Response mode differences in perspective taking: Differences in representation or differences in retrieval?

JONATHAN W. KELLY AND TIMOTHY P. MCNAMARA
Vanderbilt University, Nashville, Tennessee

Three experiments explored whether response mode differences in perspective taking result from different spatial representations or different retrieval processes. Participants learned object locations and then, while blindfolded, pointed to or verbally described object locations from perspectives aligned or misaligned with their facing direction and aligned or misaligned with the learning perspective. Pointing was facilitated from the perspective aligned with the body during testing. Similar facilitation occurred when verbally labeling, but only when conducted in the context of pointing (e.g., after pointing). Without this pointing context, or after third-person strategy instructions, the effect of body alignment was eliminated for verbal responses. Pointing was less responsive to context and strategy. Across all conditions, performance was facilitated for the learning perspective. Taken together, these experiments indicate that response mode differences are due to differences in the retrieval process, which varies with strategy, rather than differences in the organization of the underlying spatial memory.

In daily navigation, we regularly rely on our memories of spaces to guide our movements and decisions. Whether planning a detour to avoid road congestion or describing the location of the bookstore to a campus visitor, we are often required to imagine perspectives that we do not currently occupy. Our ability to imagine these unoccupied perspectives depends on a number of factors, including our representation of the environment in long-term memory and our current location and orientation within that environment (Mou, McNamara, Valiquette, & Rump, 2004).

The ease with which new perspectives can be imagined has been shown to depend greatly on the presence of self-motion cues during the imagined movement. Spatial updating, the process of updating the remembered locations of previously learned objects during self-movement, is typically quite accurate after physical rotations and translations with eyes closed (Philbeck, Loomis, & Beall, 1997), especially in contrast to the relatively poor performance after imagined movement (Rieser, 1989). In a seminal paper on imaginal repositioning, Rieser compared the relative ease with which people could point to updated object locations after imagined, rather than real, rotations and translations. Participants were able to point to objects equally well after imagined and physical translations, but performance after imagined rotations was worse than after physical rotations.

Presson and Montello (1994; see also Easton & Sholl, 1995, and May, 2004) theorized that the difference between imagined rotations and imagined translations was due to a reference frame conflict between imagined and actual perspectives for rotations, but not for translations. Specifically, for participants to correctly point to objects after an imagined rotation, they had to compute the correct response from the imagined perspective and map that response onto their physically occupied perspective. After an imagined translation, no such reference frame conflict occurred, because the imagined perspective was aligned with the participant's actual facing direction; the chosen response from the imagined perspective could therefore be directly executed without any further transformation.

The detrimental effect of imagined rotation on perspective-taking ability has been widely reported and replicated. Interestingly, this effect may be specific to certain response modalities employed in those experiments—namely, body-based responses, such as aiming a joystick or other pointing device (Easton & Sholl, 1995; Kelly, Avraamides, & Loomis, 2007; May, 2004; Mou et al., 2004; Presson & Montello, 1994; Rieser, 1989), or turning one's body to indicate the desired direction (Klatzky, Loomis, Beall, Chance, & Golledge, 1998; Waller, Montello, Richardson, & Hegarty, 2002). When verbal responses (e.g., "front left") are used to indicate object locations after imagined rotation, the difficulties encountered when a body-based response is used can be reduced (Avraamides, Ioannidou, & Kyranidou, 2007; Wraga, 2003) or eliminated altogether (Avraamides, Klatzky, Loomis, & Golledge, 2004; de Vega & Rodrigo, 2001; Wang, 2004).

At least two different possibilities have been proposed to explain the relative ease associated with ver-

J. W. Kelly, jonathan.kelly@vanderbilt.edu

bal responses compared with body-based responses after imagined rotation. As suggested by Wang (2004), unique representations and/or processes may subserve actions (e.g., body-based pointing) versus judgments (e.g., verbal descriptions). Regarding the separate-representations hypothesis, Wang proposed that body-based responses might be based on an orientation-dependent representation, whereas verbal responses could be based on an orientation-independent representation. Although the preponderance of evidence suggests that spatial memories are orientation-dependent (Diwadkar & McNamara, 1997; Kelly et al., 2007; Kelly & McNamara, 2008; McNamara, Rump, & Werner, 2003; Mou & McNamara, 2002; Mou et al., 2004; Roskos-Ewoldsen, McNamara, Shelton, & Carr, 1998; Shelton & McNamara, 1997, 2001; Werner & Schmidt, 1999), these data have generally been gathered using body-based pointing responses. Such studies indicate that spatial memories are organized around one or two primary reference directions, and retrieval of interobject relationships (e.g., "Imagine you are standing at *x* facing *y*, point to *z*," where *x*, *y*, and *z* represent learned object locations) is facilitated along those encoded reference directions, relative to other directions. The reference directions are selected through a combination of egocentric experience and environmental structure (Diwadkar & McNamara, 1997; Kelly & McNamara, 2008; McNamara et al., 2003; Shelton & McNamara, 1997, 2001), and the resulting facilitation for aligned perspectives (compared to misaligned perspectives) has been shown under remote (after removal from the remembered environment; Shelton & McNamara, 2001) and situated (when located within the remembered environment; Kelly et al., 2007; Mou et al., 2004) testing conditions. However, verbal responses have not been used with the explicit purpose of measuring the orientation dependence of long-term spatial memory, so the generality of these conclusions to other response modes is unclear.

Regarding the separate-processes hypothesis, a common representation might underlie both body-based and verbal judgments, but the response mode could modify the manner in which that representation is accessed. Specifically, Avraamides et al. (2007) proposed that body-based responses require an additional step during response computation: Once the correct spatial relationship is inferred, it must be mapped onto body coordinates before the pointing response can be executed. This response mapping from the imagined perspective onto the body in its actual perspective demands cognitive effort. In contrast, verbal responses can be made directly from the imagined perspective, without implication of the body and its egocentric frame of reference. In this way, verbal responses avoid the reference-frame conflict that plagues body-based responses.

Recent work demonstrates the potential to distinguish between orientation dependency in long-term spatial memory and conflicts resulting from body-based responses, two factors critical to assessing the separate-representations and separate-processes hypotheses, respectively. Mou et al. (2004; see also Kelly et al., 2007) had participants learn a spatial layout from a single perspective, a procedure that has

previously been shown to produce orientation-dependent spatial memories with privileged access to spatial relations from imagined perspectives aligned with the learning perspective (e.g., Shelton & McNamara, 2001). After learning, blindfolded participants were asked to imagine different perspectives within the learned layout; these imagined perspectives could be aligned or misaligned with the learning perspective. Additionally, participants physically turned to assume different facing directions prior to performing the perspective-taking task, so that the imagined perspective could also be aligned or misaligned with their actual perspective during retrieval. In this way, Mou et al. found separate evidence of facilitation for imagined perspectives aligned, compared with misaligned, with (1) the reference direction used to organize the spatial memory (established at the learning perspective), and (2) the reference frame of the body in its current orientation.

