from viewing just the head.

Similar to Experiment 1, bodies and block-head figures were associated with *higher* correlation values than block-only figures. Again, this is inconsistent with the proposal of Amorim et al. that body stimuli encourage motoric embodiment, which reduces the dependence of RT on orientation.

General Discussion

How are bodies special? In two visuospatial tasks, we found evidence that observers process human bodies more quickly than blocks arranged in similar spatial configurations. This body-related facilitation was independent of the orientation of the figures, suggesting that knowledge about one's own body enables faster recognition and encoding of other human bodies, rather than speeding up transformation of the resulting spatial representations.

[1]Overall, participants responded more slowly in this study than in Experiment 1. This could have come about as a result of deleting the colored dot from the stimuli, or perhaps because only the same-different task was administered. However, we also noted the proportion of male and female participants differed between Experiments 1 and 2. To test whether the difference in gender representation accounted for the overall RT difference, we performed an ANOVA on RT with gender and experiment number as independent variables. Both main effects were significant (sex: $F(1, 57) = 26$, $p < .001$; experiment: $F(1, 57) = 35$, $p < .001$), but the interaction was not ($F(1, 57) = 2.3$, $p = .134$). In a similar ANOVA with accuracy as the dependent variable, the main effect of sex was significant ($F(1, 57) = 13$, $p = .001$) while the main effect of experiment was not ($F(1, 57) = 1.0$, $p = .329$). The interaction was not significant ($F(1, 57) = 0.1$, $p = .742$). Thus, the different gender composition does not fully account for the slower performance in Experiment 2.

Embodiment is a term encompassing a wide variety of body-related influences on cognition. In the present context, we have distinguished between three cognitive shifts that could make up “embodiment:” improving recognition of orientation cues, increasing transformation efficiency, and encouraging the use of perspective transformations. Other studies have aimed at parsing the mechanisms of perspective transformations more finely: Zacks and Tversky (2005) distinguished between the mechanism of updating a head-centered egocentric reference frame and performing a motor simulation that updates other egocentric reference frames such as those of the hands or feet. Kessler and Rutherford (2010) proposed that left-right judgments about objects arrayed in depth on a picture of a table were likely to elicit motor simulations of the perspective-taking target’s posture whereas visibility judgments were not. Using a similar paradigm, Michelon and Zacks (2006) found that visibility judgments did not elicit perspective transformations. Could increased use of motor simulations account for the present effects of body features on judgment times? This is not likely; increased use of motor simulations would be expected to change the slope of the orientation-response time function, not just its intercept.

Another way that body features could modulate processing is by increasing the sensorimotor interference between the side of the response and the side marked on the stimulus figure (May & Wendt, 2012). For example, when viewing an upright figure that is facing the viewer, the figure’s left hand is on the viewer’s right and vice versa. In the present studies, could body features have increased the amount of interference with one’s own egocentric reference frame? Again, this is not likely. If so, this should have differentially affected upright figures because for upright figures viewer’s left-right conflicts with the figure’s left-right, whereas for inverted figures the viewer’s left-right is the same as the figure’s left-right.

These results bear directly on the distinction of Amorim, Isableu and Jarraya (2006) between spatial and motor embodiment. We conclude that our manipulation of body features did not result in a different (and differently embodied) transformation process; instead it appears that the reduced RTs to animate figures derive from the familiarity of the human body and especially the head. We believe that the source of our effect corresponds more closely to the process of spatial embodiment put forth by Amorim et al., which involves accessing and extending structural information about one’s own body in order to reason about another being.

In all cases, we found that that presenting a human head on top of a block figure resulted in performance that was intermediate between performance with bodies and with objects that did not include heads. Sayeki (1997) exposed participants to abnormally-positioned heads, placed so as to disrupt the analogy between the block figure and a participant’s body. He found that this manipulation greatly increased the RT-orientation slope for picture plane rotations relative to the “normal” block-head condition. He also found that the RT intercept was lower compared to a condition where the head was replaced with an arrow. These findings are consistent with results from studies of the *body inversion effect*, which is the fact that matching and identification judgments about bodies are often gravely disrupted by inverting the stimuli (Reed, Stone, Bozova, & Tanaka, 2003). Removing the head from body stimuli reduces the body inversion effect in some situations (Yovel, Pelc, & Lubetzky, 2010), though not always (Robbins & Coltheart, 2012). Functional MRI data suggest that some of the processing that is specialized for upright bodies draws on face-selective areas of the ventral temporal cortex (Brandman & Yovel, 2010). Thus, part of the spatial transformation advantage for bodies may arise from specialized processing of the head in its normal structural relationship to the rest of the body.

