Perception & Psychophysics
2006, 68 (2), 327-337

Two kinds of visual perspective taking

PASCALE MICHELON and JEFFREY M. ZACKS
Washington University, St. Louis, Missouri

Visual perspective taking can be used to determine where objects are located relative to another agent, or whether the agent can see a particular object. Four experiments indicated that different processes provide these different kinds of information. When participants were asked to report whether an object was to the left or to the right of another agent, response times (RTs) increased with increasing angular distance between the participant and the agent, suggesting that participants mentally transformed their perspective to align it with that of the agent. For visibility judgments, RTs were independent of the angle between the participant and the agent but increased with the distance between the agent and the object, suggesting that participants traced the agent's line of sight. Together, these data suggest that perspective taking encompasses at least two qualitatively different computational processes: one that updates the viewer's imagined perspective, and one that traces a line of sight.

Visual perspective taking (VPT) is the ability to predict the visual experience of another agent. This ability is valuable in contexts as diverse as avoiding predators, reasoning about what others know, navigation, and spatial problem solving. Depending on the situation and goals at hand, VPT allows one to predict qualitatively different kinds of information. In particular, one can predict (1) whether another person can see an object at all and (2) where objects are located relative to another person's egocentric reference frame. One possibility is that a general process of visual perspective taking is involved in both situations. However, an alternative hypothesis is that different operations are performed to calculate these two different types of information.

Studies of visual perspective taking in children suggest that more than one type of knowledge is necessary to achieve successful imaginary perspective changes. Specifically, a distinction has been made between knowledge about which objects are visible from another viewpoint (Level 1 knowledge) and knowledge about the visual aspects of a scene relative to an imagined viewpoint (Level 2 knowledge—Flavell, Everett, Croft, & Flavell, 1981; Salatas & Flavell, 1976). Level 1 knowledge is reflected by performance in hiding or occlusion tasks in which the child is asked either to position an object so that another person cannot see it or to decide whether or not another person can see a target object. Level 2 knowledge is usually tested using tasks in which children are asked to predict how an object or scene would look from another position. These two types of information are clearly different, but both require representing the fact that the other person has a perspective and calculating information about the difference between that person's perspective and one's own (Salatas & Flavell, 1976; Yaniv & Shatz, 1990). Level 2 knowledge typically appears later in development than Level 1 knowledge. However, this developmental progression does not reveal what processes support these two kinds of VPT.

Research on spatial transformations in adults provides one possibility. In these studies, paradigms have been used in which participants predict what a scene would look like if they were at a specified position (possibly different from their actual position). Typically, response times (RTs) are longer if the to-be-imagined position is misaligned with the participant's actual position (Amorim & Stucchi, 1997; Creem, Downs, Wraga, Proffitt, & Downs, 2001; Presson, 1982; Presson & Montello, 1994; Simons & Wang, 1998; Wang & Simons, 1999; Wraga, Creem, & Proffitt, 2000; Zacks, Vettel, & Michelon, 2003). This has been interpreted as reflecting the use of analogue *perspective transformations*, which update the location and/or orientation of one's egocentric perspective. The few studies of adult cognition that have been conducted to directly examine the degree to which participants take the perspective of another person suggest that perspective transformations are also used in this situation (Amorim, 2003).

Very few studies of adult VPT have addressed situations in which it is necessary to predict whether or not an object is at all visible from another agent's viewpoint (but see Kelly, Beall, & Loomis, 2004). This is somewhat surprising, given the importance in the developmental literature of Level 1 knowledge. Perspective transformations may not be necessary in such situations; judgments of visibility could be made on the basis of information about the line of sight of the other agent. Avoiding perspective transformations when possible may be adaptive for two reasons. First, tracing a line of sight may sometimes be easier or more accurate than performing a perspective transformation. Second, performing an imagined

---

We thank Ting-Ting Avis Lau for assistance with data collection and stimulus creation. Correspondence concerning this article should be addressed to P. Michelon, Department of Psychology, Washington University, Campus Box 1125, St. Louis, MO 63130 (e-mail: pmichelo@artsci.wustl.edu).

perspective transformation puts the viewer's actual reference frame in conflict with that of the other agent, which may induce errors and slow performance. Tracing a line of sight avoids this difficulty. Yaniv and Shatz (1990) proposed that computing the line of sight of another agent is the imaginary analogue of actually drawing a line from the agent to the target object. This would make this process similar to other visual imagination processes, such as mental scanning. In mental scanning tasks, participants are instructed to imagine tracing a path in a mental image. RTs are generally proportional to the length of the path (Denis, Gonçalves, & Memmi, 1995; Kosslyn, 1973).

On the basis of these considerations, we hypothesized that observers would make use of line-of-sight tracing to perform judgments about the visibility of objects, but they would use perspective transformations to make judgments about the locations of objects relative to another person. Two tasks were used to test this hypothesis, both of which involved making judgments about pictures of scenes containing objects, occluding walls, and a representation of a person (which we will refer to as an *avatar*). In the *visibility task*, participants estimated which objects the avatar could "see." In the *left–right task*, they estimated the location of objects in the scene relative to the avatar.

This design allowed us to make specific predictions regarding the variables that would affect response latency in each task. First, we hypothesized that to solve the left–right task, participants would perform perspective transformations. This strategy requires one to (1) locate the avatar in an egocentric space, (2) perform a perspective transformation so that one's imagined position matches the position of the avatar, (3) locate the target object in the transformed spatial representation of the scene, and (4) read off the coordinates of the object. If such an analogue transformation strategy is used, latencies are expected to increase with increasing angular disparity between the participant's and the avatar's positions.

Second, we hypothesized that the visibility task would be performed solely relative to the participant's egocentric reference frame. This strategy requires that one locate the avatar, locate the target object, and trace a mental line from one to the other. Because this involves no transformation that depends on the orientation of the avatar, RT was not expected to increase when the avatar was misaligned with the participant. However, because mental scanning time is proportional to the distance traversed (Kosslyn, 1973), it was hypothesized that for the visibility task RT would increase with increasing distance between the avatar and the target object.

Note that for both tasks, participants could in principle use either a perspective transformation or a line-of-sight tracing strategy. For the left–right task, participants could trace a line of sight from the avatar and then use one of several rules to classify regions on either side of that line as "left" and "right." (One such rule is, "If the avatar is on my right, then right is above the line of sight.") However, we predicted that such strategies would not be used because they require complicated verbal or mathematical computations, whereas the perspective transformation strategy allows one to read off the correct answer directly from a transformed spatial representation. For the visibility task, one could perform a perspective transformation to align one's egocentric reference frame with that of the avatar, generate a detailed mental image of the scene from that perspective, and then verify whether or not the object is in the resulting spatial representation of the scene. However, we predicted that this strategy would not be used because it requires a spatial transformation and the generation of a detailed visuospatial image, whereas line-of-sight tracing is likely to be more efficient and more accurate.