The experimental design used by Mou et al. (2004) provides a promising tool for uncovering the nature of response mode differences. First, it should prove useful in evaluating the separate-representations hypothesis proposed by Wang (2004). Given prior evidence, pointing responses should be based on an orientation-dependent representation, and the learning conditions in the present experiments have been previously shown to produce orientation-dependent spatial memories with a preferred reference direction parallel to the learning perspective. Similar to the findings of Mou et al., pointing responses should be facilitated when imagining perspectives aligned, compared with misaligned, with this reference direction in long-term memory. If verbal responses are based on an orientation-independent representation, there should be no benefit for perspectives aligned with the learning perspective during a verbal response. The same predictions should hold, whether participants are oriented to their surrounds (i.e., when they are aware of their position and orientation within the immediately surrounding environment) or disoriented. Second, the Mou et al. design can be employed to evaluate the separate-processes hypothesis. Based on prior evidence (Kelly et al., 2007; May, 2004; Mou et al., 2004; Presson & Montello, 1994; Rieser, 1989), pointing responses should also be facilitated when the participant imagines perspectives aligned, compared with misaligned, with the body in its current orientation. If verbal responses activate a different retrieval process that does not rely on the body's actual orientation, there should be no benefit for perspectives aligned with the body under verbal response conditions. After disorientation, performance should be similar on perspectives aligned and misaligned with the body, for both pointing and verbal responses.

EXPERIMENT 1

Method

Participants. Sixteen undergraduate students (8 males) at Vanderbilt University participated in exchange for course credit.

Stimuli and Design. The experiment was conducted in a 5×7 m room (see Figure 1) containing the test objects and a laptop computer used to present the experimental trials and collect responses. Each object set consisted of eight objects evenly spaced around a circle (3 m in diameter), centered in the middle of the room. Four
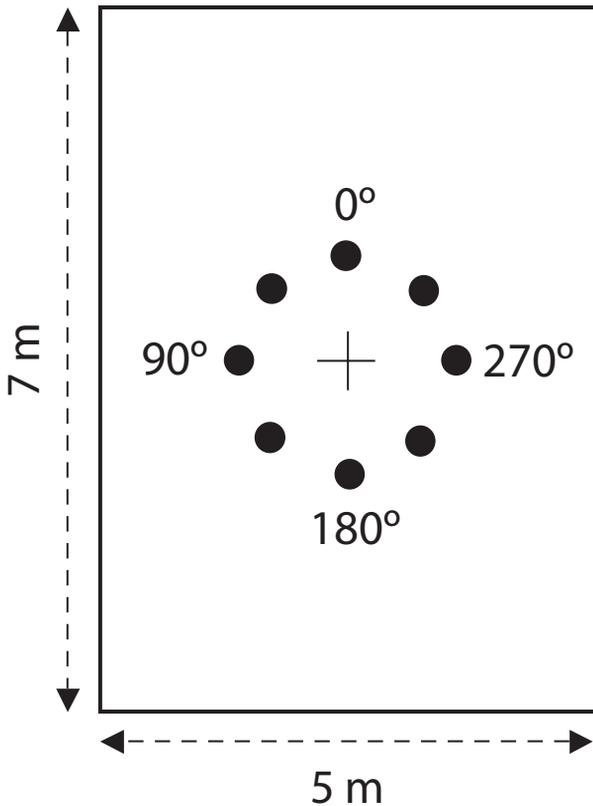
Figure 1. Plan view of the stimuli and room environment used in Experiments 1–3. Circles represent object locations and the cross represents the participant's location.

such sets of objects were learned over the course of the experiment, and each set was chosen from a different semantic category.

Participants faced 0º (see Figure 1) during learning, and faced either 90º or 270º during blindfolded testing. They arrived at 90º or 270º directly or indirectly, depending on the disorientation condition (explained below). After participants were positioned for testing, each subsequent trial consisted of two objects selected from the eight-object array, and required participants to locate one object as if they were facing a second object (e.g., "Face the pear, find the apple"). Trials were presented via wireless headphones (HDR-130 from Sennheiser, Old Lyme, CT).

The three primary independent variables were response mode, disorientation, and imagined perspective. All three variables were manipulated within participants. Response mode was either verbal labeling or pointing, and participants either remained oriented or were disoriented prior to testing. The imagined perspective was either aligned with the learning perspective (0º, termed the "learning" perspective), aligned with the participant's facing direction at test (90º when facing 90º, or 270º when facing 270º, termed the "body-aligned" perspective), or 180º misaligned with the participant's facing direction at test (e.g., 90º when facing 270º, termed the "misaligned" perspective).

Factorial combinations of response mode and disorientation were blocked and block order was counterbalanced using a balanced Latin-square design. A new set of objects was learned before each block. Within each block of trials, imagined perspective was pseudorandomized so that the same object never appeared on two consecutive trials, whether as an orienting object or as a target object. Additionally, a single object never served as both the orienting and target object for the same trial. Three imagined perspectives were combined with seven pointing directions, resulting in 21 trials per block. The dependent measures, defined in detail below, were decision time, response time, and absolute angular error. Data were recorded on a laptop computer, using Vizard software (WorldViz, Santa Barbara, CA).

Procedure. After providing informed consent, participants were presented with a sample set of eight objects for training purposes. Prior to training, participants were outfitted with a wireless joystick (Freedom 2.4 by Logitech, Freemont, CA) affixed to a small table, which was suspended in front of their waists, using shoulder straps. In this way, the joystick was always in front of the participant. During training, the experimenter explained the perspective taking task and described the two response modes. Participants were told to imagine facing one object and then to decide where a second object would be from that imagined perspective. They were shown a button on the joystick, which they were told to press just prior to making their responses, regardless of the response mode. They were instructed to press the button only when they were ready to respond, and to then immediately give their response. For the pointing response, they were told to deflect the joystick in the direction of the target object from the imagined perspective. For the verbal response, they were told to verbally describe the direction of the target object from the imagined perspective, using one of eight response options: front, front right, right, back right, back, back left, left, or front left. These verbal labels were chosen because they provided sufficient precision to perform the task. Participants were shown how these verbal labels matched up with the regularly spaced object array. Participants then performed four practice trials for both response modes using the training objects, which were perceptually available at all times during training. Any errors during practice were corrected.