Apropos of the inversion effect, it is worth noting the differences between studies of the inversion effect and the spatial transformation tasks used here. The inversion effect is conceived as a visual recognition phenomenon; accordingly, researchers ask participants to make matching judgments about stimuli that are either both upright or both inverted, or to name stimuli that are upright or inverted (Carey & Diamond, 1994; Reed et al., 2003; Robbins & Coltheart, 2012). For both of these tasks, responses are usually slower and less accurate for inverted stimuli, and the effects are particularly large for faces and bodies. In contrast, the tasks used here asked participants make matching judgments about pairs of stimuli at *different* orientations (the same-different task) or to make a judgment about the spatial configuration of a stimulus without reference to its identity (the left-right task). Whereas the tasks used to study the inversion effect in visual recognition produce large and consistent inversion effects, the left-right task (with front-facing figures) usually does not produce substantially slower response times for inverted bodies (Parsons, 1987; Zacks & Tversky, 2005). The same-different task does produce slower response times when one of the two figures is inverted—this is the classical mental rotation effect (Shepard & Cooper, 1982). However, it is important to distinguish this from the inversion effect: In the same-different task, only one of the two pictures is inverted. The mechanism responsible for the effect in the same-different task is thought to be an object-based transformation, whereas in the inversion effect paradigms it is thought to be components of object recognition.

Stimulus characteristics

In the same-different task, our stimuli differed only by their handedness. Other researchers have elicited object-based transformations using stimulus pairs mismatched in other ways, such as perturbation of polygon vertices (Cooper & Podgorny, 1976), or changing only sub-parts of the figure (Takano, 1989). We suspect that such characteristics could place greater demands on structural knowledge of the figures and hence could show greater embodiment effects, perhaps even changing the efficiency of object-based transformations. For example, one could imagine using stimulus pairs that varied only in the placement of a single small element such as the flexion or extension of a human body's elbow, or the angle of one intersection in a Shepard-Metzler figure. We predict that this task would be more demanding of visual recognition processes, and that this would encourage participants to adopt a motor simulation strategy. If so, it could be that depicting the same geometric configuration as a human figure would facilitate the motor simulation compared to depicting the configuration as an array of blocks; this would correspond to the motoric embodiment described by Amorim, Isableu, and Jarraya (2006).

Our stimulus set included only picture-plane rotations; this contrasts with that of Amorim, Isableu, and Jarraya (2006), which also included depth-plane rotations (where objects were always upright but were rotated around their vertical axis). This difference may explain the relatively fast RTs, though it does not explain our failure to find effects of body features on RT slopes. Moreover, Amorim and colleagues found that there were in fact reduced body-related differences for some planes of rotation—specifically in trials with figures rotated around the vertical axis. They explained that for such figures the top and bottom of the figures were always clear (as they were always at the top and bottom of the screen), eliminating the advantage for identifying the orientation of human-like figures with heads. This explanation is consistent with our account of body features facilitating the initial stimulus recognition and encoding processes.

Task characteristics

Our version of the same-different task used simultaneous presentation of the paired figures, whereas other investigators have required participants to judge figures presented successively (Sayeki, 1997; Shepard & Metzler, 1971). Simultaneous presentation may reduce the difficulty associated with comparing two figures, in that it is unnecessary to construct and maintain a representation of the first figure across the inter-stimulus interval in order to compare it to the second figure. Conversely, successive presentation of the two comparison figures would place greater demands on the ability to maintain the representation of the first figure. We suspect that successive presentation would amplify body feature effects in same-different task performance.

Visual familiarity vs. embodied experience

Reed, Nyberg, and Grubb (2012) contrasted visual familiarity with embodied experience by constructing stimuli with either familiar or unfamiliar visual forms (human vs. canine) and familiar or unfamiliar postures (human-typical posture such as taking a bow vs. canine-typical posture such as begging for food with paws outstretched). Their participants were asked to make same-different judgments to pairs of similar stimuli of presented sequentially. Reed and colleagues proposed that if embodied expertise facilitated recognition, this should be found for both upright and inverted figures; however, if facilitation was due to expertise in visual experience effects should be found only for upright figures because they are visually familiar. They found that both visual familiarity and embodied experience can contribute to recognition performance. Specifically, their results revealed an overall performance benefit for judgments of upright human figures relative to inverted human figures and canine figures, but also similar inversion effects for canine figures adopting human-typical postures. They interpreted this as indicating that when pictures of canines showed them in a human-like pose, people were able to use embodied recognition mechanisms to recognize them, not just visual experience. Similarly, we found body feature effects were similar at all orientations. This suggests that to the extent that familiarity contributes to faster recognition of body features, it is embodied familiarity rather than purely visual experience.

