Differential effects of object orientation on imaginary object/viewer transformations

ROB VAN LIER
University of Nijmegen, Nijmegen, The Netherlands

Given a specific view of a simple symmetrical object, participants were asked whether a certain imaginary transformation could result in a second viewed image. An experiment was conducted in which the participants had either to mentally rotate an object or to imagine themselves looking at the object from another position (i.e., the object-based condition and the viewer-based condition, respectively). In the experiment, combinations of these imagery tasks (i.e., the combined conditions) were also included. The symmetrical objects could be oriented horizontally or vertically. The performance in the object-based conditions was generally equal to or better than the performance in the viewer-based conditions. In addition, there were more confusions for shapes with a horizontal orientation, especially when viewer-based upside-down rotations were involved, with an apparent mediating role of object rotation in the combined conditions.

We usually look at an object from one specific viewpoint. There are two obvious actions an observer can carry out to obtain other views of an object. One consists in a repositioning of the object, as when we rotate it in 3-D space; the other consists in a repositioning of the observer, as when we walk around the object. Given a specific view of an object, observers can often reliably derive a different view of the object without physically performing one or the other action. In the past, such viewpoint changes have most frequently been studied by means of object transformations (see, e.g., Biederman & Gerhardstein, 1993; Cooper, 1975; Pani, 1993: Rock & DiVita, 1987; Shepard & Metzler, 1971; Tarr, 1995; van Lier & Wagemans, 1998, 1999; Yuille & Steiger, 1982). In most experimental tasks, different 2-D projections of 3-D objects must be compared. Yet although pairs of such images reflect differences in viewpoint, the manipulation itself is actually based on differences in object orientation, because the viewpoint of the observer remains the same. This manipulation contrasts with events in the real world, where viewpoint changes usually occur through self-motion (cf. Simons & Wang, 1998). The difference between these two has often been neglected, possibly because it has been assumed that equivalence in change of the retinal projection would require equivalent mental transformations (cf. Wang & Simons, 1999). Recently, various investigations of these differences have been performed. For example, Simons and Wang showed that performance was much better when the viewpoint changes were caused by viewer movements. Generally they found that changes in the spatial layout of an array of objects was detected reliably following an observer's motion, whereas many more errors were made when the array of objects rotated.

Given these two ways of achieving different views of an object in the real world, other studies have focused on differences between imaginary object transformations and imaginary viewer transformations (Amorim & Stucchii, 1997; Carpenter & Proffitt, 1991; Creem, Wraga, & Proffitt, 1991; Huttonlocher & Presson, 1979; Presson, 1982; Wraga, Creem, & Proffitt, 1999, 2000). On the whole, such imagery studies have revealed that observers perform better when imaginary viewer transformations are requested. For example, Amorim and Stucchii showed that for such imaginary transformations observers experienced greater difficulties when they had to imagine a display rotating than when they had to imagine themselves moving to another position, a result that resembles the "real-world" results found by Simons and Wang (1998). Recently, Proffitt and colleagues have performed an impressive series of studies to further investigate the differences between imagined viewer and object rotations. Their studies (Carpenter & Proffitt, 1991; Creem et al., 2001; Wraga et al., 2000) basically confirm the notion that participants perform better on imagined viewer rotations than on imagined object rotations. In their experiments, which were based on variations of an experimental paradigm used by Presson and collaborators (Huttonlocher & Presson, 1979; Presson, 1982), the participants faced life-size arrays of objects. The participants were first asked to memorize the position of the objects and were then asked to perform either an imaginary self-rotation or an imaginary array rotation. In addition, they had to answer specific questions about the objects. Besides finding the general viewer ad-
vant age, Wraga et al. (2000) showed that this advantage could be attributed partly to a difference between the cohesiveness of the viewer frame of reference and that of the object frame of reference. Indeed they found some improvement on imagined object rotation when they used highly familiar objects or when participants could use haptic information; but generally, the imagined viewer rotation advantage persisted. Wraga et al.’s (2000) suggestion that this viewer advantage has ecological plausibility because most observers will have more experience with motion in a static environment has been strengthened further by Carpenter and Proffitt (2001), who showed a strong influence of the plane in which the rotation had to be established. Most noticeably, the viewer advantage disappeared when the rotation had to be performed not in the familiar horizontal plane, but in planes orthogonal to that plane. In addition, Creem et al. (2001) showed in a series of experiments that the orthogonality of the relationship between the viewer’s orientation and the imagined rotation is more important than gravity with respect to the viewer advantage. The present experiment was designed as an attempt to give an additional twist to the object/viewer discussion in mental imagery tasks by investigating whether the difficulty or ease with which an object-based or viewer-based transformation is performed depends on the orientation of the shape. In past decades, the importance of object orientations on mental rotation has been established in various studies on object rotations (e.g., Pani, 1993; Pani, William, & Shipley, 1995; Parsons, 1995). Whereas these studies established the influence of the orientation of the object axis on object rotation, the present goal was to investigate whether the direction of the object axis would be found to have different effects on the object-based and viewer-based conditions. The stimuli consisted of simple V-like shapes that had an unambiguous orientation (an axis of symmetry and a varying width at both sides of the axis). Given a specific view of such a shape, participants were asked to perform a certain imaginary transformation on that shape and to subsequently respond whether a second image correctly displayed the result of the requested transformation. In the object-based condition, they had to mentally reposition the object with respect to themselves. In the viewer-based condition, the participants had to mentally reposition themselves with respect to the object. In this way, an object-based y-axis rotation (180°) would result in the same image as an upright view of the object from the rear side (to be referred to as a rear/upright view). Similarly, an object-based z-axis rotation (180°) would equal a front/upside-down view, and an x-axis rotation (180°) of the object would equal a rear/upside-down view.