After training with the pointing and verbal response procedures, participants were escorted to the experiment room on a different floor of the same building. Prior to entering the room, participants donned the blindfold and were led directly into the center of the object array, facing 0º. Once positioned, participants removed their blindfolds, and learning began. They were instructed to study the object locations for 60 sec and then, with eyes closed, to point to each object called out in a random order by the experimenter. Participants were instructed not to move their feet during learning, and to rotate at the neck and waist to view all of the objects. The study–test sequence ended when participants were able to successfully point to all object locations.

After learning the layout, participants donned their blindfolds. In the disorientation condition, participants were told to rotate in place for 60 sec. At random times during rotation, the experimenter instructed them to change directions. Participants were told that the experimenter would walk around them as they rotated, so that the experimenter's location would not be a stable orientation cue. After the disorientation procedure, participants were turned to face 90º or 270º. In the other condition, oriented participants turned 90º to their right or left to face 90º or 270º, respectively, and then stood idle for 60 sec without going through the disorientation procedure.

Once participants were positioned, they were told which type of response to use on the ensuing block of trials. Sound files for each trial were prerecorded and presented via wireless headphones. *Decision time* was defined as the time between the termination of the sound file and the participant's buttonpress on the joystick, indicating readiness to respond. *Response time* was defined as the time between the buttonpress and completion of the response. Decision time and response time were measured separately, because it could take longer to produce verbal responses than pointing responses, and this difference in response production could obfuscate the effects of the independent variables. For pointing trials, the response was completed when the joystick was deflected by 30º from vertical. For verbal trials, the response was completed when the experimenter pressed a key corresponding to the participant's verbal response. Decision time was expected to be a more informative measure than response time. After completing a block of trials, oriented participants turned directly to face 0º, whereas disoriented participants underwent the disorientation procedure again before being returned to 0º. This was done to

prevent disoriented participants from receiving feedback about their facing direction during testing, in light of the repeated measures design. After participants were returned to 0º, a new set of objects was laid on the floor and the procedure began again.

Analysis. Facilitation for perspectives aligned, compared with misaligned, with body orientation at test should be revealed by a performance difference between the body-aligned and misaligned perspectives. Additionally, facilitation due to alignment with an orientation-dependent spatial memory should be revealed by a performance difference between the learning and misaligned perspectives. These two indicators will serve as the primary evidence of the processes and representations used when performing the perspective-taking task under verbal and pointing conditions.

Although the joystick measured pointing responses continuously between 0º and 360º, pointing responses were quantized in 45º increments. This was done to make pointing responses more comparable to verbal responses, which were limited to the same 45º intervals. Participants were made aware of the layout's spatial regularity during training, and so presumably would never intend to produce response angles in anything other than multiples of 45º. Absolute pointing error was calculated by computing the absolute value of the difference between the correct responses and participants' quantized pointing responses.

Results

Latency. Decision time (as shown in Figure 2) was analyzed in a 2 (gender) $\times$ 2 (response mode: verbal labeling or pointing) $\times$ 2 (orientation: disoriented or oriented) $\times$ 3 (imagined perspective: learning, body-aligned, or misaligned perspective) mixed-model ANOVA. The analysis revealed a significant main effect of perspective [$F(2,28) = 15.25, p < .001, \eta_p^2 = .52$], qualified by a significant interaction between orientation and perspective [$F(2,28) = 7.55, p = .002, \eta_p^2 = .35$]. The three-way interaction between orientation, response mode, and imagined perspective was not significant [$F(2,28) = 0.64, p = .54, \eta_p^2 = .04$].

To further evaluate a priori hypotheses, performance on the misaligned perspective was compared with performance on the body-aligned and original perspectives for each combination of response mode and orientation. These contrasts indicated that when participants used the pointing response and remained oriented to the environment, performance on the learning and body-aligned perspectives was faster than it was on the misaligned perspective [$F(1,14) = 15.08, p = .002, \eta_p^2 = .52$, and $F(1,14) = 13.58, p = .002, \eta_p^2 = .49$, respectively]. When pointing after disorientation, judgments were faster for the learning perspective than for the misaligned perspective [$F(1,14) = 20.04, p < .001, \eta_p^2 = .59$], but there was no difference between the body-aligned and misaligned perspectives. When participants used the verbal labeling response and remained oriented to the environment, performance on the learning and body-aligned perspectives was faster than it was on the misaligned perspective [$F(1,14) = 7.56, p = .016, \eta_p^2 = .35$, and $F(1,14) = 5.53, p = .034, \eta_p^2 = .28$, respectively]. When verbally labeling after disorientation, performance was better on the learning perspective than on the misaligned perspective [$F(1,14) = 5.93, p = .029, \eta_p^2 = .30$], but there was no difference between the body-aligned and misaligned perspectives.

Response time was also analyzed in a 2 (gender) $\times$ 2 (response mode: verbal labeling or pointing) $\times$ 2 (orienta-
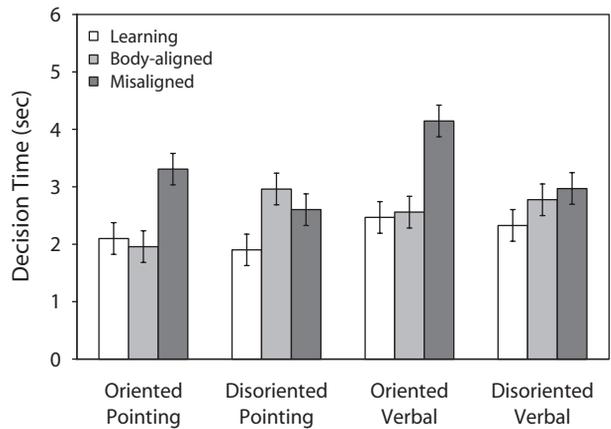

Figure 2. Decision time as a function of response mode, disorientation, and imagined perspective in Experiment 1. Error bars are standard errors estimated from the ANOVA.

tion: disoriented or oriented) $\times$ 3 (imagined perspective: learning, body-aligned, or misaligned perspective) mixed-model ANOVA. Results showed only a main effect of response mode [$F(1,14) = 268.15, p < .001, \eta_p^2 = .95$], with faster responses overall for pointing ($M = 0.53$ sec, $SE = 0.06$) than for verbal labeling ($M = 1.64$ sec, $SE = 0.06$).

