Journal of Experimental Psychology:
Human Perception and Performance
2018, Vol. 44, No. 3, 337–355

0096-1523/18/$12.00 http://dx.doi.org/10.1037/xhp0000464

Spatial Perspective Taking Mediated by Whole-Body Motor Simulation

Hiroyuki Muto, Soyogu Matsushita, and Kazunori Morikawa
Osaka University

Humans can envision the world from other people's viewpoints. To explore the embodied process of such spatial perspective taking, we examined whether action related to a whole-body movement modulates performance on spatial perspective-taking tasks. Results showed that when participants responded by putting their left/right foot or left/right hand forward, actions congruent with a movement's direction (clockwise/counterclockwise) reduced RTs relative to incongruent actions. In contrast, actions irrelevant to a movement (a left/right hand index-finger response) did not affect performance. Furthermore, we demonstrated that this response congruency effect cannot be explained by either spatial stimulus-response compatibility or sensorimotor interference. These results support the involvement of simulated whole-body movement in spatial perspective taking. Moreover, the findings revealed faster foot responses than hand responses during spatial perspective taking, whereas the opposite result was obtained during a simple orientation judgment task without spatial perspective taking. Overall, our findings highlight the important role of motor simulation in spatial perspective taking.

Public Significance Statement

Spatial perspective taking is a human ability to envision the world from other people's viewpoints. Our five behavioral experiments show that during spatial perspective taking, people mentally simulate whole-body movement as if they moved to a position from which they took a new perspective. Specifically, we demonstrated that actions congruent with a movement's direction facilitated spatial perspective taking compared with incongruent actions. This response congruency effect was observed only when the action was relevant to whole-body movement. Furthermore, we also demonstrated that foot responses were faster than hand responses for spatial perspective taking although hand responses were faster than foot responses for a task for which spatial perspective taking was unnecessary. These findings highlight the important role of motor processing in spatial perspective taking, suggesting that spatial cognition is closely related to bodily movement.

Keywords: spatial perspective taking, embodied cognition, spatial cognition, mental rotation, motor simulation

Humans are capable of understanding the world from other people's viewpoints, for example, you can ask a friend to pass you a glass on his or her right side, even when the glass is not on the right side from your perspective. This type of spatial problem can be solved readily or even sometimes automatically (Tversky & Hard, 2009); however, other primates seem to be incapable of such spatial perspective taking[1] (e.g., Tomasello, Carpenter, Call, Behne, & Moll, 2005). Previous studies have shown that spatial perspective-taking ability relates closely to a variety of other important abilities, such as navigation (Wolbers & Hegarty, 2010, for a review), theory of mind (Hamilton, Brindley, & Frith, 2009), and empathic perspective taking (Erle & Topolinski, 2015). However, cognitive processes underlying spatial perspective taking have not yet been adequately elucidated. The present study addresses this issue using an embodied cognition approach.

Object-Based and Perspective Transformations

Pioneering studies on spatial perspective taking by Presson et al. focused on comparing object-based and perspective transformations (e.g., Huttenlocher & Presson, 1973, 1979; Presson, 1982). While object-based transformations refer to operating a mental

This article was published Online First August 10, 2017.

Hiroyuki Muto, Soyogu Matsushita, and Kazunori Morikawa, School of Human Sciences, Osaka University.

Soyogu Matsushita is now at Department of Beauty and Fashion Studies, Faculty of Liberal Arts, Osaka Shoin Women's University.

This work was supported by a JSPS Grant-in Aid for JSPS Fellows (16J00012).

Correspondence concerning this article should be addressed to Hiroyuki Muto, School of Human Sciences, Osaka University, 1-2 Yamadaoka, Suita-shi 565-0871, Japan. E-mail: h_muto@hus.osaka-u.ac.jp

[1] Previous studies (e.g., Surtees, Apperly, & Samson, 2013) have reported two forms of spatial perspective taking; one is related to an understanding of whether another person can see a particular object (e.g., visibility or front/behind judgments) and the other is related to an understanding of where an object is located from another person's viewpoint (e.g., left/right judgments). Because the former can be performed by drawing a line between another person and an object (Kessler & Rutherford, 2010; Michelon & Zacks, 2006; Surtees et al., 2013) and thus does not require perspective transformations, we focus only on the latter form, referring to it as "spatial perspective taking" for convenience.

image of an object or an array, perspective transformations refer to operating a mental image of the self, and spatial perspective taking has been assumed to be one form of perspective transformations (see Zacks & Michelon, 2005, for a review). Presson et al. found that the two transformations were processed differently.