METHOD

Participants

There were 12 participants in the experiment, all were students at the University of Nijmegen who received either course credits or a small fee (approx. 5€). All participants were naïve with respect to the purpose of the experiment.

Stimuli

The stimuli consisted of simple line-drawn V-like shapes. Such shapes obviously have only one axis of symmetry. The axis of symmetry could be aligned with the environment’s vertical or with the environment’s horizontal. In this way, the V-like shapes could point up, down, left, or right. During the experiment, the participants had to mentally rotate the shape or to imagine themselves looking at the shape from another position (i.e., the object-based condition and the viewer-based condition, respectively). In each trial, the drawings in the display indicated the mental operation to be performed.

In Figure 1A, a typical stimulus display is depicted. The left half of this display, the initial situation is shown. The line-drawn “puppet” observer in front of the rectangular outline at the top left was always aligned with the environment’s vertical. The participants were told that they should imagine themselves to be this puppet-observer watching a “transparent piece of paper” (represented by the rectangular outline) on which a specific shape was drawn. This shape was always presented at the bottom left of the display. Thus, in the initial

![Figure 1](image-url)
situation, the observer was always seen from the back, looking at a shape from the same viewpoint (i.e., with the same 2-D projective properties), as that from which the participant saw the shape. In the right half of the screen, the required mental transformation is shown together with a second shape, which could be (or could not be) the image after that transformation. The participants were told that the rotation was always a 180° turn, which could proceed in either direction around the specific axes. There were six single transformations: three object-based transformations (180° x-axis, 180° y-axis, and 180° z-axis, with the x-axis aligned with the environment’s horizontal, the y-axis with the vertical, and the z-axis orthogonal to both) and three viewer-based transformations (rear/upside-down, front/upside-down). The transformation required in Figure 1A, for example, comprises a 180° rotation of the object around the y-axis. In Figure 1B, the display for the corresponding viewer-based condition (i.e., the rear/upright view) is given. In addition to these single-transformation trials, combined conditions were included as well, in which both an object-based transformation and a viewer-based transformation were required. Figure 1C shows a display with such a double irrotation transformation. In fact, all possible combinations of the object-based and viewer-based transformations were included in the experiment. In Figure 2, the icons representing all transformation conditions are shown.