Finally, we tested the importance of the actual presence of an external agent for the performance of the two VPT tasks. As reviewed earlier, previous studies suggest that participants can perform perspective transformations in response to either a depiction of a human figure or a symbolic cue. Thus, it seemed likely that the left–right task can be performed whether or not an agent is shown and that perspective transformations will be used in both cases. For the visibility task, we hypothesized that line-of-sight computations would be involved when an agent is shown. However, it was not clear that such computations also would be used in the absence of an anthropoid agent. If not, then participants might adopt a strategy in which perspective transformations are used, which would produce an increase in response latency with increasing angular disparity between the participant's position and the to-be-imagined position.

EXPERIMENT 1

In this experiment, we tested the hypothesis that people use perspective transformations when they need to judge where objects are located relative to another person's egocentric reference frame, but they rely on tracing a line of sight to determine which objects that person can see. The visibility task and the left–right task described above were used. The angular disparity between the avatar and the participant was varied from trial to trial.

Method

Participants. Twenty-four Washington University students (4 male, mean age = 19.42 years) participated in the experiment in return for course credit.

Stimuli. The participants made judgments about color photographs of a square table on which eight objects and four occluders were positioned (Figure 1A). The avatar was a female doll. The angular disparity between the participant's position and the avatar's position varied from 0º to 315º in increments of 45º. Four views of the display were used, one from each side of the table.

Design and Procedure. In the visibility task, the participants judged whether or not the avatar could see a target object. The target object was occluded on half of the trials. In the left–right task, the participants judged whether a target object was on the avatar's left side. All target objects were visible to the avatar, and half were on the left of the avatar. In both tasks, the participants responded by using buttons marked "Yes" and "No." To avoid stimulus–response incompatibility in the left–right task, the "Yes" button was always on the participant's left. The order of the two tasks was counterbalanced across participants. Each task involved 64 trials (left–right

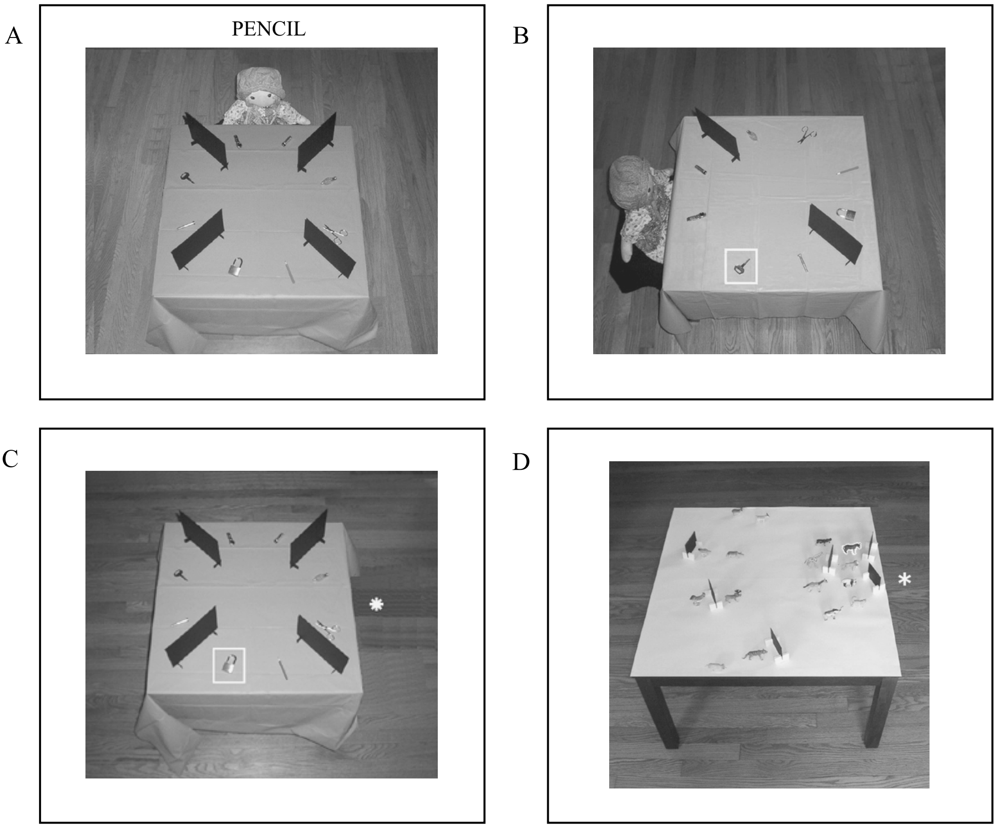

Figure 1. Example of the displays used in Experiment 1 (panel A), Experiment 2 (panel B), Experiment 3 (panel C), and Experiment 4 (panel D). In the left–right task, participants decided whether or not the target object was on the left side of the doll avatar. In the visibility task, they decided whether or not the avatar could see the target object. Photographs of the display were originally color photographs. Target objects were indicated by a yellow square (Experiments 2 and 3) or outlined in yellow (Experiment 4), and a yellow asterisk represented the avatar in Experiments 3 and 4.

task, 4 views × 8 angles × 2 conditions [left vs. right]; visibility task, 4 views × 8 angles × 2 conditions [visible vs. occluded]). These trials were repeated twice, for a total of 128 trials per task.

Each trial started with a central crosshair (500 msec) followed by a photograph of the display with the name of the target object presented above the photograph (Figure 1A). The display remained on the screen until the participant responded. The intertrial interval was 1,000 msec. After the tasks were performed, each participant completed a debriefing questionnaire inquiring about the strategy used in each task.

Results

In all analyses reported here, clockwise and counterclockwise positions were collapsed (e.g., 90º and 270º were combined). RTs were trimmed so that outlying RTs (more than two *SD*s from the participant's mean for that condition) were excluded.

The error rate was low in both the visibility task ($M = 7.81\%$, $SD = 8.79\%$) and the left–right task ($M = 4.27\%$, $SD = 5.40\%$). Therefore, no further analyses were performed on errors. On the basis of the trimming criteria, 4.36% of the RTs in the left–right task and 7.68% of those in the visibility task were deleted. Trimmed correct RTs were analyzed by computing means for each participant for each condition and performing a 2 × 5 ANOVA with task and angular disparity between avatar and participant as within-subjects factors. The results are shown in Figure 2. RTs were longer in the visibility ($M = 2{,}094$ msec, $SD = 392$) than in the left–right ($M = 1{,}912$ msec, $SD = 554$) task, although this main effect failed to reach significance [$F(1,23) = 1.91$, $p = .13$]. RTs increased with increasing angular disparity in the left–right task but not in the visibility task, leading to a significant main effect of angular disparity [$F(4,93) = 12.70$, $p < .001$] and a significant task × angular disparity interaction [$F(4,93) = 23.62$, $p < .001$].