The most studied object-based transformation is mental rotation of an object. In the initial experiment of Shepard and Metzler (1971), participants were presented a pair of two pictorial three-dimensional objects comprising cubes and were asked to respond as quickly as possible as to whether the two objects were the same or different. Results showed that response times (RTs) for same-different judgments increased linearly with the angular disparity between the two objects. This suggested that mental imagery can be rotated just like a real object. Analogous to the mental rotation of an object, perspective transformations have been extensively studied in terms of mental rotation of the self or viewer rotation (e.g., Amorim & Stucchi, 1997; Carpenter & Proffitt, 2001; Creem, Downs, et al., 2001; Creem, Wraga, & Proffitt, 2001; Lambrey, Doeller, Berthoz, & Burgess, 2012; Wraga, Creem, & Proffitt, 2000). Most previous research has shown that performance (i.e., speed or accuracy) on both kinds of mental transformation are impaired with increasing angles of rotation; this implies the existence of mental spatial transformations analogous to physical ones.

Regarding different mental spatial transformations, Zacks et al. proposed a multiple systems framework (e.g., Zacks & Michelon, 2005; Zacks & Tversky, 2005). This framework assumes that the two forms of mental spatial transformations are implemented to some degree by distinct neural substrates, which are hypothesized to have been shaped by natural selection. This means that unique neural and cognitive mechanisms underlie each form of transformation, and they lead to unique physiological or behavioral consequences. Several empirical studies have provided evidence for the multiple systems framework. For example, some studies have shown that object rotation and viewer rotation depend on different neural structures (Lambrey et al., 2012; Wraga, Shephard, Church, Inati, & Kosslyn, 2005; Zacks, Vettel, & Michelon, 2003), and, in fact, viewer rotation can usually be performed more efficiently than object or array rotation (e.g., Amorim & Stucchi, 1997; Presson, 1982; Wraga et al., 2000), particularly when the rotational axis is perpendicular to the horizontal plane (Carpenter & Proffitt, 2001; Creem, Wraga, et al., 2001). Furthermore, humans can select an appropriate transformation for a given situation, and instructions to use an inappropriate transformation adversely affect task performance (Zacks & Tversky, 2005). Other studies have shown that psychometric tests can measure abilities related to each transformation as two separable factors (Hegarty & Waller, 2004; Kozhevnikov & Hegarty, 2001) and that the ability of object-based transformation develops earlier in childhood and declines with age later than that of perspective transformation (Huttenlocher & Presson, 1973; Inagaki et al., 2002). These findings are all consistent with the multiple systems framework.

Spatial Perspective Taking as a Perspective Transformation

Thus far, spatial perspective taking has been naively (or perhaps implicitly) thought of as a form of perspective transformation because the results of typical experiments on spatial updating or perspective change have shown monotonic increases in RT or error with the rotational angle (e.g., Easton & Sholl, 1995; Rieser, 1989). However, some researchers have proposed a different interpretation of the angle effect in terms of sensorimotor interference (e.g., Brockmole & Wang, 2003; May, 2004; Wang, 2005). According to this account, impaired performance associated with an angle is attributed not to an additional cognitive effort of mental transformations but to interference conflict between real and imagined perspectives. For example, May (2004) provided empirical evidence favoring the sensorimotor interference account. He compared angle effects of self-translation and self-rotation while controlling for the amount of angular disparity between real and imagined perspectives. In the self-translation condition, efforts of mental transformations were the same regardless of angular disparities because the distance between real and imagined positions was constant. Thus, if the transformation was needed, the angle effect would appear only in the self-rotation condition. However, results showed monotonic increases of RT and error as a function of angular disparity for both translation and rotation conditions. Furthermore, the angle effect was observed even when extra time was given so that participants could complete, if any, a mental transformation in advance (May, 2004; Wang, 2005). These findings seem to contradict the transformation account.

Nonetheless, these findings do not necessarily deny the transformation account. First, as indicated by Kessler and Thomson (2010), tasks used by May (2004) and Wang (2005) imposed a heavy cognitive load on working memory. During their tasks, participants had to maintain simultaneously a complicated array of four or five objects and the self's updated location. This might have motivated participants to use another strategy (e.g., simply wait and do nothing during the extra time) against researchers' expectations. Second, most previous research on perspective change has used a task that can be solved largely based on knowledge from long-term memory, for example, a previously remembered array (e.g., May, 2004; Wang, 2005) or a familiar environment (e.g., Brockmole & Wang, 2003). Such knowledge-based offline processes might be helpful in some situations, such as route planning or giving navigational directions.