**Procedure**

Within each trial, all aspects of the whole stimulus (the initial situation, the required transformation, and the shapes) appeared on the screen simultaneously. This display remained on the screen until a response was given. All participants received an elaborate instruction about the stimulus display, the icons used, and the specific mental imagery tasks. In the instruction phase, it was stressed that they really had to try to perform the required imaginary object-based or viewer-based rotation task. Before the experiment was run, the experimental procedure and the various conditions were practiced with a different set of shapes. During this practice, feedback was given with respect to the accuracy of the participant’s response immediately after each trial. After the instruction and the practice trials, all participants indicated that they understood the task. Responses were given by pushing one of two indexed keys on a keyboard (“Yes” if the image on the right could be the result of the required transformation, starting with the image on the left, and “No” if it could not be the result of the required transformation).

All participants received all possible trials in a random order. There were equal numbers of matches and nonmatches. Nonmatch trials always consisted of the same shape and the same horizontal/vertical symmetry (only the pointing direction of the V-like shapes left/right or top-bottom was opposite to the correct direction). In total there were 16 different initial configurations (2 sets of V-like shapes with 2 horizontal orientations and 2 vertical orientations). 16 transformation conditions (see Figure 2), and both a match and a nonmatch condition for each combination, resulting in 256 trials per participant. The room was dimly lit, and the participants were seated about 1 m from the monitor.

**RESULTS**

A repeated measures analysis of variance was performed on the following design: object rotation × viewer rotation
× object axis (horizontal vs. vertical orientation). The dependent variable was the response accuracy (number correct trials per cell/total number trials per cell). The analysis was performed on within-cell averages. There was a main effect of viewer rotation \( F(3,33) = 13.54, p < .0001 \) and a main effect of object axis \( F(1,11) = 7.34, p < .05 \). There was no main effect of object rotation \( F(3,33) = 2.26, p = .0999 \). There were two significant two-way interactions: object rotation × object axis \( F(3,33) = 3.69, p < .05 \) and viewer rotation × object axis \( F(3,33) = 8.58, p < .0005 \).

We first take a closer look at trials in which just a single imagined rotation had to be performed, which could be either an object rotation or a viewer rotation. In Figure 3A, the graph is plotted for the object rotation trials for both object axis orientations. For these data, the partial interaction of object rotation × object axis appears not to be significant \( F(2,22) = 1.29, p = .29 \). An additional least significant difference (LSD) post hoc analysis on the data of Figure 3A revealed no significant differences between accuracies for the same object transformation but different object axis orientation. In Figure 3B, the graph is plotted for the viewer rotation trials, again for both object axis orientations. For these data, the partial interaction of viewer rotation × object axis was significant \( F(2,22) = 4.23, p < .05 \). An LSD post hoc analysis on the data of Figure 3B revealed a significant difference between the accuracies for the horizontal and vertical object axis orientation of the rear/upside-down view (\( p < .0001 \), indicated with an asterisk below the graph). Notice that in this
condition, the imagined view for shapes with a horizontal axis was performed almost at chance level. In addition, we have compared the differences between the corresponding object-based transformations and viewer-based transformations (i.e., z-axis vs. front/upside-down; y-axis vs. rear/upright; x-axis vs. rear/upside-down) for both object axis orientations with an LSD post hoc analysis, revealing significant differences (at least \( p < .05 \)) for the z-axis versus front/upside-down (for both the vertical and the horizontal object axis orientations), and the y-axis versus rear/upside-down (again for both the vertical and the horizontal object axis orientations).

These data suggest that the imagined upside-down views cause the most confusion, most clearly so for the shapes with a horizontal axis in the rear/upside-down condition.

The difficulty of this imaging task becomes even more prevalent on trials in which both a viewer transformation and an object transformation had to be performed. For these data, the partial interaction of object rotation \( \times \) object axis and the partial interaction of viewer rotation \( \times \) object axis were both significant \( (F(2,22) = 5.10, p < .05, \) and \( F(2,22) = 9.20, p < .005, \) respectively). In Figures 3C-1 to 3C-3, the accuracy for each of the three object transformations in combination with one of the three viewer transformations have been plotted. In each of the graphs, significant differences according to an LSD post hoc analysis (at least \( p < .05 \)) between accuracies for the same object/viewer transformation but different object axis orientation are indicated by means of an asterisk; all but one upside-down view revealed significant differences between the two object axis orientations, such that the accuracy on the horizontal object axis orientation was (much) lower than the accuracy on the vertical object axis orientation. In the only case in which there was no significant difference, both the viewer and the object had to be rotated in the picture plane (i.e., a front/upside-down rotation combined with a z-axis rotation).