To further characterize the task × angular disparity interaction, separate ANOVAs were performed for each task with angular disparity as a within-subjects factor. In the left–right task, there was a significant effect of angular disparity [$F(4,92) = 19.98$, $p < .001$]. A planned contrast indicated that RTs increased linearly with increasing angular disparity [$t(23) = 6.39$, $p < .001$]. In the visibility task, RTs also varied with angular disparity [$F(4,92) = 14.68$, $p < .001$]. However, as can be seen in Figure 2, this was mostly due to slower responses when the doll was positioned at the corners of the table (45º and 135º) rather

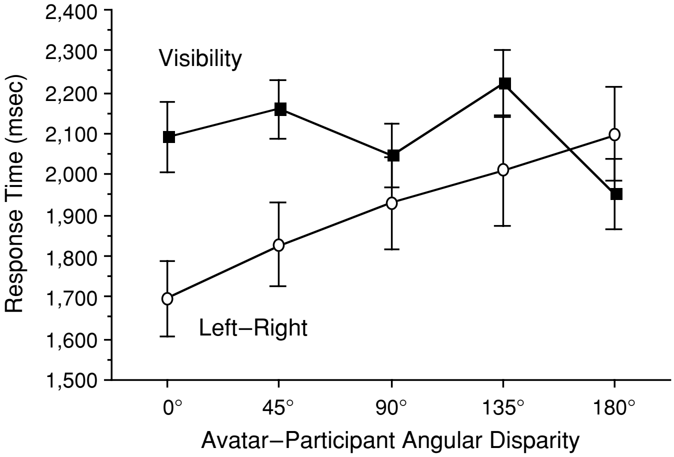

Figure 2. Experiment 1: Response times (in milliseconds) in the left–right and visibility tasks as a function of the angular disparity between the avatar's and the participant's positions. Vertical bars represent standard errors of the means.

than at the sides (0º, 90º, and 270º) [$t(23) = 4.72$, $p < .001$]. To the extent that there was a linear relationship between angular disparity and RT, this relationship was decreasing rather than increasing [$t(23) = 2.97$, $p < .008$].

Discussion

The fact that RT increased with the angle between the participant and the avatar for the left–right task suggests that in this task the participants performed perspective transformations. As was described in the introduction, this requires the participants to transform their egocentric frame of reference to align it with that of the avatar. Once this is done, they have to locate the target object relative to their imagined egocentric reference frame.

Latencies in the visibility task did not increase with increasing angular disparity, suggesting that perspective transformations were not used. This finding is consistent with the hypothesis that the participants traced the avatar's line of sight in order to perform the visibility task. However, during the postexperimental debriefing, 11 of the 24 participants reported that over the course of the task they attempted to memorize the positions of the objects the avatar could see depending on its position. We tried to prevent the participants from using such strategies by using four views of the display, which allowed us to vary the position of the objects throughout the task. However, the use of only four arrangements may be limited enough to allow for the emergence of mnemonic strategies. The use of such strategies could provide an alternative explanation for the fact that the latencies in the visibility task did not increase with angular distance.

Inspection of the pattern of response latencies in the visibility task was consistent with the use of memory-based strategies by some participants. RTs to corner positions (45º and 135º) were greater than RTs to side positions (90º and 180º). Because the table itself provides a salient environmental reference frame, the alignment of this reference frame with the avatar's position in the 90º and 180º angular disparity conditions may have afforded easier memorization of the positions of objects relative to those positions (Shelton & McNamara, 2001).

In sum, the data from Experiment 1 are consistent with the hypothesis that two distinct spatial processes were used to perform the two different perspective-taking tasks. However, part of the observed pattern could be attributed to the use of a memory-based strategy in the visibility task. Experiment 2 was designed to eliminate the use of memory-based strategies and to test a second prediction of the hypothesis that people tend to use line-of-sight tracing to make visibility judgments.

EXPERIMENT 2

Our first goal in this experiment was to prevent participants from relying on a memory-based strategy in the visibility task. To this end, we varied the number of occluders from trial to trial. This change substantially reduced the correlation between the avatar's position and which objects were visible, rendering a memory-based strategy unreliable.

Our second and primary goal was to explore whether or not the line-of-sight computation involved in the visibility task engaged an analogue process. If this were the case, the time to compute the line of sight would increase with the distance between the target and the avatar (i.e., with the length of the line). To test this hypothesis, we manipulated the distance between the avatar and the target object. We did not have strong predictions for the effect of this manipulation in the left–right task. Previous studies found that the distance between two objects did not affect RT when participants judged whether or not an arrangement of letters fit a particular spatial configuration (Carlson &

Logan, 2001). However, in those studies all distances were relatively small, and the task configuration was quite different from that used in the present study.

Finally, the paradigm was slightly modified to reduce the contribution of visual search processes to task performance. In Experiment 1, the participants had to look for the target object once they had read its name. To make sure that bias in visual search processes, such as the tendency to scan arrays of objects from left to right and from top to bottom (Abed, 1991; Nachshon, Shefler, & Samocha, 1977), did not interfere with the distance effects tested here, the target was indicated by a yellow square rather than by its name. We assumed that attention would be automatically captured by the yellow square. The contrast between the light color of the square and the homogeneous blue background formed by the table was expected to create a pop-out effect (Baldassi & Burr, 2004), reducing voluntary visual search processes.

Method

Participants. Twenty-four Washington University students (8 female, mean age = 19.25 years) participated in the experiment in return for course credit.

Stimuli. The stimuli were the same as in Experiment 1, with three exceptions: (1) The number of occluders varied from two to four, (2) the angular disparity between the participant's position and the doll avatar's position was 0º, 90º, 180º, or 270º; and (3) a yellow square was drawn around the target object instead of the object's name being presented at the top of the screen (Figure 1B). In the two-occluders condition, the bottom left and top right occluders were removed. In the three-occluders condition, the top left occluder was removed. We defined *far* objects as the four objects more than halfway across the table from the avatar (two of these objects were 26 in. away from the doll, and two were 27.5 in. away) and *near* objects as those less than halfway across the table from the avatar (two of these objects were 8 in. away from the doll, and two were 17 in. away).

Design and Procedure. The procedure was similar to that of Experiment 1. In both tasks, half of the target objects were far from and the other half were near the avatar. Each task involved 96 trials (left–right task, 3 occluder conditions × 4 views × 4 angles × 2 conditions [left vs. right]; visibility task, 3 occluder conditions × 4 views × 4 angles × 2 conditions [visible vs. occluded]). These trials were repeated twice, for a total of 192 trials per task.

Results

The error rate was low in both the visibility task (M = 4.24%, SD = 6.76%) and the left–right task (M = 3.24%, SD = 6.92%). Correct RTs were trimmed as described earlier; on average, 3.64% of the RTs in the left–right task and 6.36% in the visibility task were deleted. Trimmed correct RTs were analyzed by computing means for each participant for each condition and subjecting these to a 2 (task) × 2 (distance between the avatar and the target object) × 3 (angular disparity between the avatar and the participant) repeated measures ANOVA.

RTs were similar in the visibility task (M = 771 msec, SD = 173) and the left–right task (M = 749 msec, SD = 196) [$F(1,23) < 1$]. They were shorter for near objects (M = 733 msec, SD = 173) than for far objects (M = 787 msec, SD = 193) [$F(1,23) = 67.75$, $p < .001$]. Overall, RTs increased with increasing angular disparity [$F(2,46) = 37.35$, $p < .001$].