Animacy detection

Some researchers have argued for the existence of neural structures that are efficient at processing biologically-relevant features such as faces (Kanwisher, McDermott, & Chun, 1997) or bodies (Downing, Jiang, Shuman, & Kanwisher, 2001). Furthermore, there is evidence for the existence of specialized neural circuits to facilitate motor responses to natural scenes containing animals (Drewes, Trommershäuser, & Gegenfurtner, 2011; Kirchner & Thorpe, 2006) and human faces (Crouzet, Kirchner, & Thorpe, 2010). These studies employed saccadic responses to show that the ventral stream processing mechanisms presumably used to identify animate things can rapidly influence motor programming for saccade generation. Animacy-specific detection mechanisms could explain our observation that block-head figures were associated with faster responses than block-teapot figures, even though they both presented a similar level of basic information about the spatial arrangement of the figure.

Conclusion

In two spatial judgment tasks, we found that people responded more quickly to stimuli with more features of a human body. The present results for the same-different task diverge from previous reports in that the facilitative effect of body features was limited to non-orientation-dependent processes (Amorim et al., 2006; Sayeki, 1997). These results clarify what it means for a spatial transformation to be "embodied." Bringing to bear knowledge about one's own body may increase the efficiency of some spatial transformations. It may change which spatial transformation one performs to solve a problem. But for some important classes of spatial reasoning problems, body knowledge may simply facilitate the perception of orientation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by a United States Department of Defense Science, Mathematics, and Research for Transformation (SMART) scholarship to the first author, and NIMH grant RO1MH70674 and NIA grant R01AG031150 to the second author.

References

Amorim M-A, Isableu B, Jarraya M. Embodied spatial transformations: "Body analogy" for the mental rotation of objects. Journal of Experimental Psychology: General. 2006; 135:327–347. [PubMed: 16846268]

Brandman T, Yovel G. The body inversion effect is mediated by face-selective, not body-selective, mechanisms. The Journal of Neuroscience. 2010; 30:10534–10540. [PubMed: 20685996]

Carey S, Diamond R. Are faces perceived as configurations more by adults than by children? Visual Cognition. 1994; 1:253–274.

Cooper LA, Podgorny P. Mental transformations and visual comparison processes: Effects of complexity and similarity. Journal of Experimental Psychology: Human Perception and Performance. 1976; 2:503–514. [PubMed: 1010999]

Cooper LA, Shepard RN. The time required to prepare for a rotated stimulus. Memory & Cognition. 1973; 1(3):246–250. [PubMed: 24214553]

Crouzet SM, Kirchner H, Thorpe SJ. Fast saccades toward faces: face detection in just 100 ms. Journal of Vision. 2010; 10:1–17. [PubMed: 20465335]

Davidenko N, Flusberg SJ. Environmental inversion effects in face perception. Cognition. 2012; 123:442–447. [PubMed: 22425642]

Downing PE, Jiang Y, Shuman M, Kanwisher N. A cortical area selective for visual processing of the human body. Science. 2001; 293:2470–2473. [PubMed: 11577239]

Drewes J, Trommershäuser J, Gegenfurtner KR. Parallel visual search and rapid animal detection in natural scenes. Journal of Vision. 2011; 11:1–21.

Gregory E, McCloskey M. Mirror-image confusions: Implications for representation and processing of object orientation. Cognition. 2010; 116(1):110–129. [PubMed: 20466358]

Huttenlocher J, Presson CC. Mental rotation and the perspective problem. Cognitive Psychology. 1973; 4:277–299.

Kanwisher N, McDermott J, Chun MM. The fusiform face area: a module in human extrastriate cortex specialized for face perception. Journal of Neuroscience. 1997; 17:4302–4311. [PubMed: 9151747]

Kessler K, Rutherford H. The two forms of visuo-spatial perspective taking are differently embodied and subserve different spatial prepositions. Frontiers in Psychology. 2010; 1:213. [PubMed: 21833268]

Kirchner H, Thorpe SJ. Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. Vision Research. 2006; 46:1762–1776. [PubMed: 16289663]

May M, Wendt M. Separating mental transformations and spatial compatibility effects in the own body transformation task. Cognitive Processing. 2012; 13:257–260.