Accuracy. Absolute error (Figure 3) was analyzed in a 2 (gender) $\times$ 2 (response mode: verbal labeling or pointing) $\times$ 2 (orientation: disoriented or oriented) $\times$ 3 (imagined perspective: learning, body-aligned, or misaligned perspective) mixed-model ANOVA. Verbal responses ($M = 7.65º, SE = 2.73$) were more accurate overall than were pointing responses [$M = 24.20º, SE = 3.55; F(1,14) = 88.85, p < .001, \eta_p^2 = .86$]. The main effect of perspective [$F(2,28) = 10.07, p = .001, \eta_p^2 = .42$] was qualified by an orientation $\times$ perspective interaction [$F(2,28) = 4.14, p = .027, \eta_p^2 = .23$].

Contrasts conducted to evaluate a priori hypotheses showed that when participants pointed while oriented to the environment, performance was more accurate on the
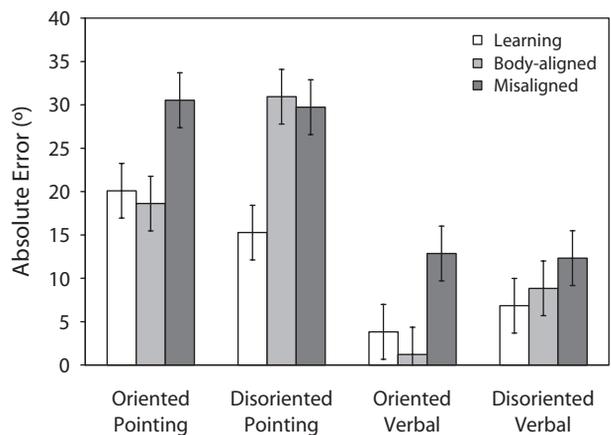

Figure 3. Absolute angular error as a function of response mode, disorientation, and imagined perspective in Experiment 1. Error bars are standard errors estimated from the ANOVA.

learning and body-aligned perspectives than it was on the misaligned perspective [$F(1,14) = 4.37$, $p = .055$, $\eta_p^2 = .24$, and $F(1,14) = 5.94$, $p = .029$, $\eta_p^2 = .30$, respectively]. When pointing after disorientation, performance was better for the learning perspective than for the misaligned perspective [$F(1,14) = 14.40$, $p = .002$, $\eta_p^2 = .51$], but there was no difference between the body-aligned and misaligned perspectives. When participants verbally responded and remained oriented to the environment, performance was more accurate on the learning perspective and the body-aligned perspective than it was on the misaligned perspective [$F(1,14) = 4.78$, $p = .046$, $\eta_p^2 = .25$, and $F(1,14) = 9.76$, $p = .007$, $\eta_p^2 = .41$, respectively]. When verbally labeling after disorientation, performance was better on the learning perspective than on the misaligned perspective [$F(1,14) = 3.47$, $p = .084$, $\eta_p^2 = .20$], but there was no difference between the body-aligned and misaligned perspectives.

Discussion

The separation of decision time and response time proved to be an effective means of accounting for the increased latency associated with producing verbal compared with joystick responses. The fact that response time was unaffected by the manipulations of imagined perspective and disorientation suggests that participants followed instructions, and did not expend further cognitive effort after indicating their preparedness to respond. Accordingly, only decision time and accuracy are discussed. The overall larger errors in pointing, compared with verbal labeling, are attributed to added noise in the joystick response. Whereas verbal responses allowed for high precision, joystick responses were performed with only proprioceptive feedback (participants were blindfolded during responding). Although attempts were made to reduce this noise by quantizing the pointing responses in 45° intervals, this was not sufficient to equate the two responses, presumably because noise in the joystick response often exceeded the 45° quantization.

Pointing and verbal labeling responses were faster and more accurate when participants imagined the learning perspective than when they imagined the misaligned perspective. This evidence for an orientation-dependent representation occurred under both oriented and disoriented conditions, for both verbal and pointing responses. Previous research on long-term spatial memory predicted this finding, not only because of the saliency of the learning view (Diwadkar & McNamara, 1997; Kelly et al., 2007; Shelton & McNamara, 1997) but also because of its alignment with the long axis of the room (Kelly & McNamara, 2008; Shelton & McNamara, 2001). One or both of these factors most likely caused participants to organize their memories for the objects around the 0º–180º axis. The fact that both response modes showed evidence of orientation-dependent memories casts doubt on the hypothesis that pointing responses are based on orientation-dependent representations and verbal labeling responses are based on orientation-independent representations; instead, it appears that a common orientation-dependent representation underlies both response modalities.

Additionally, when participants remained oriented to the environment, their pointing and verbal responses were faster and more accurate when they imagined their current perspective, compared with the misaligned perspective. Although the pointing result replicates much of the previous work on imagined rotations (e.g., Presson & Montello, 1994; Rieser, 1989), the verbal labeling result fails to replicate previous work, indicating equivalent access to occupied and nonoccupied perspectives (Avraamides et al., 2007; de Vega & Rodrigo, 2001; Wang, 2004; Wraga, 2003). This finding indicates that the same retrieval process governed both response modes in the present task. To explain the superior performance for imagined perspectives aligned with the body, Presson and Montello suggested that the misaligned perspective required participants to calculate the response from the imagined perspective and then map that response onto their actual perspective, and that this remapping process results in interference. Although this is a reasonable explanation when a pointing response is used, it is not clear why the same interference would occur for verbal responses, which are not dependent on the body.

Before concluding that verbal labeling and pointing responses share common representations as well as common retrieval processes, it is important to understand the differences between this experiment and those that do report response mode differences. Previous research has typically manipulated response mode between participants, compared with the within-participants manipulation in Experiment 1, where participants received training on both response modes before beginning the experiment. In previous studies, different groups of participants might have used different retrieval strategies (resulting in potentially different retrieval processes) for verbal labeling and pointing, whereas participants in the present study may have elected to use a consistent strategy for the two tasks. Verbal labeling responses, but not pointing responses, are commonly used both egocentrically and nonegocentrically, depending on one's goal. In the former case, an observer can verbally describe object locations relative to him/herself, where verbal labels correspond to egocentric directions (e.g., "The chair is 3 m to my right"). Alternatively, verbal labeling can also be done nonegocentrically, from a third-person framework, by describing object locations relative to another person or object (e.g., "The chair is 3 m to your right"). This third-person strategy is not consistent with the typical definition of a perspective-taking task, where the observer is expected to imagine egocentrically occupying the new perspective, rather than imagining someone or something else occupying that perspective. In contrast to the potential flexibility of the verbal response, pointing is reserved for indicating egocentric locations.[1]

It is possible that the discrepancy between the results of Experiment 1 and those reported elsewhere is due to differences in participant strategy. The within-participants manipulation of response mode in Experiment 1 may have encouraged participants to use the same egocentric strategy for both verbal and pointing responses, a more parsimonious solution to the two tasks. In contrast, participants in the verbal labeling condition of previous experiments

may have used a third-person strategy, which does not incur the same interference costs as an egocentric strategy does. In Experiment 2, participants were instructed to use a third-person strategy for both response modes. This nonegocentric strategy should eliminate the cost associated with imagining nonoccupied perspectives when using a verbal response, which can be used to indicate nonegocentric directions. The third-person strategy should not, however, reduce the costs associated with pointing from a nonoccupied perspective, because the correct response will still need to be transformed into body coordinates.