The effect of angular disparity differed between the two tasks [$F(2,46) = 50.79$, $p < .001$]. In contrast, the effect of distance between the avatar and the target object was similar in both tasks [$F(1,23) = 2.17$, $p = .15$]. Both of these interactions are illustrated in Figure 3. No other interactions were significant (highest $F = 0.34$).

To clarify the differences between the tasks, follow-up ANOVAs were performed for each task, with angular disparity and distance as within-subjects factors. In the left–right task, RTs increased with increasing angular disparity [$F(2,46) = 50.34$, $p < .001$]. A planned contrast indicated a significant linear trend [$t(23) = 8.12$, $p < .001$]. Near objects were responded to faster than far objects [mean difference, 44 msec; $F(1,23) = 48.12$, $p < .001$]. The two-way interaction was not significant [$F(2,46) < 1$]. In the visibility task, RTs were shorter for near objects than for far objects [mean difference, 64 msec; $F(1,23) = 31.82$, $p < .001$] and varied slightly but significantly with angular disparity [$F(2,46) = 5.00$, $p < .02$]. This variation did not correspond to a linear increase [$t(23) = 1.59$, $p = .12$]. There was a significant increase between 0º and 90º [$t(23) = 3.53$, $p < .01$] and then a nonsignificant de-

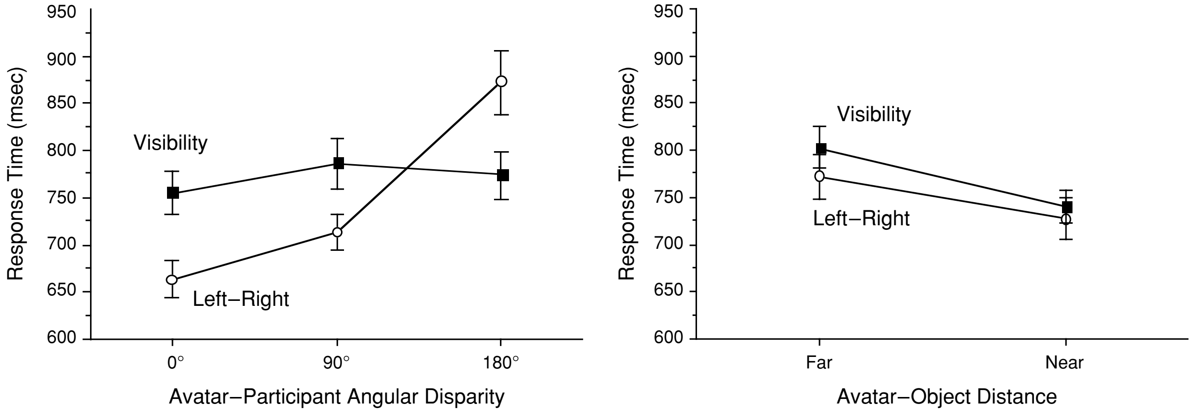

Figure 3. Experiment 2: Response times (in milliseconds) in the left–right and visibility tasks as a function of the angular disparity between the avatar's and the participant's positions (left panel) and the distance between the avatar and the target object (right panel). Vertical bars represent standard errors of the means.

crease between 90º and 180º [$t(23) < 1$]. The two-way interaction was not significant [$F(2,46) < 1$].

Discussion

The results of Experiment 2 replicated and extended those of Experiment 1. In the left–right task, RTs again increased with angular disparity between the avatar and the participant, suggesting the use of perspective transformations. In the visibility task, RTs were again only weakly dependent on angular disparity, and this dependence was not well fit by a linear contrast. These results support the view that the participants were relying more on perspective transformations for left–right judgments and more on line-of-sight computations for visibility judgments. In Experiment 2, only 3 of the 24 participants reported using memory-based strategies, suggesting that these differences cannot be accounted for by the use of such strategies.

The distance between the target object and the avatar influenced performance in the visibility task. This supports the hypothesis that computing a line of sight is the imaginary analogue of actually drawing a line between another agent and an object (Yaniv & Shatz, 1990). The distance between the avatar and the target object also influenced performance in the left–right task. This is notable, since previous studies had found no effect of the distance between two objects when participants judged whether one object was below or above the other (Carlson & Logan, 2001; Logan & Compton, 1996). However, a recent study in which larger distances were used did find distance effects in a task in which left–right judgments were made (Carlson & van Deman, 2004). Carlson and van Deman proposed that such a distance effect is caused by attention moving from one object to the other: The farther away the objects, the longer it takes. An alternative explanation of the distance effect found in the present study is that left–right judgments about near objects were faster because near objects were farther from the avatar's midline, in angular distance, than far objects, which may have made them more typical instances of the spatial categories *left of* and *right of*. Previous studies suggest a similar effect for the spatial categories *above* and *below* (Logan & Sadler, 1996). To test this hypothesis, we performed a post hoc ANOVA for each task with angular distance between the target object and the avatar's midline as a factor. The angular distance from midline was defined as either "near" (i.e., objects that were either 7º or 25º away from midline) or "far" (i.e., objects that were either 30.5º or 40º away from midline). Results supported the hypothesis: Responses to objects far from the avatar's midline were faster than responses to objects close to midline in the left–right task [$t(23) = 4.89$, $p < .001$] but not in the visibility task [$t(23) = 1.08$, $p = .28$]. (See also Experiment 4, below.)

EXPERIMENT 3

Is the presence of an external agent necessary for different kinds of VPT? The results of Experiments 1 and 2 suggest that when an agent is shown, participants either transform their perspective or compute the line of sight of the avatar, depending on the task. Previous studies suggest that people can perform perspective transformations without having a physical avatar present (Amorim, Glasauer, Corpinot, & Berthoz, 1997; Amorim & Stucchi, 1997; Amorim, Trumbore, & Chogyen, 2000; Creem et al., 2001; Juurmaa & Lehtinen-Railo, 1994; Presson & Montello, 1994; Wraga et al., 2000; Zacks, Mires, Tversky, & Hazeltine, 2000; Zacks et al., 2003). However, it is less clear that people can easily trace a line of sight without perceiving an agent. Kelly et al. (2004) have shown that the difficulty in predicting what another agent can see increases as the distance between oneself and the agent increases, suggesting that easy line-of-sight computation may be limited to certain situations. For example, imagine you are standing in one of the corners of a small town plaza. If you were asked whether Sally, who is standing at another corner of the plaza, can see the bench next to the church, it seems intuitive that you would try to trace her line of sight. However, if you are asked whether it is possible for a person to see the bench from that corner when no one is currently standing there, you might imagine yourself at that corner rather than try to trace the line of sight of an imaginary person whose precise position you don't really know and whose face you cannot see. If it is indeed difficult to trace a line of sight from a symbolically cued location, one would expect that when participants performed the visibility task without an anthropoid agent as the cue to the target location, response latency would reflect the use of perspective transformation—that is, response latency would be sensitive to the angular disparity between the participant's position and the to-be-imagined position. To test this, in this experiment we replicated Experiment 2 but replaced the anthropoid doll with an abstract symbol.