McCloskey M. Spatial representation in mind and brain. What Deficits Reveal about the Human Mind/Brain: A Handbook of Cognitive Neuropsychology. 2001:101–132.

Michelon P, Zacks JM. Two kinds of visual perspective taking. Perception and Psychophysics. 2006; 68:327–337. [PubMed: 16773904]

Parsons LM. Imagined spatial transformation of one's body. Journal of Experimental Psychology: General. 1987; 116:172–191. [PubMed: 2955072]

Presson CC, Montello DR. Updating after rotational and translational body movements: coordinate structure of perspective space. Perception. 1994; 23:1447–1455. [PubMed: 7792134]

Reed CL, Nyberg AA, Grubb JD. Contributions of visual and embodied expertise to body perception. Perception. 2012; 41:436–446. [PubMed: 22896916]

Reed CL, Stone VE, Bozova S, Tanaka J. The body-inversion effect. Psychological Science. 2003; 14:302–308. [PubMed: 12807401]

Robbins RA, Coltheart M. The effects of inversion and familiarity on face versus body cues to person recognition. Journal of Experimental Psychology: Human Perception and Performance. 2012; 38:1098–1104. [PubMed: 22642217]

Rock, I. Orientation and form. New York: Academic Press; 1973.

Sayeki, Y. “Body analogy” and the cognition of rotated figures. In: Cole, M., et al., editors. Mind, Culture, and Activity. 1997. p. 90-99.

Shepard, RN.; Cooper, LA. Mental images and their transformations. Cambridge, MA: MIT Press; 1982.

Shepard RN, Metzler J. Mental rotation of three-dimensional objects. Science. 1971; 171:701–703. [PubMed: 5540314]

Takano Y. Perception of rotated forms: a theory of information types. Cognitive Psychology. 1989; 21:1–59. [PubMed: 2920537]

Tarr MJ, Pinker S. Mental rotation and orientation-dependence in shape recognition. Cognitive Psychology. 1989; 21:233–282. [PubMed: 2706928]

Yovel G, Pelc T, Lubetzky I. It’s all in your head: Why is the body inversion effect abolished for headless bodies? Journal of Experimental Psychology: Human Perception and Performance. 2010; 36:759. [PubMed: 20515202]

Yu AB, Zacks JM. The role of animacy in spatial transformations. Memory & Cognition. 2010; 38:982–993. [PubMed: 20921110]

Zacks JM, Michelon P. Transformations of visuospatial images. Behavioral and Cognitive Neuroscience Reviews. 2005; 4:96–118. [PubMed: 16251727]

Zacks JM, Mires J, Tversky B, Hazeltine E. Mental spatial transformations of objects and perspective. Spatial Cognition & Computation. 2002; 2:315–322.

Zacks JM, Tversky B. Multiple systems for spatial imagery: Transformations of objects and bodies. Spatial Cognition and Computation. 2005; 5:271–306.