EXPERIMENT 2

Participants in Experiment 2 were instructed to imagine an arrow pointing toward one object and to then indicate the location of a second object relative to that arrow. This modified instruction set was expected to induce a third-person strategy when solving the task. The disorientation condition was removed in Experiment 2 because it did not greatly aid in the identification of response mode differences, since the independent effects of alignment with the body and the learning perspective were both apparent within the oriented test conditions.

Method

Participants. Sixteen undergraduate students (8 males) at Vanderbilt University participated in exchange for course credit.

Stimuli, Design, and Procedure. Stimuli, design, and procedure were similar to those of Experiment 1. Due to the reduced number of conditions, only two of the four object sets from Experiment 1 were used. During training, participants were presented with eight objects used only during the training session. A 25-cm cardboard arrow was placed on the ground, emanating from the center of the circle and pointing toward one object, and participants were instructed to indicate, through pointing or verbal labeling, the direction of a second object relative to the arrow. They were told that the arrow would not be physically present during the experiment, but that they were to imagine the arrow facing the first object and then indicate the direction of the second object relative to the arrow (e.g., "The arrow is facing the pear, find the apple"). The primary independent variables were response mode (pointing or verbal labeling) and imagined perspective (learning, body-aligned, or misaligned). Response mode was blocked and order was counterbalanced. Each block contained 21 trials consisting of seven pointing responses from three imagined perspectives, and trial order was pseudorandomized, as in Experiment 1.

Results

Latency. Decision time (shown in Figure 4) was analyzed in a 2 (gender) $\times$ 2 (response mode: verbal labeling or pointing) $\times$ 3 (imagined perspective: learning, body-aligned, or misaligned perspective) mixed-model ANOVA. Main effects of response type [$F(1,14) = 3.75$, $p = .073$, $\eta_p^2 = .21$] and imagined perspective [$F(2,28) = 21.87$, $p < .001$, $\eta_p^2 = .61$] were qualified by a significant response $\times$ perspective interaction [$F(2,28) = 9.66$, $p = .001$, $\eta_p^2 = .41$]. Interaction contrasts indicated that the latency difference between trials testing the body-aligned and misaligned perspectives was reduced for verbal labeling, relative to pointing [$F(1,14) = 11.90$, $p = .004$, $\eta_p^2 = .46$].

Further contrasts showed that pointing decision times were faster on trials testing the learning and body-aligned perspectives than they were on those testing the misaligned perspective [$F(1,14) = 23.71$, $p < .001$, $\eta_p^2 = .63$, and $F(1,14) = 28.35$, $p < .001$, $\eta_p^2 = .67$, respectively]. Verbal decision times were faster for the learning perspective than for the misaligned perspective [$F(1,14) = 26.55$, $p < .001$, $\eta_p^2 = .66$], and marginally faster for the body-aligned than for the misaligned perspective [$F(1,14) = 3.45$, $p = .085$, $\eta_p^2 = .20$].

Similar analysis of the response time data showed only a main effect of response type [$F(1,14) = 128.23$, $p < .001$, $\eta_p^2 = .90$], with faster pointing responses ($M = 0.60$ sec, $SE = 0.03$) than verbal responses ($M = 2.19$ sec, $SE = 0.14$).

Accuracy. Absolute pointing error (shown in Figure 5) was also analyzed in a 2 (gender) $\times$ 2 (response mode: verbal labeling or pointing) $\times$ 3 (imagined perspective: learning, body-aligned, or misaligned perspective) mixed-model ANOVA. Main effects of response mode [$F(1,14) = 46.39$, $p < .001$, $\eta_p^2 = .77$] and perspective [$F(2,28) = 5.27$, $p = .011$, $\eta_p^2 = .27$] were qualified by a marginal interaction between response mode and perspective [$F(2,28) = 2.84$, $p = .076$, $\eta_p^2 = .17$]. Interaction contrasts indicated that the accuracy difference between trials testing the body-aligned and misaligned perspectives was reduced for verbal labeling relative to pointing responses [$F(1,14) = 9.07$, $p = .009$, $\eta_p^2 = .39$].

Planned contrasts showed that pointing was more accurate for trials testing the learning and body-aligned perspectives than for those testing the misaligned perspective [$F(1,14) = 5.62$, $p = .033$, $\eta_p^2 = .29$, and $F(1,14) = 9.87$, $p = .007$, $\eta_p^2 = .41$, respectively]. Accuracy for verbal responses did not differ among the three imagined perspectives.
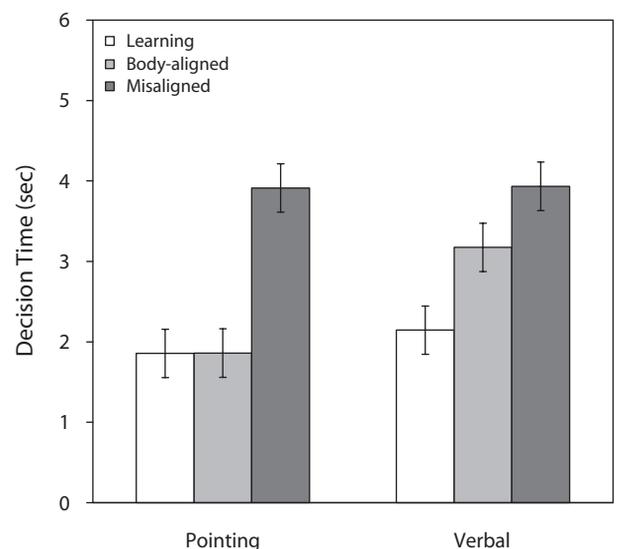