Method

Participants. Twenty-two Washington University students (5 male, mean age = 19.27 years) participated in the experiment as a course credit requirement.

Stimuli. The stimuli were the same as those in Experiment 2, except that the doll was replaced by a yellow asterisk (Figure 1C).

Design and Procedure. The design and procedure were identical to those used in Experiment 2. Only the instructions were slightly modified, to "Would an observer be able to see the target object from the asterisk's location?" and "Would the target object be on the observer's left side if the observer were at the asterisk's location?" An example of the display with the doll was shown before the beginning of each task to give the participant an idea of the position of the absent observer.

Results

The error rate was low in both the visibility task ($M = 7.7\%$, $SD = 9.97\%$) and the left–right task ($M = 4.25\%$, $SD = 7.44\%$). Correct RTs were trimmed as described previously; on average, 4.38% of the RTs in the left–right task and 8.71% of those in the visibility task were deleted. Trimmed correct RTs were analyzed by computing means for each participant for each condition and subjecting these to a $2 \times 2 \times 3$ ANOVA, with task, distance between the asterisk and the target object, and angular disparity between the asterisk and the participant as within-subjects factors.

RTs were longer in the left–right task ($M = 927$ msec, $SD = 263$) than in the visibility task ($M = 817$ msec, $SD = 188$) [$F(1,21) = 5.71, p < .03$]. They were shorter for near ($M = 825$ msec, $SD = 197$) than for far ($M = 919$ msec, $SD = 259$) objects [$F(1,21) = 44.94, p < .001$]. Overall, they increased with angular disparity [$F(2,42) = 56.94, p < .001$]. The effect of angular disparity on RT varied depending on the distance between the asterisk and the target object [$F(2,42) = 4.08, p < .03$].

RT increased with increasing angular disparity for the left–right task but not for the visibility task, leading to a significant two-way interaction [$F(2,42) = 46.39, p < .001$]. The effect of the distance between the asterisk and the target object was similar in the two tasks but larger in the left–right than in the visibility task [$F(1,21) = 5.47, p < .03$]. Both of these interactions are illustrated in Figure 4. The three-way interaction was not significant [$F(2,42) = 2.52, p = .09$].

Further ANOVAs were performed for each task with angular disparity and distance as within-subjects factors. In the left–right task, RTs increased with increasing angular disparity [$F(2,42) = 56.98, p < .001$]. A planned contrast indicated a significant linear trend [$t(21) = 8.80, p < .001$]. RTs were shorter for near than for far objects [mean difference, 122 msec; $F(1,21) = 26.14, p < .001$]. The two-way interaction was significant [$F(2,42) = 3.77, p < .04$]. In the visibility task, near objects led to shorter RTs than far objects [mean difference, 66 msec; $F(1,21) = 40.09, p < .001$]. There was no main effect of angular disparity on RT [$F(2,42) < 1$]. There was no difference between 0º and 90º or between 90º and 180º (highest t value $= 0.42$). The two-way interaction was not significant [$F(2,42) < 1$].

Discussion

In short, the data from Experiment 3 replicated the substantive features of Experiment 2: Latencies increased with increasing angular disparity in the left–right task but not in the visibility task. This suggests that these patterns do not depend on the presence of a human-like avatar.

This was the only experiment in which the overall RT varied significantly between the two tasks: The participants were faster in the visibility task than in the left–right task. The comparison of Experiments 2 and 3 shows that the cost of the absence of the avatar is larger in the left–right task (in which the participants were 157 msec slower in the absence of the avatar than in its presence) than in the visibility task (68-msec difference). However, the error rate increase caused by the absence of the avatar was larger in the visibility task (4.5%) than in the left–right task (1%), so a speed–accuracy trade-off cannot be ruled out.

As in Experiment 2, the distance between the avatar and the target object influenced performance in both tasks. However, in Experiment 3 this effect was larger in the left–right task than in the visibility task. This may reflect in part the overall slower performance in the left–right task. In Experiment 2, the two tasks did not differ in terms of overall latencies and the distance effect was similar in both tasks. In Experiment 3, given that latencies were longer in the left–right task than in the visibility task, a scaling effect may have amplified the distance effect in the left–right task.

As was noted in the Discussion of Experiment 2, the finding that left–right judgments about near objects were faster than those about far objects was not predicted but could be due to the fact that near objects were, on average, further than far objects from the avatar's midline and, hence, better instances of the categories *left* and *right*. As for Experiment 2, we performed a post hoc ANOVA with the angular distance between the target object and the avatar's midline as a factor, and the results supported the hypothesis. Responses to objects far from the avatar's midline were faster than responses to objects close to midline in the left–right task only [$t(21) = 4.10, p < .001$]. Surprisingly, in the visibility task, the opposite effect was observed [$t(21) = 2.42, p < .02$]. In the visibility task,

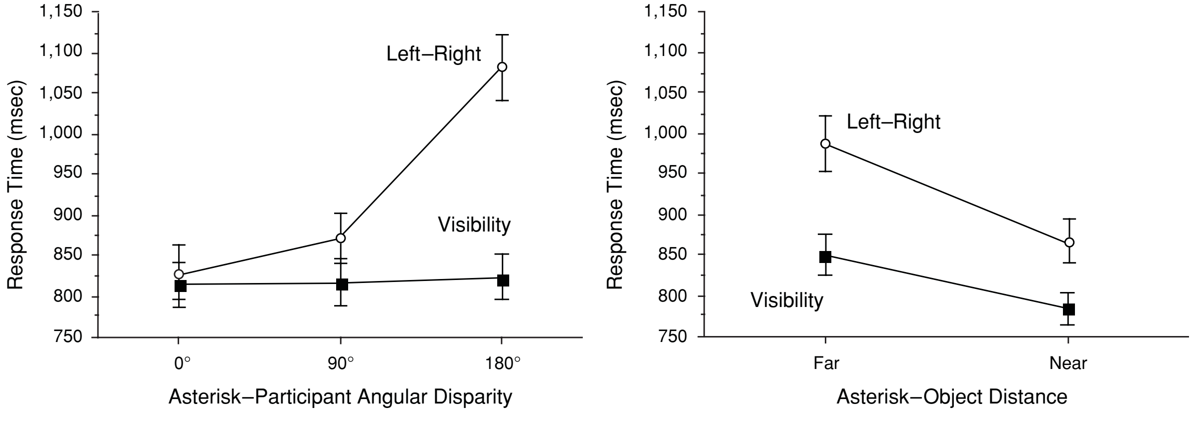

Figure 4. Experiment 3: Response times (in milliseconds) in the left–right and visibility tasks as a function of the angular disparity between the asterisk's and the participant's positions (left panel) and the distance between the asterisk and the target object (right panel). Vertical bars represent standard errors of the means.

better performance for objects close to midline than for objects far from midline may reflect the fact that visual acuity is generally higher for locations on the vertical axis than for those on oblique axes (Heeley, Buchanan-Smith, Cromwell, & Wright, 1997). As a consequence, judgments based on tracing a line of sight that departs from vertical orientation may be more difficult than judgments based on tracing a line of sight close to a vertical orientation (i.e., close to the avatar's midline).