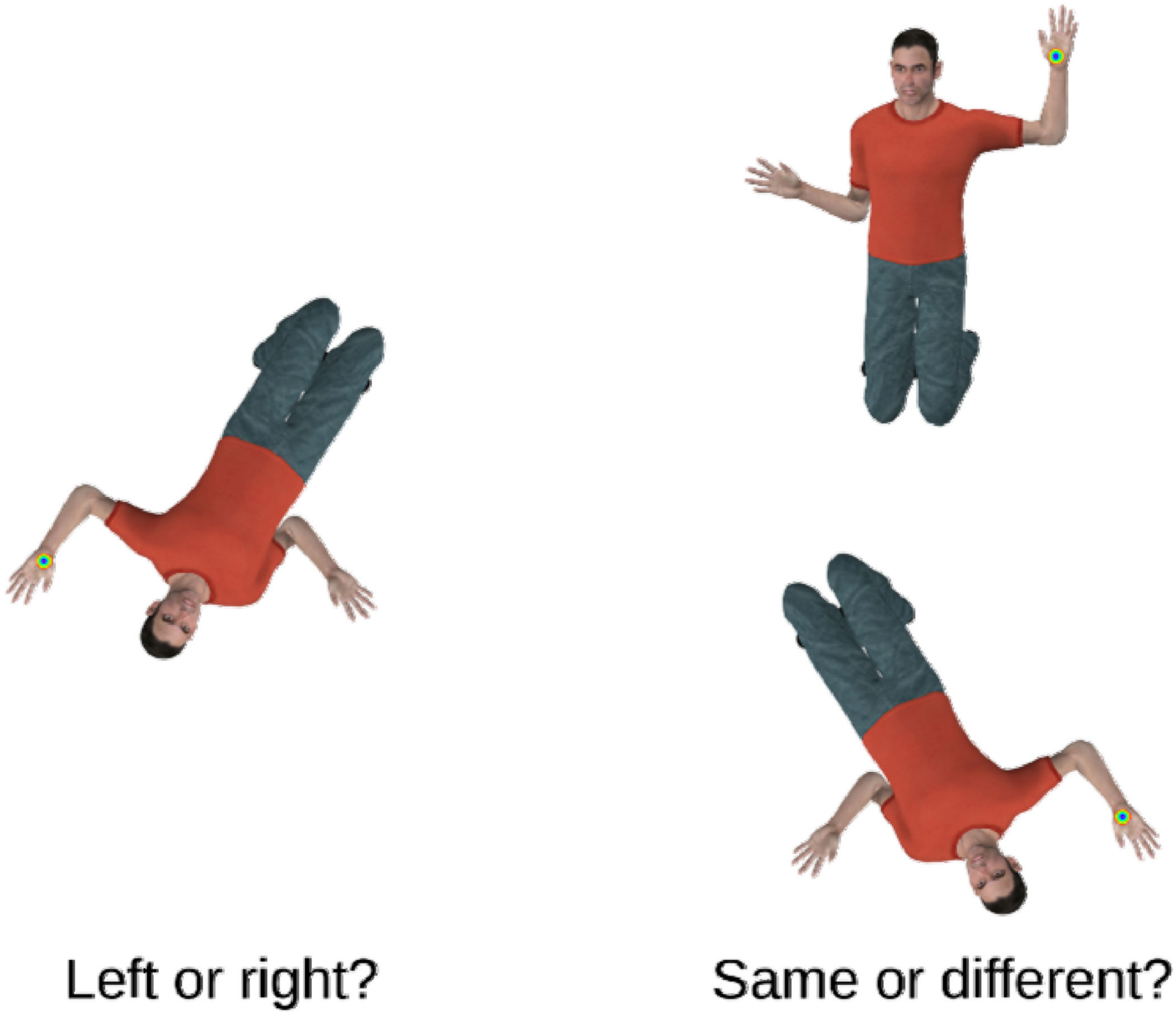

Figure 1.

Examples of the left-right and same-different tasks. For the left-right task, participants are asked to identify which side of the figure is marked, from the figure's point of view. (The correct answer here is "left.") For the same-different task, participants are asked whether the same side is marked in the two figures. (The correct answer here is "different.")