Figure 4. Decision time as a function of response mode and imagined perspective in Experiment 2. Participants were instructed to use a third-person perspective taking strategy. Error bars are standard errors estimated from the ANOVA.
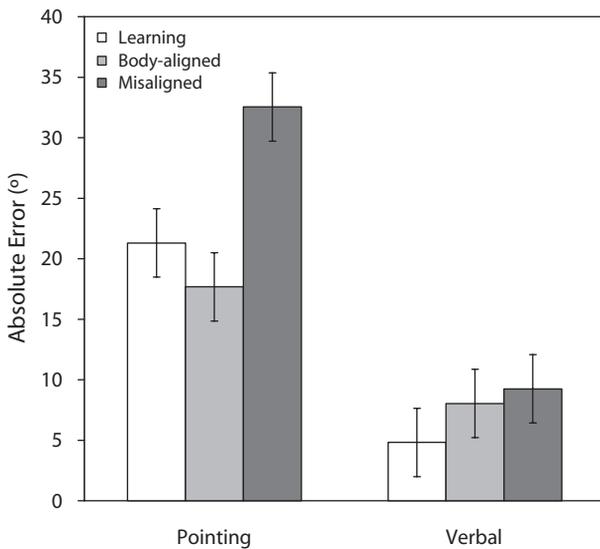
Figure 5. Absolute angular error as a function of response mode and imagined perspective in Experiment 2. Participants were instructed to use a third-person perspective taking strategy. Error bars are standard errors estimated from the ANOVA.

Discussion

As in Experiment 1, the separation of decision time and response time circumvented the longer latencies associated with producing and recording the verbal responses. Response times were longer overall for verbal labeling, but were unaffected by imagined perspective; therefore, we only interpret the decision-time data.

When using a third-person strategy, participants in Experiment 2 pointed faster and more accurately from imagined perspectives aligned, compared with misaligned, with their actual facing direction. The difficulty associated with pointing from the misaligned perspective is attributed to a reference frame conflict between the third-person perspective and the participant's actual perspective. In contrast, when using a verbal response, participants were slightly faster on body-aligned perspectives than on misaligned ones, and equally accurate. The interference that impaired pointing from the misaligned perspective was substantially reduced for verbal responses, which could be made directly from the imagined third-person perspective.

The latency data again indicate that pointing and verbal labeling responses were based on an orientation-dependent representation, organized with respect to the 0°–180° axis. However, this finding was not supported in the accuracy data, which indicate an orientation-dependent representation for pointing responses only. Thus, although the data generally support the separate-processes hypothesis of response mode differences, they are equivocal regarding the separate-representations hypothesis.

The next experiment was designed to more closely replicate previous work, which has manipulated response mode between participants. Experiments 1 and 2 suggest that response mode differences are based on different retrieval processes rather than on different representations. Moreover, strategy appears to play a critical role in the specific

retrieval process employed by participants. Experiment 3 was conducted to determine whether participants would naturally use egocentric and nonegocentric strategies for pointing and verbal labeling, respectively, when response mode is manipulated between participants.

EXPERIMENT 3

In Experiment 3, participants again completed two blocks of trials corresponding to the two response modes. Unlike Experiments 1 and 2, instructions for each response mode were given just prior to participation in the appropriate response condition. By manipulating response mode and order (i.e., pointing then verbal vs. verbal then pointing), the first test block from both order conditions is comparable with previous between-participants manipulations of response mode. Including the second test block allows for replication and comparison with data from Experiment 1, where egocentric response strategies used in the pointing task may have influenced the subsequent verbal labeling task.

Method

Participants. Thirty-two undergraduate students (16 males) at Vanderbilt University participated in exchange for course credit.

Stimuli, Design, and Procedure. Stimuli, design, and procedure were similar to those of Experiments 1 and 2. The primary independent variables were response mode (pointing or verbal labeling), order (pointing, then verbal labeling, or vice versa), and imagined perspective (learning, body-aligned, or misaligned). Order was manipulated between participants, and response mode and imagined perspective were manipulated within participants. Upon arrival at the lab, participants were instructed on the response mode to be used in the first block of testing. After completing the first block, they were then instructed on the other response mode to be used in the second block of testing.

Results

Latency. Decision time (see Figure 6) was analyzed in a 2 (gender) × 2 (response mode: verbal labeling or pointing) × 2 (order: pointing, then verbal labeling, or vice versa) × 3 (imagined perspective: learning, body-aligned, or misaligned perspective) mixed-model ANOVA. Only main effects of response mode [$F(1,28) = 4.46, p = .044, \eta_p^2 = .14$] and imagined perspective [$F(2,56) = 28.02, p < .001, \eta_p^2 = .39$] were significant. Decisions were faster for pointing responses ($M = 2.45$ sec, $SE = 0.25$) than for verbal responses ($M = 2.86$ sec, $SE = 0.32$), and faster for the learning ($M = 2.19$, $SE = 0.22$) and body-aligned ($M = 2.36$, $SE = 0.23$) perspectives than for the misaligned perspective ($M = 3.41$, $SE = 0.41$). No other main effects or interactions were significant.

To further evaluate a priori hypotheses, performance on the misaligned perspective was compared with performance on the body-aligned and original perspectives for each combination of response mode and response order. When verbal labeling occurred before pointing (Block 1 of the verbal-then-pointing order), contrasts showed that verbal labeling decision time was faster when testing the learning perspective than when testing the misaligned perspective [$F(1,14) = 26.56, p < .001, \eta_p^2 = .66$], but did not differ between the body-aligned and misaligned
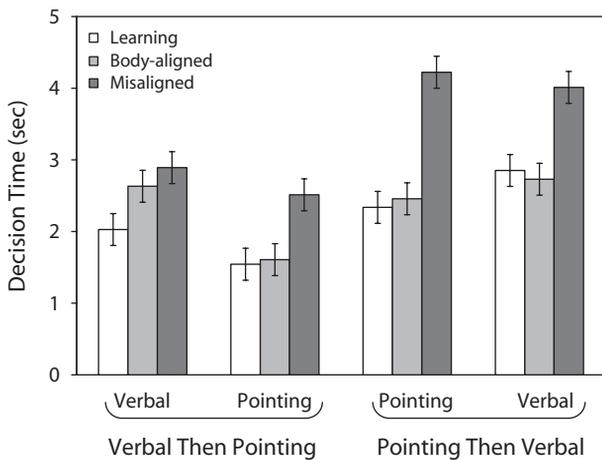
Figure 6. Decision time as a function of response mode, imagined perspective, and response mode order in Experiment 3. Error bars are standard errors estimated from the ANOVA.

perspectives. On the ensuing pointing task (Block 2 of the verbal-then-pointing order), pointing decision time was faster on trials testing the learning and body-aligned perspectives than on those testing the misaligned perspective [$F(1,14) = 8.40, p = .012, \eta_p^2 = .38$, and $F(1,14) = 10.41, p = .006, \eta_p^2 = .43$, respectively]. When pointing occurred before verbal labeling (Block 1 of the pointing-then-verbal order), pointing decision time was faster on trials testing the learning and body-aligned perspectives than on those testing the misaligned perspective [$F(1,14) = 10.10, p = .007, \eta_p^2 = .42$, and $F(1,14) = 7.19, p = .018, \eta_p^2 = .34$, respectively]. Subsequent verbal labeling decision time (Block 2 of the pointing-then-verbal order) was also faster on trials testing the learning and body-aligned perspectives than on those testing the misaligned perspective [$F(1,14) = 6.67, p = .022, \eta_p^2 = .32$, and $F(1,14) = 5.87, p = .03, \eta_p^2 = .30$, respectively].