Thus, post hoc analyses of the object–avatar distance effects in the left–right task in Experiments 2 and 3 are consistent with the hypothesis that participants categorize better instances of *left* and *right* faster than less prototypical category members. (As was noted previously, these effects also could arise due to attention-shifting costs.) If this account is correct, then it should be possible to eliminate distance effects in left–right judgments by controlling how good a category member each left and right location is, which is largely a question of angular distance from midline (Logan & Sadler, 1996). We therefore hypothesized that controlling the angular distance of the objects from the avatar's midline would eliminate effects of the distance between the object and the avatar in the left–right task but not in the visibility task. Experiment 4 was designed to test this hypothesis.

EXPERIMENT 4

The goal of this experiment was to tease apart the effects of the distance between the avatar and the target object and the effects of the angular distance between the target object and the avatar's midline. In Experiments 2 and 3, these two factors were confounded because objects near the avatar were also further in angular distance from the avatar's midline than were objects far from the avatar. In Experiment 4, we manipulated these two factors directly in both the left–right and visibility tasks. We hypothesized that the avatar–object distance would affect only line-of-sight computation, and thus should affect RT only in the visibility task. As in the previous experiments, we also manipulated the angular disparity between the avatar's position and the participant's position. We expected to replicate the previous finding that RT would increase with increasing angular disparity between the avatar and the participant only in the left–right task. Thus, we expected that in this modified design we would observe a double dissociation: RT would increase with distance between the avatar and the object more in the visibility task than in the left–right task, whereas RT would increase with angular distance between the avatar and the participant more in the left–right task than in the same–different task.

Method

Participants. Twenty-two Washington University students (5 male, mean age = 19.8 years) participated in the experiment as a course credit requirement.

Stimuli. The participants made judgments about color photographs of a square table on which 16 toy animals and six occluders were positioned (Figure 1D). The avatar was represented by a yellow asterisk that was positioned 3 in. away from the table. Half of the objects were close to the asterisk (8.125 or 11 in.), and the other half were far from the asterisk (22 or 24.75 in.). Half of the objects close to the asterisk were close to the avatar's midline (20º of angular disparity between the avatar's midline and the location of the object), and the other half were far from midline (32º). The same was true for objects far from the asterisk. The angular distance between the asterisk's position and the participant's position was 0º, 90º, 180º, or 270º. Target objects were outlined in yellow.

For each asterisk position, each animal could be the target object and could be either occluded or visible. As a consequence, a total of 128 photographs of the display was used. For each photograph, the positions of the occluders were randomly chosen with the constraint that the target object be either visible or occluded, depending on the condition.

Design and Procedure. The visibility task and the left–right task used in Experiment 3 were used again, with the following differences. First, the display was different (see above). Second, the number of trials was different. There were 128 unique stimuli in the visibility task (4 angles × 2 conditions [visible vs. occluded] × 16 animals). Given that only visible objects were tested in the left–right task, there were only 64 unique stimuli in that task (4 angles × 16 animals). Stimuli were repeated so that each task included a total of 256 trials. Finally, in each task half of the target objects were far from the asterisk and half were near the asterisk, and half of the target objects were far from the avatar's midline and half were near the avatar's midline.

Results

The error rate was low in both the visibility task (*M* = 4.19%, *SD* = 3.36%) and the left–right task (*M* = 2.46%, *SD* = 2.01%). Correct RTs were trimmed as was explained previously; on average, 2.46% of the RTs in the left–right task and 4.19% in the visibility task were deleted. Correct trimmed RTs were analyzed by computing means for each participant for each condition and submitting these to a 2 × 2 × 2 × 3 ANOVA, with task, distance between the avatar/asterisk and the target object, angular distance between the midline of the avatar/asterisk and the target object, and angular disparity between the avatar/asterisk and the participant as within-subjects factors.

RTs were longer in the visibility task (*M* = 1,052 msec, *SD* = 278) than in left–right task (*M* = 862 msec, *SD* = 181) [$F(1,23) = 27.35, p < .001$]. Overall, they increased with the angular disparity between the avatar and the participant [$F(2,46) = 97.50, p < .001$]. Trials in which objects were far from the avatar took longer (*M* = 973 msec, *SD* = 263) than did trials in which objects were close to the avatar (*M* = 941 msec, *SD* = 242) [$F(1,23) = 13.62, p < .01$]. There was no main effect of angular midline distance [$F(1,23) = 3.07, p = .09$].

As can be seen in Figure 5, angular disparity between the avatar and the participant had a larger effect on RTs in the left–right task than in the visibility task [$F(2,46) = 63.41, p < .001$]. Conversely, distance between the avatar and the target object had a larger effect on RTs in the visibility task than in the left–right task [$F(1,23) = 7.16, p < .03$]. As is shown in Figure 6, the effect of the angular midline distance also differed between the two tasks [$F(1,23) = 92.82, p < .001$]. In the visibility task, objects near midline were responded to more quickly, whereas in the left–right task, objects far from midline were re-

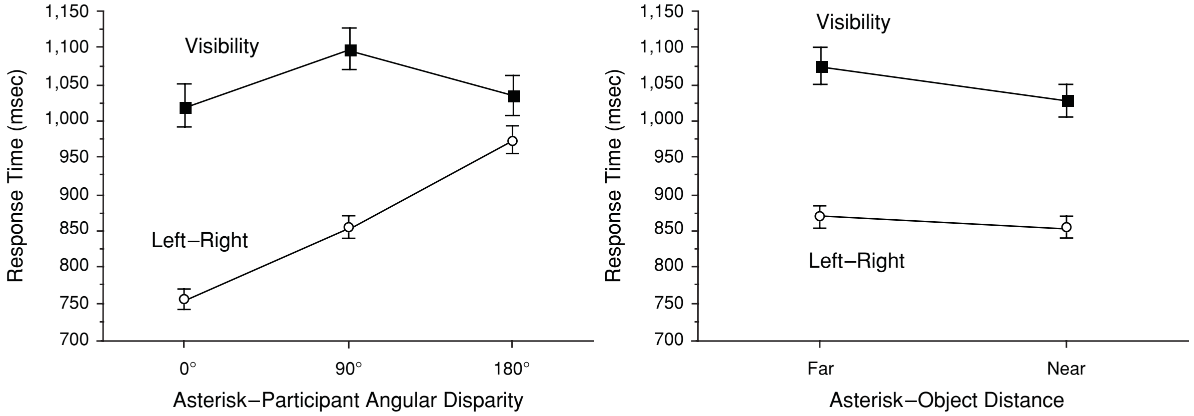

Figure 5. Experiment 4: Response times (in milliseconds) in the left–right and visibility tasks as a function of the angular disparity between the asterisk's and the participant's positions (left panel) and the distance between the asterisk and the target object (right panel). Vertical bars represent standard errors of the means.

sponded to more quickly. This pattern replicates that of Experiment 3.