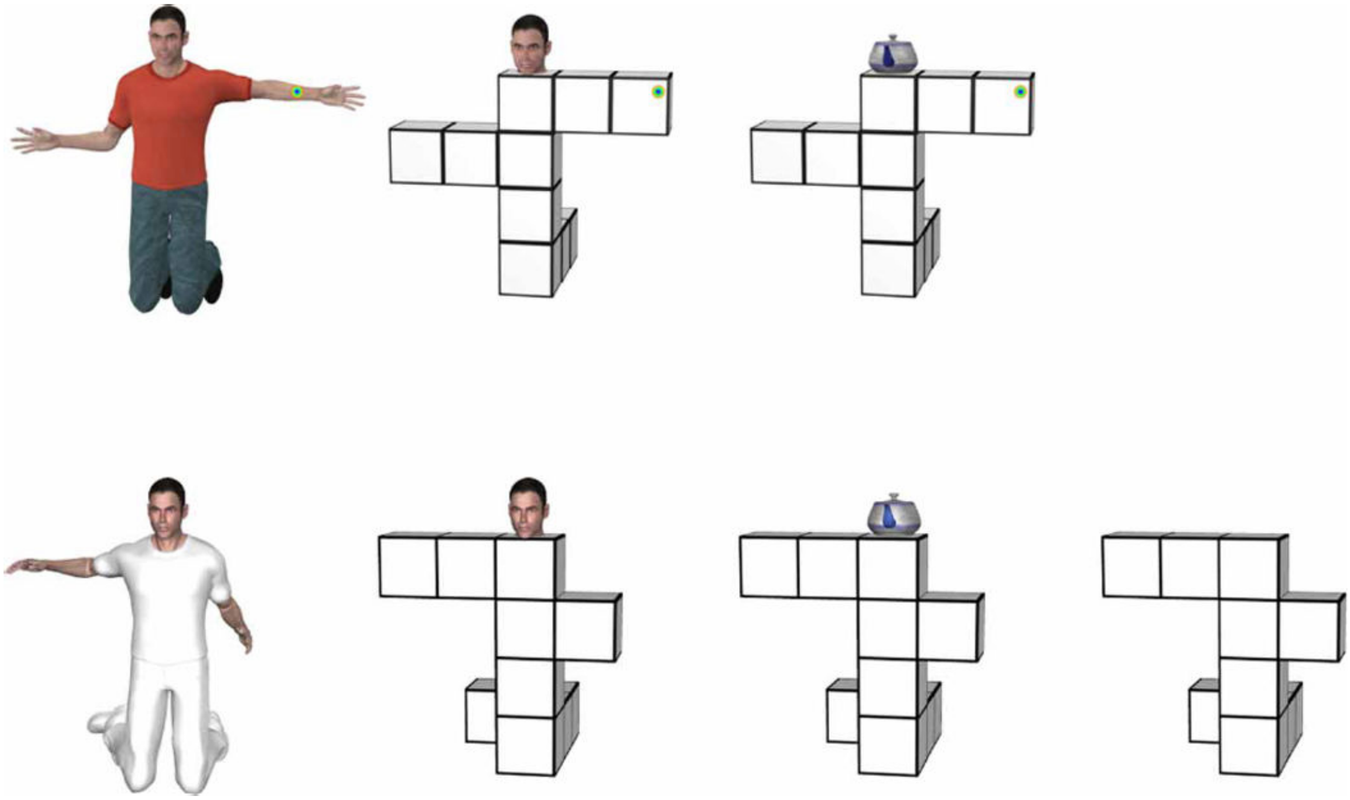

Figure 2.

Upper row: Example of the body, block-head, and block-teapot figures used in Experiment 1. Lower row: Example of the body, block-head, block-teapot, and block-only figures used in Experiment 2.