**DISCUSSION**

Generally, the results revealed that the imagined object rotations might very well lead to better performances than would imagined viewer rotations: especially when viewer upside-down rotations were required, participants made many more errors. In addition, the object axis orientation influenced the performances. More specifically, it appeared that the performance on horizontal object axis orientations was poorer than the performance on vertical object axis orientations, most likely so when upside-down viewer-based rotations were involved.

The overall advantage for object-based transformations in the present study contrasts with the findings of Carpenter and Proffitt (2001), who obtained a viewer-based advantage for rotations in the \( x, z \) plane (i.e., the current upright viewer conditions), which disappeared for rotations in the \( x, y \) plane and the \( y, z \) plane (which the current upside-down viewer conditions are part of). However, notice that, notwithstanding this overall difference, both studies revealed the best viewer results in the upright conditions. The overall difference between imagined object transformations and imagined viewer transformations in these studies might be related to differences in the experimental paradigm and the specific stimulus layout. For example, the experiments of Carpenter and Proffitt, but also various other studies that focused on the difference between imaginary viewer and imaginary object transformations (e.g., Creem et al., 2001; Presson, 1982; Wraga et al., 2000) used a "predicted outcome task." It has been established that observers may experience great difficulties on object transformations with such tasks (see, e.g., Pani, 1993; Parsons, 1995), especially when misalignments occur (between, e.g., environment and object, as in oblique object orientations). As might be clear, these kinds of misalignments did not occur in the present experiment. In addition, the distinction between a predicted outcome task and the present matching task has been minimized here, because in fact the image properties remained exactly the same, except that the V-like shapes could point in the opposite direction after the imaginary transformation. Presumably more important here is the difference in stimulus layout. As already mentioned, Carpenter and Proffitt (2001) used a life-size array of four different objects with the participant standing in the middle. In the present experiment, the whole configuration was displayed on a computer screen. The present flat-display set-up maintained the analogy with much previous research in which mental rotation of objects had been studied. Nevertheless, it may also have caused the generally poorer performances on the viewer rotations—perhaps just as, in a real-world setting, rotations of life-size multiple-object arrays may be more difficult to imagine. In addition, the fact that we used small and simple shapes (see also Wraga et al., 2000) presumably also enhanced object rotations relative to viewer rotations.

Notwithstanding these differences, one can of course still conclude that, given the present task with simple symmetrical shapes, imagined viewer transformations appear more difficult than imagined object transformations for instances in which the viewer was to be imagined upside-down, and that, most noticeably for these upside-down views, object orientation played an important additional role. That is, when viewer-based transformations were to be made, there were many more errors for the horizontally oriented shapes than for the vertically oriented shapes. In the single-transformation conditions, this difference was significant for the rear/upside-down view. In the combined conditions, there was a significant difference for five out of six viewer upside-down conditions. There were no differences between the performances on the two object orientations when the viewer front/upside-down rotation was combined with a rotation of the object around the z-axis, and whenever the viewer remained upright (i.e., rear/upright condition). Notice that in the former combination of a viewer front/upside-down rotation and an object z-axis rotation, the rotations actually cancel each other out, because the required rotations comprise twice the same ob-
ject/viewer-relative 180° turn. Interestingly, this aspect as such cannot account for this non-significant difference for that specific viewer upside-down rotation. Consider, for example, the performances on the two object orientations in the combined condition when the viewer was to be imagined rear/upside-down and the shape was imagined to be rotated around the x-axis. Again, in this combination, the rotations would cancel each other out. However, this time the difference between the horizontally and vertically oriented shapes is noticeably the largest of all. In the graphs on the combined conditions (Figure 3C), it can be seen that the x-axis object rotation induces the largest proportion of errors for both viewer upside-down rotations (front and rear). This apparent differential effect of object rotation when viewer upside-down rotations are required also appears from the partial object rotation × object axis interactions, additionally calculated for each of the viewer rotations: front/upside-down, F(2,22) = 4.18, p < .05; rear/upright, F(2,22) = 0.25, p = .78; rear/upside-down, F(2,22) = 3.83, p < .05.