A similar ANOVA using the response time data again showed only a main effect of response mode [$F(1,28) = 164.99, p < .001, \eta_p^2 = .86$], with faster responses for pointing ($M = 0.57$ sec, $SE = 0.05$) than for verbal labeling ($M = 1.79$ sec, $SE = 0.10$). No other main effects or interactions were significant.

Accuracy. Absolute pointing error (shown in Figure 7) was analyzed in a 2 (gender) × 2 (response mode: verbal labeling or pointing) × 2 (order: pointing, then verbal labeling or vice versa) × 3 (imagined perspective: learning, body-aligned, or misaligned perspective) mixed-model ANOVA. Main effects of gender [$F(1,28) = 4.66, p = .04, \eta_p^2 = .14$], response mode [$F(1,28) = 37.21, p < .001, \eta_p^2 = .57$], order [$F(1,28) = 6.65, p = .015, \eta_p^2 = .19$], and imagined perspective [$F(2,56) = 18.70, p < .001, \eta_p^2 = .40$] were qualified by the following interactions: response mode and gender [$F(1,28) = 4.14, p = .051, \eta_p^2 = .13$], where males were more accurate than females, but only for pointing responses; response mode and order [$F(1,28) = 9.60, p = .004, \eta_p^2 = .26$], where pointing responses were more accurate when they occurred in the

second test block than in the first block; response mode and imagined perspective [$F(2,56) = 3.93, p = .025, \eta_p^2 = .12$], where pointing responses, but not verbal responses, were facilitated when aligned, compared with misaligned, with the body; gender, order, and perspective [$F(2,56) = 2.66, p = .079, \eta_p^2 = .09$], where males showed a greater effect of body alignment when pointing preceded verbal labeling [$F(1,28) = 3.18, p = .085, \eta_p^2 = .10$], and females showed a larger benefit for the learning perspective when verbal labeling preceded pointing [$F(1,28) = 3.46, p = .073, \eta_p^2 = .11$].

To further evaluate a priori hypotheses, performance on the misaligned perspective was compared with performance on the body-aligned and original perspectives for each combination of response mode and response order. When verbal labeling occurred before pointing (Block 1 of the verbal-then-pointing order), contrasts indicated that verbal labeling was more accurate when testing the learning perspective than when testing the misaligned perspective [$F(1,14) = 9.42, p = .058, \eta_p^2 = .18$], but did not differ between the body-aligned and misaligned perspectives. Subsequent pointing responses (Block 2 of the verbal-then-pointing order) were more accurate on trials testing the learning and body-aligned perspectives than on those testing the misaligned perspective [$F(1,14) = 4.98, p = .043, \eta_p^2 = .26$, and $F(1,14) = 6.68, p = .022, \eta_p^2 = .32$, respectively]. When pointing occurred before verbal labeling (Block 1 of the pointing-then-verbal order), pointing was more accurate on the learning and body-aligned perspectives than it was on the misaligned perspective [$F(1,14) = 13.69, p = .002, \eta_p^2 = .49$, and $F(1,14) = 11.58, p = .004, \eta_p^2 = .45$, respectively]. Subsequent verbal labeling responses (Block 2 of the pointing-then-verbal order) were also more accurate on trials testing the learning and body-aligned perspectives than on those testing the misaligned perspective [$F(1,14) = 9.40, p = .008, \eta_p^2 = .40$, and $F(1,14) = 6.37, p = .024, \eta_p^2 = .31$, respectively].
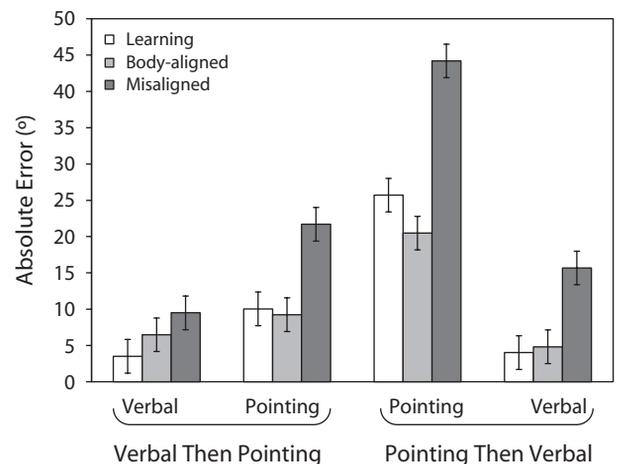

Figure 7. Absolute angular error as a function of response mode, imagined perspective, and response mode order in Experiment 3. Error bars are standard errors estimated from the ANOVA.

Discussion

In Experiment 3, participants were provided with instructions for each response mode just before participating in the appropriate response condition. Therefore, the verbal labeling and pointing responses that occurred in the first block of testing can be compared with previous between-participants manipulations of response mode. To that end, results from the first test block replicate previous findings (e.g., Wang, 2004), showing an advantage for the body-aligned over the misaligned perspective for pointing, but not for verbal labeling. This finding indicates that response mode differences are due at least in part to different retrieval processes, since both body-aligned and misaligned perspectives were equally misaligned with the learning perspective. On the basis of Experiments 1 and 2, these retrieval differences may be grounded in strategy differences, where participants in the pointing condition used an egocentric strategy and participants in the verbal labeling condition used a nonegocentric strategy. Results from the first test block also indicate that pointing and verbal labeling responses are both based on orientation-dependent representations. For both response modes, the learning perspective was easier to imagine than the misaligned perspective was. Consistent with the previous two experiments, findings from the first block of testing in Experiment 3 indicate that response mode differences in perspective taking are due to different retrieval processes, not to different representations.