Several other interactions were significant (see Table 1 for a summary of the relevant means), including angular disparity × angular midline distance $[F(2,46) = 8.08, p < .01]$, task × angular midline distance × avatar–object distance $[F(1,23) = 28.66, p < .001]$, task × angular disparity × angular midline distance $[F(2,46) = 5.51, p < .01]$, and angular disparity × angular midline distance × avatar–object distance $[F(2,46) = 4.67, p < .02]$. None of these interactions were of theoretical interest, so they were not pursued.

Further ANOVAs were performed for each task with angular disparity, angular midline distance, and avatar–object distance as within-subjects factors. In the left–right task, RTs increased with increasing angular disparity $[F(2,46) = 109.43, p < .001]$. A planned contrast indicated a significant linear trend $[t(23) = 12.58, p < .001]$. Objects that were far from midline were responded to faster than objects close to midline $[F(1,23) = 35.38, p < .001]$. The main effect of the avatar–object distance approached but did not reach statistical significance $[F(1,23) = 4.40, p = .05]$. Finally, the midline distance × avatar–object distance interaction was significant $[F(2,46) = 17.59, p < .001]$.

In the visibility task, RTs varied with angular disparity $[F(2,46) = 26.31, p < .001]$. However, this increase was not linear $[t(23) = 1.01, p = .32]$. Additional *t* tests showed that RTs increased from 0º to 90º $[t(23) = -5.79, p < .001]$ but then decreased from 90º to 180º $[t(23) = 7.71, p < .001]$. Objects that were close to midline were responded to faster than objects far from midline $[F(1,23) = 69.72, p < .001]$, in contrast to what was found in the left–right task. Objects that were close to the

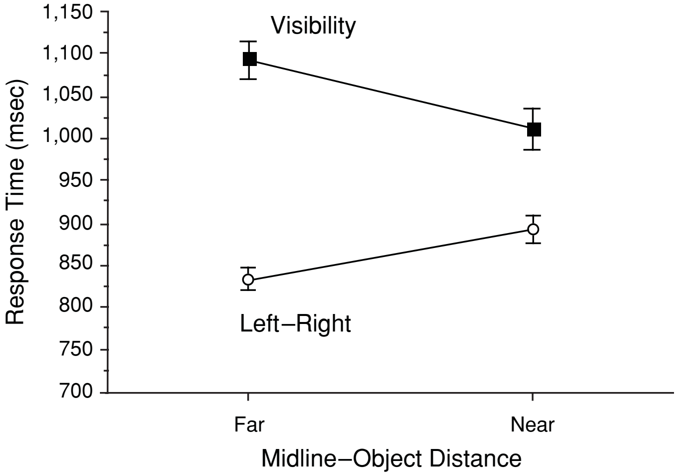

Figure 6. Experiment 4: Response times (in milliseconds) in the left–right and visibility tasks as a function of the angular distance between the avatar's or the asterisk's midline and the target object. Vertical bars represent standard errors of the means.

Table 1
Experiment 4: Mean Response Times (RTs, in Milliseconds) and Standard Deviations (*SD*s) in the Left–Right and Visibility Tasks as a Function of Angular Disparity Between the Avatar's and the Participant's Positions, the Angular Distance Between the Avatar's Midline and the Target Object (Far From Midline vs. Near Midline), and the Distance Between the Avatar and the Target Object (Distant vs. Near)

	Left–Right Task								Visibility Task							
	Far From Midline				Near Midline				Far From Midline				Near Midline			
	Distant		Near		Distant		Near		Distant		Near		Distant		Near	
Angular Disparity	RT	*SD*	RT	*SD*	RT	*SD*	RT	*SD*	RT	*SD*	RT	*SD*	RT	*SD*	RT	*SD*
0º	749	112	705	98	787	148	783	133	1,059	269	1,041	275	1,015	301	969	316
90º	832	146	803	112	883	177	902	154	1,119	285	1,131	289	1,135	297	1,012	213
180º	979	180	921	154	987	187	1,009	226	1,124	307	1,091	240	1,003	279	927	206

avatar were responded to faster than objects that were far from the avatar [$F(1,23) = 13.83$, $p < .01$]. The angular disparity × midline distance [$F(2,46) = 9.70$, $p < .001$], midline distance × avatar distance [$F(1,23) = 10.53$, $p < .01$], and angular disparity × midline distance × avatar distance [$F(2,46) = 4.29$, $p < .02$] interactions were also significant.

Discussion

As in Experiment 2, results showed that increasing angular disparity between the participant's position and the avatar's position led to increasing latencies in the left–right task, whereas in the visibility task latencies were only weakly dependent on angular disparity and this relationship was not linear. The distance between the target object and the avatar influenced performance more strongly in the visibility task than in the left–right task.

In contrast to the results of Experiments 2 and 3, the distance between the target object and the avatar did not strongly influence performance in the left–right task (although there was a nonsignificant trend for RT to increase with increasing distance). This supports the proposal that the effects of distance between the target object and the avatar in Experiments 2 and 3 reflected a confound between avatar–object distance and angular distance from midline. This explanation received further support from the finding that objects at a great angular distance from midline were responded to faster than objects closer to midline in the left–right task only. This suggests that objects far from the avatar's midline were easier to categorize as *left* or *right* because they were clear instances of those categories.

An effect of the angular midline distance was also observed in the visibility task. Replicating the results of the post hoc analysis of Experiment 3, it was found that objects closer to midline were responded to faster than objects farther from midline—the opposite of what was found in the left–right task. As we proposed earlier, this may be caused by the fact that visual acuity is higher for locations on the vertical axis than for those on oblique axes.

GENERAL DISCUSSION

The data clearly suggest that different VPT processes are used depending on the task and goals at hand. When the participants had to report the location of an object relative to another person, RTs increased as the angle between the participants' current position and that of the other person increased, which is consistent with performance of a spatial transformation that would map their egocentric reference frame onto the other's position. When the participants had to report whether or not another person could see an object, the relationship between angular disparity and RT was weak or nonexistent, in consistency with the act of tracing a line of sight from the other person to the object.

Transforming one's egocentric reference frame seems to be an efficient process in most situations in which one wonders about the relative location or appearance of objects from another person's viewpoint (Amorim & Stucchi, 1997; Creem et al., 2001; Huttenlocher & Presson, 1979; Presson & Montello, 1994; Wraga et al., 2000; Zacks et al., 2003). This process allows one to read off of the location/appearance of objects in one's transformed egocentric reference frame. The present results suggest that a different process is engaged when one wonders whether or not another person can see an object of interest. This process includes computing the line of sight of the other person. Line-of-sight tracing may be preferable to transformation processes in such situations, because it does not require updating of any spatial reference frame and does not engender a conflict between one's own reference frame and the imagined reference frame.