Obviously, the horizontally oriented shapes have a unique left-right direction, whereas the vertically oriented shapes have a unique top-bottom direction. Generally, it can be said that there was a greater vulnerability to left-right errors when viewer upside-down rotations were required and that there was a mediating role of object rotation in the differential effect between the left-right/top-bottom confusions when combined with viewer upside-down rotations. The overall viewer-related left-right confusions fit with results of other studies. For example, Franklin and Tversky (1990) found decreased left-right performances in certain viewer-related position questions, and Parsons (1987) found impoverished left-right judgments when participants were asked to imagine rotating their bodies and then locate specific body parts. Carpenter and Pratt (2001) found, using the previously outlined task and multiple object arrays, a general advantage for top-bottom questions as compared with left-right questions for both viewer and object rotations. The present results reveal a clear object/viewer dissociation concerning the influence of an object’s orientation.

In the past, findings on mental rotations have been related to the use of multiple frames of reference, such as the object-related frame, the viewer-related frame, and the environmental frame (cf. Hinton & Parsons, 1988; Just & Carpenter, 1985; Wraba et al., 1999, 2000). In the present experiment, the first two frames were most distinctive, both of them aligned with the environmental frame. In the present imagery task, each mental transformation at least affected the object-related frame or the viewer related frame. Wraba et al. (2000) suggested that, in line with their viewer-advantage findings, the viewer-related reference frame might be transformed in a more holistic way than would the object-related reference frame. They further suggested that object rotations might operate in a less piecemeal fashion with increased cohesiveness of the object (similar results in this respect, albeit in a different context, have been found by van Lier & Wagemans, 1998). Wraba et al. (2000) restricted themselves to viewer rotations in the transverse plane (in which the viewer remained upright). Apparently, if other viewer rotations are performed, this viewer advantage may disappear (Carpenter & Proffitt, 2001), or may even turn into a severe disadvantage, as has been shown here. Now, just as the rotation of large object arrays may be established in a rather piecemeal way, it may well be that the viewer rotation becomes less cohesive and more piecemeal for upside-down rotations. In an evolutionary context, the difference between holistic object rotations and holistic viewer rotations seems to make sense. Taking various perspectives on a small cohesive object often occurs through one’s simply grasping the object and turning it around by hand in various possible orientations. Upside-down views are no exception to this. In contrast, we take other perspectives on arrays of objects while remaining upright mostly proceeds as we move our bodies in the environment, rather than turn the array of objects (cf. Simons & Wang, 1998). Upside-down views are rare in this case. In the present experiment, the viewer upside-down rotations might therefore have been performed in sequential steps. This could specifically have been the case for the rear/upside-down view, for which a viewer rotation about the y-axis plus a viewer rotation about the z-axis might have been performed sequentially. The double left-right inversion in this sequence (as opposed to a single top-bottom inversion) could account for the larger number of errors on the horizontal object orientations. For the front/upside-down view, however, such a decomposition in sequential steps is less obvious (which apparently fits with the approximately equal performance on both object orientations in Figure 3B). Nevertheless, the additional differentiation between the performances on both object orientations for the various combined conditions reveals a more complex error source—with a mediating role of object rotation for both upside-down views.

The present results can be regarded as an existence proof that imagined viewer transformations can be more error prone than imagined object transformations and that there can be differential effects of object orientation on imagined object/viewer transformations, at least for the present simple symmetrical shapes (by means of which left-right errors and top-bottom errors could clearly be discerned). However, it might be clear from the literature so far that mental rotation outcomes depend on various factors inherent to the specific experimental set-up, including the specific way of presentation, task, and type of objects. Therefore, care should be taken in generalizing results for mental rotation to other settings. This is of course true for the present study but also for studies in which a general advantage for imagined viewer rotations has been found. The contrasting experimental results once more point at the vulnerability of mental transformations to the specific experimental settings—which in fact only adds to the subtlety of these imagery phenomena.

REFERENCES


(Manuscript received August 23, 2001; revision accepted for publication April 22, 2002.)