Results from the second test block replicate findings from Experiment 1, showing that verbal responses conducted in the context of pointing result in superior performance on the body-aligned perspective. In contrast, the pointing response results were unaffected by the context of the preceding verbal response task.

GENERAL DISCUSSION

Previous studies have found that perspective-taking performance varies with the response mode used to make spatial judgments from the imagined perspective (Avraamides et al., 2007; Avraamides et al., 2004; de Vega & Rodrigo, 2001; Wang, 2004; Wraga, 2003). Wang proposed that these response mode differences could be due to differences in the underlying spatial memory representation, or to differences in the spatial memory retrieval process; the experiments reported here directly address these two possibilities. First, these experiments indicate that verbal and pointing responses are both based on an orientation-dependent representation. Across all three experiments, there was evidence that the learning perspective was facilitated for both pointing and verbal responses, despite the fact that the learning perspective was actually misaligned with participants' facing directions during testing. This suggests that participants represented the layout relative to a reference direction selected during learning, and that this organization persisted even after body rotation and disorientation. The selected reference direction may have been influenced by the learning perspective and/or the long axis of the room, both of which have been previously shown to influence spatial memory organization (Shelton & Mc-

Namara, 2001). This corroborates a large body of work reporting orientation-dependent spatial memories when using pointing responses, and indicates the generality of this finding.

Second, these experiments indicate that the retrieval processes associated with pointing and verbal responses depend greatly on experimental conditions. Pointing responses are reserved for indicating egocentric directions, and pointing performance across all three experiments was better for perspectives aligned rather than misaligned with participant facing direction. The difficulty in imagining misaligned perspectives is attributed to a reference frame conflict between the imagined perspective and the participant's body, which is used to execute the pointing response. In contrast, verbal responses can be used either egocentrically or nonegocentrically, and facilitation of the body-aligned perspective was found to vary across experimental conditions. In Experiments 1 and 3, prior experience with pointing led participants to use verbal responses egocentrically, resulting in facilitated access to imagined perspectives aligned, compared with misaligned, with their body orientation. In Experiment 2, instructions to use a nonegocentric strategy were sufficient to override the effect of previous pointing experience, and the resulting verbal labeling performance was comparable for perspectives aligned and misaligned with the body. Collectively, these findings emphasize the greater flexibility of verbal responses over pointing responses, which are not amenable to nonegocentric strategies.

It is unclear exactly why participation in the pointing task affected the strategy used in the subsequent verbal labeling task. Participants in Experiment 3 who performed verbal labeling in the first block used a nonegocentric strategy, but surely some of them had manually pointed to things in the days preceding the experiment. It is possible that elapsed time between the different response modes is sufficient to prevent the observed strategy carryover between the pointing and verbal labeling tasks, which occurred within minutes of each other in the present experiments. The two tasks also shared many features apart from their temporal proximity, such as the organization of the learned spatial layouts, the room environment, and the buttonpress preceding each response. Understanding the specific causes of strategy carryover may be important in interpreting research on spatial memory retrieval, where different response modes and different experimental designs can support considerably different conclusions.

REFERENCES

Avraamides, M. N., Ioannidou, L. M., & Kyranidou, M. N. (2007). Locating targets from imagined perspectives: Comparing labeling with pointing responses. *Quarterly Journal of Experimental Psychology*, **60**, 1660-1679.

Avraamides, M. N., Klatzky, R. L., Loomis, J. M., & Golledge, R. G. (2004). Use of cognitive vs. perceptual heading during imagined

locomotion depends on response mode. *Psychological Science*, **15**, 403-408.

DE VEGA, M., & RODRIGO, M. J. (2001). Updating spatial layouts mediated by pointing and labelling under physical and imaginary rotation. *European Journal of Cognitive Psychology*, **13**, 369-393.

DIWADKAR, V. A., & MCNAMARA, T. P. (1997). Viewpoint dependence in scene recognition. *Psychological Science*, **8**, 302-307.

EASTON, R. D., & SHOLL, M. J. (1995). Object-array structure, frames of reference, and retrieval of spatial knowledge. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 483-500.

KELLY, J. W., AVRAAMIDES, M. N., & LOOMIS, J. M. (2007). Sensorimotor alignment effects in the learning environment and in novel environments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **33**, 1092-1107.

KELLY, J. W., & MCNAMARA, T. P. (2008). Spatial memories of virtual environments: How egocentric experience, intrinsic structure, and extrinsic structure interact. *Psychonomic Bulletin & Review*, **15**, 322-327.

KLATZKY, R. L., LOOMIS, J. M., BEALL, A. C., CHANCE, S. S., & GOLLEDGE, R. G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science*, **9**, 293-298.

MAY, M. (2004). Imaginal perspective switches in remembered environments: Transformation versus interference accounts. *Cognitive Psychology*, **48**, 163-206.

MCNAMARA, T. P., RUMP, B., & WERNER, S. (2003). Egocentric and geocentric frames of reference in memory of large-scale space. *Psychonomic Bulletin & Review*, **10**, 589-595.

MOU, W., & MCNAMARA, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 162-170.

MOU, W., MCNAMARA, T. P., VALIQUETTE, C. M., & RUMP, B. (2004). Allocentric and egocentric updating of spatial memories. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 142-157.

PHILBECK, J. W., LOOMIS, J. M., & BEALL, A. C. (1997). Visually perceived location is an invariant in the control of action. *Perception & Psychophysics*, **59**, 601-612.

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

RIESER, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 1157-1165.

ROSKOS-EWOLDSEN, B., MCNAMARA, T. P., SHELTON, A. L., & CARR, W. (1998). Mental representations of large and small spatial layouts are orientation dependent. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 215-226.

SHELTON, A. L., & MCNAMARA, T. P. (1997). Multiple views of spatial memory. *Psychonomic Bulletin & Review*, **4**, 102-106.

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

WALLER, D., MONTELLO, D. R., RICHARDSON, A. E., & HEGARTY, M. (2002). Orientation specificity and spatial updating of memories for layouts. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 1051-1063.

WANG, R. F. (2004). Action, verbal response, and spatial reasoning. *Cognition*, **94**, 185-192.

WERNER, S., & SCHMIDT, K. (1999). Environmental reference systems for large-scale spaces. *Spatial Cognition & Computation*, **1**, 447-473.

WRAGA, M. (2003). Thinking outside the body: An advantage for spatial updating during imagined versus physical self-rotation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 993-1005.

NOTE

1. Pointing responses to egocentrically defined targets need not be executed from the same egocentric framework. For example, pointing with the hand may be conducted with respect to the shoulder frame of reference, and pointing with a joystick may be conducted from the joystick frame of reference.