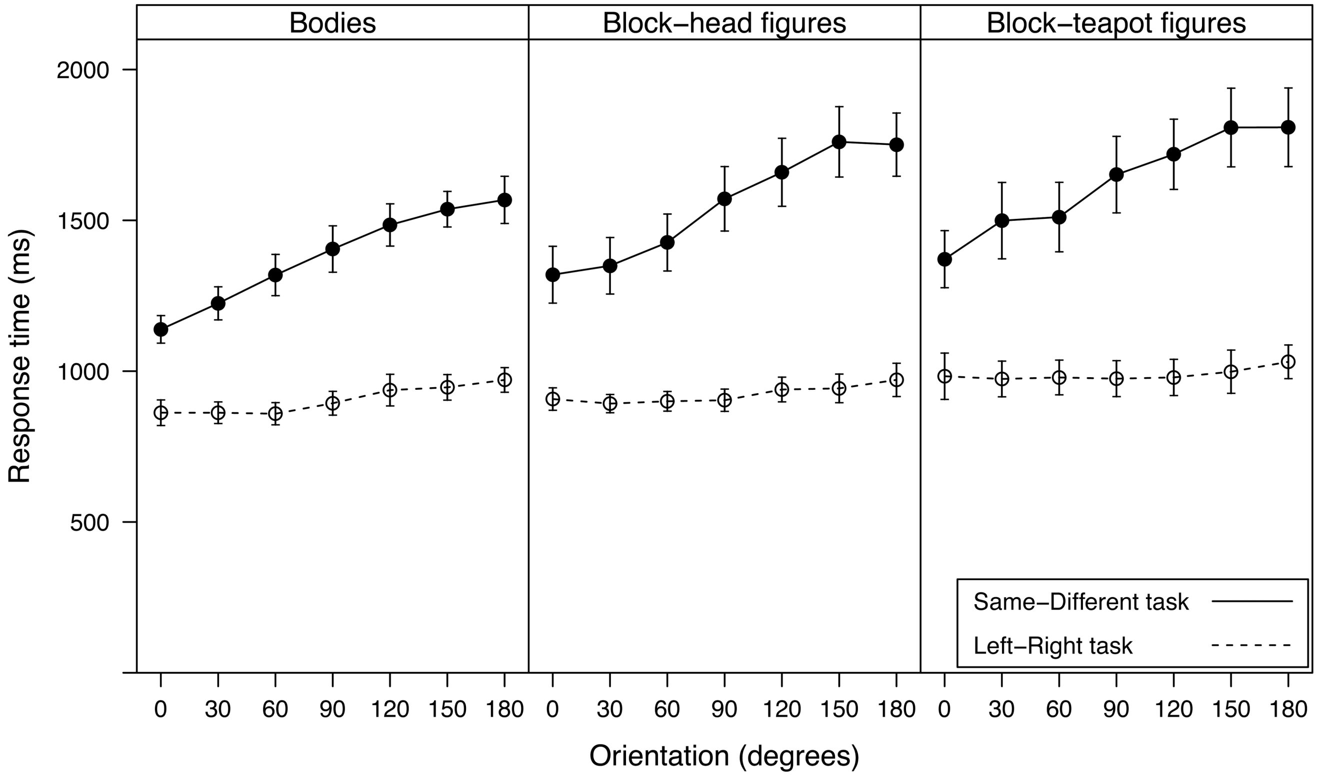

Figure 3.
Trimmed response times in Experiment 1. Error bars represent the standard error of the mean.

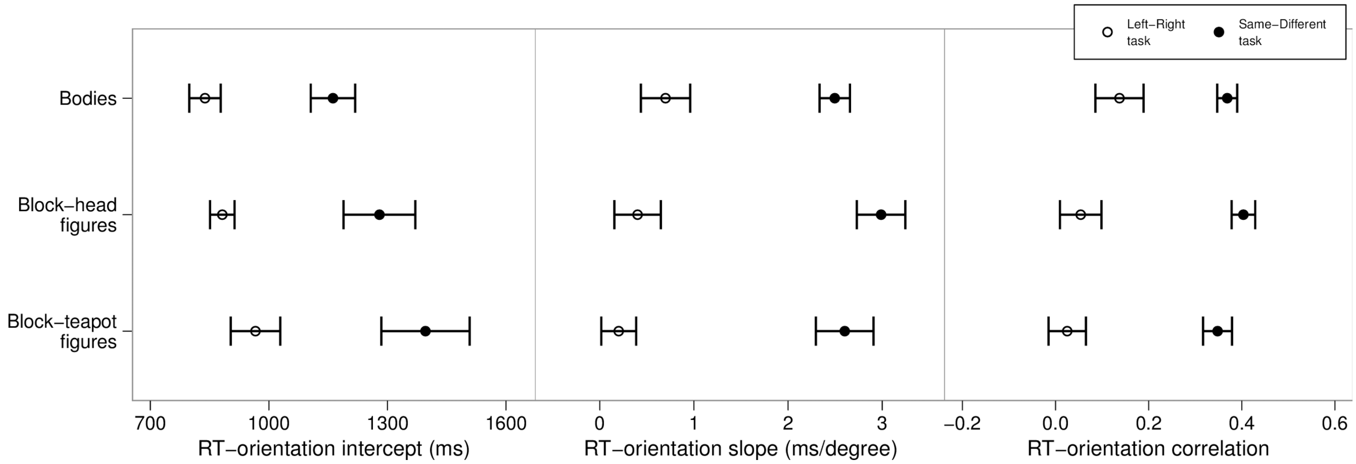

Figure 4.

Response time component analyses for Experiment 1. Error bars represent standard error of the mean across subjects.