The computation of the line of sight of an agent may be an analogue process similar to mental scanning. As would be the case if an actual line were drawn between the avatar and the target object, RTs increased with the distance between the two. The distance between the avatar and the target object also affected performance in the left–right task in Experiments 2 and 3, but, as was shown in Experiment 4, this can be attributed to angular disparity effects between the avatar's midline and the target object.

The adoption of different processes for judgments of relative location and visibility does not appear to depend on the presence of a human-like avatar. This finding suggests that the presence of the head of an agent is not a necessary condition for the computation of a line of sight. In natural situations, computing the line of sight in the absence of an agent, as in selecting a place to hide from an approaching predator, may be critical.

Neuroimaging evidence suggests that both common and specialized mechanisms subserve different types of perspective taking: imagining what it's like to perform an action in another's position, what another person knows, and how another person feels (Brunet, Sarfati, Hardy-Bayle, & Decety, 2000; Decety & Chaminade, 2003; Ruby & Decety, 2001; Sebanz & Frith, 2004; Vogeley et al., 2004). The results reported here argue for the presence of different mechanisms *within* purely spatial perspective taking. Some mechanisms allow us to transform our perspective to answer general questions about the visuospatial aspects of a scene viewed from a different viewpoint. Others allow us to compute the line of sight of a third agent. The present study suggests that this computation is used in determining whether or not another agent can see an object of interest. Future research may investigate whether a similar line-of-sight computation is used to answer another crucial question for survival: Can this agent see me?

REFERENCES

Abed, F. (1991). Cultural influences on visual scanning patterns. *Journal of Cross-Cultural Psychology*, **22**, 525-534.

Amorim, M.-A. (2003). "What is my avatar seeing?": The coordination of "out-of-body" and "embodied" perspectives for scene recognition across views. *Visual Cognition*, **10**, 157-199.

Amorim, M.-A., Glasauer, S., Corpinot, K., & Berthoz, A. (1997). Updating an object's orientation and location during nonvisual navigation: A comparison between two processing modes. *Perception & Psychophysics*, **59**, 404-418.

Amorim, M.-A., & Stucchi, N. (1997). Viewer- and object-centered mental explorations of an imagined environment are not equivalent. *Cognitive Brain Research*, **5**, 229-239.

Amorim, M.-A., Trumbore, B., & Chogyen, P. L. (2000). Cognitive repositioning inside a desktop VE: The constraints introduced by first- versus third-person imagery and mental representation richness. *Presence: Teleoperators & Virtual Environments*, **9**, 165-186.

Baldassi, S., & Burr, D. C. (2004). "Pop-out" of targets modulated in luminance or colour: The effect of intrinsic and extrinsic uncertainty. *Vision Research*, **44**, 1227-1233.

Brunet, E., Sarfati, Y., Hardy-Bayle, M. C., & Decety, J. (2000). A PET investigation of the attribution of intentions with a nonverbal task. *NeuroImage*, **11**, 157-166.

Carlson, L. A., & Logan, G. D. (2001). Using spatial terms to select an object. *Memory & Cognition*, **29**, 883-892.

Carlson, L. A., & van Deman, S. R. (2004). The space in spatial language. *Journal of Memory & Language*, **51**, 418-436.

Creem, S. H., Downs, T. H., Wraga, M., Proffitt, D. R., & Downs, J. H., III (2001). An fMRI study of imagined self-rotation. *Cognitive, Affective, & Behavioral Neuroscience*, **1**, 239-249.

Decety, J., & Chaminade, T. (2003). Neural correlates of feeling sympathy. *Neuropsychologia*, **41**, 127-138.

Denis, M., Gonçalves, M.-R., & Memmi, D. (1995). Mental scanning of visual images generated from verbal descriptions: Towards a model of image accuracy. *Neuropsychologia*, **33**, 1511-1530.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1–Level 2 distinction. *Developmental Psychology*, **17**, 99-103.

Heeley, D. W., Buchanan-Smith, H. M., Cromwell, J. A., & Wright, J. S. (1997). The oblique effect in orientation acuity. *Vision Research*, **37**, 235-242.

Huttenlocher, J., & Presson, C. C. (1979). The coding and transformation of spatial information. *Cognitive Psychology*, **11**, 375-394.

Juurmaa, J., & Lehtinen-Railo, S. (1994). Visual experience and access to spatial knowledge. *Journal of Visual Impairment & Blindness*, **88**, 157-170.

Kelly, J. W., Beall, A. C., & Loomis, J. M. (2004). Perception of shared visual space: Establishing common ground in real and virtual environments. *Presence: Teleoperators & Virtual Environments*, **13**, 442-450.

Kosslyn, S. M. (1973). Scanning visual images: Some structural implications. *Perception & Psychophysics*, **14**, 90-94.

Logan, G. D., & Compton, B. J. (1996). Distance and distraction effects in the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 159-172.

Logan, G. D., & Sadler, D. D. (1996). A computational analysis of the apprehension of spatial relations. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (pp. 493-529). Cambridge, MA: MIT Press.

Nachshon, I., Shefler, G. E., & Samocha, D. (1977). Directional scanning as a function of stimulus characteristics, reading habits, and directional set. *Journal of Cross-Cultural Psychology*, **8**, 83-99.

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

Presson, C. C., & Montello, D. R. (1994). Updating after rotational and translational body movements: Coordinate structure of perspective space. *Perception*, **23**, 1447-1455.

Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: A PET investigation of agency. *Nature Neuroscience*, **4**, 546-550.

Salatas, H., & Flavell, J. H. (1976). Perspective taking: The development of two components of knowledge. *Child Development*, **47**, 103-109.

Sebanz, N., & Frith, C. (2004). Beyond simulation? Neural mechanisms for predicting the actions of others. *Nature Neuroscience*, **7**, 5-6.

Shelton, A. L., & McNamara, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology*, **43**, 274-310.

Simons, D. J., & Wang, R. F. (1998). Perceiving real-world viewpoint changes. *Psychological Science*, **9**, 315-320.

Vogeley, K., May, M., Ritzl, A., Falkai, P., Zilles, K., & Fink, G. R. (2004). Neural correlates of first-person perspective as one constituent of human self-consciousness. *Journal of Cognitive Neuroscience*, **16**, 817-827.

Wang, R. F., & Simons, D. J. (1999). Active and passive scene recognition across views. *Cognition*, **70**, 191-210.

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

Yaniv, I., & Shatz, M. (1990). Heuristics of reasoning and analogy in children's visual perspective taking. *Child Development*, **61**, 1491-1501.

Zacks, J. M., Mires, J., Tversky, B., & Hazeltine, E. (2000). Mental spatial transformations of objects and perspective. *Spatial Cognition & Computation*, **2**, 315-332.

Zacks, J. M., Vettel, J. M., & Michelon, P. (2003). Imagined viewer and object rotations dissociated with event-related fMRI. *Journal of Cognitive Neuroscience*, **15**, 1002-1018.

(Manuscript received September 15, 2004;
revision accepted for publication May 28, 2005.)