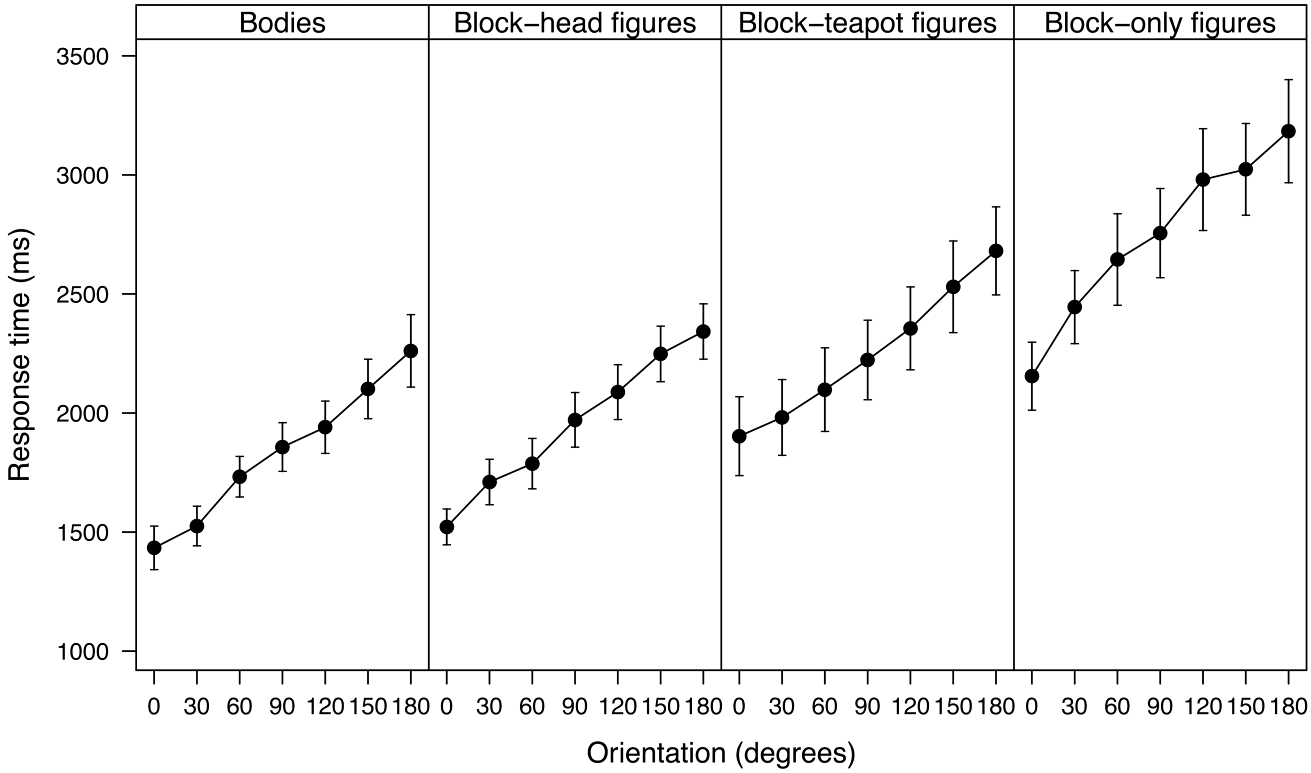

Figure 5.

Trimmed response times in Experiment 2. Error bars represent the standard error of the mean.

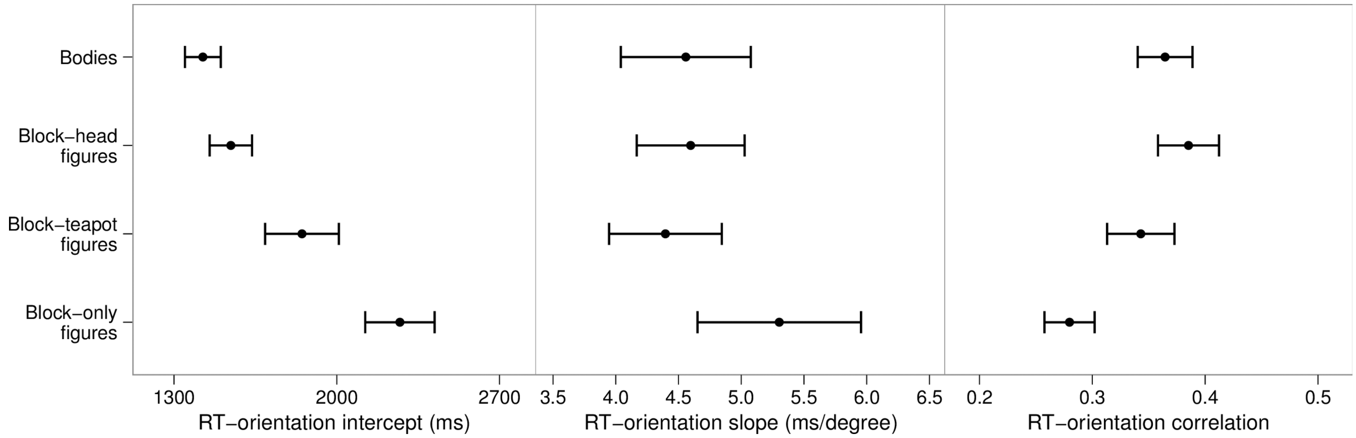

Figure 6.

Response time component analyses for Experiment 2. Error bars represent the standard error of the mean.