However, online processes are also important for real-life spatial problem solving. In fact, many spatial problems in daily life are solved by real-time processing rather than a priori knowledge because of limited time, lack of knowledge or cognitive tools for using it, or difficulty in the knowledge-level solution (Freksa & Schultheis, 2014). In addition, most studies on object-based transformations have used tasks requiring real-time processes (e.g., Shepard & Metzler, 1971). Perhaps online processes require a more concrete strategy (e.g., mental transformation) than offline processes that might prompt a more abstract strategy (e.g., calculation or verbal thought). Consistent with this view, Kessler and Thomson (2010) provided evidence that spatial perspective taking involves "embodied" transformations using a task that emphasized real-time processing (described in detail in the following section). To elucidate cognitive processes of spatial perspective taking as a mental transformation, the present study also focuses on online processes.

Embodiment in Spatial Perspective Taking

Given that human evolution covers less than 1% of the entire evolutionary history of life on Earth, high-level cognitive functions unique to humans are likely based largely on primitive functions such as motor processing (Waller, 2014). In other words, cognition is embodied. Approaches based on such embodied cognition have thus far revealed that mental object rotation is closely related to physical hand movements. For example, concurrent rotational hand movements facilitate or inhibit mental object rotation when they are congruent or incongruent with the direction of mental rotation, respectively (Wexler, Kosslyn, & Berthoz, 1998; Wohlschläger & Wohlschläger, 1998); same-different judgments via mental and physical rotations yield a similar RT pattern (Gardony, Taylor, & Brunyé, 2014; Wohlschläger & Wohlschläger, 1998); and objects difficult to move physically by hand are also difficult to move in mental imagery (Flusberg & Boroditsky, 2011). These findings suggest shared processing between mental object rotation and motor simulation of hand movements; this has been corroborated by neuroimaging studies' reports of brain activities in motor regions (Zacks, 2008, for a meta-analysis and review).

Less attention, however, has been paid to the kinesthetic aspects of spatial perspective taking, but Kessler and Thomson (2010) introduced a promising new approach. They used a round-table stimulus on which two objects (a gun and a flower) were laid in front of a sitting avatar (Experiments 1 and 4) or an empty chair (Experiment 2). Participants were asked to judge the position (left or right) of a target object indicated in advance from the avatar or the chair's perspective. Consistent with other studies, results showed monotonically increasing RTs with angular disparity between participants' actual and imagined perspectives. Ingeniously, Kessler and Thomson (2010) also manipulated the actual orientation/posture of participants' bodies in a clockwise or counterclockwise direction. They found that body posture congruent with an imagined movement's direction facilitated spatial perspective taking compared with straight body posture (baseline), and incongruent body posture hindered spatial perspective taking compared with baseline. This posture congruency effect could not be accounted for by the angle difference between a body orientation and an imagined perspective; it thus contradicted the sensorimotor interference account. Instead, the posture congruency effect depended on whether body posture was congruent or incongruent with the imagined movement direction. Therefore, Kessler and Thomson (2010) concluded the existence of embodied transformation. An interesting find was that the posture congruency effect disappeared in a comparable task that required object-based transformations instead of perspective transformations, suggesting the involvement of a whole-body schema in spatial perspective taking, not that of a specific body part (i.e., hand) as in mental object rotation (Experiment 3 in Kessler & Thomson, 2010).

Although Kessler and Thomson (2010) elegantly demonstrated that spatial perspective taking is embodied in simulated movements, its underlying mechanism remains unclear. For example, they claimed that a whole-body schema was involved in spatial perspective taking, which has yet to be proven because their manipulation of participants' body posture could affect representations of both a whole-body and specific body parts (i.e., turning the whole-body orientation also altered the position of the arms and legs). To confirm the involvement of the whole-body schema, we have to manipulate different body parts (e.g., feet and hands) separately.

It also remains unclear whether actions related to a whole-body movement affect the spatial perspective taking. A number of previous studies demonstrated the involvement of motor simulation in various tasks such as mental object rotation (e.g., Schwartz & Holton, 2000; Wexler, Kosslyn, & Berthoz, 1998; Wohlschläger & Wohlschläger, 1998) and imagined locomotion (e.g., Kunz, Creem-Regehr, & Thompson, 2009) by examining the effect of concurrent physical action on the performance. If spatial perspective taking involves motor simulation of a whole-body movement, it should be affected only by actions related to a whole-body movement. Thus, the effect of actions would be a more direct evidence of simulated whole-body movement than the posture effect (Kessler & Thomson, 2010). Although some neuroimaging studies have reported activations of brain regions associated with motor processing during perspective transformations (Creem, Downs, et al., 2001; Schwabe, Lenggenhager, & Blanke, 2009; Wraga et al., 2005), very little behavioral data exist to help interpret such neuroscientific findings. This has led to controversy regarding the involvement of motor simulation in spatial perspective taking (e.g., Wraga et al., 2005). To dissipate this controversy, we need behavioral studies that examine the effect of actions.

The Present Study

To determine whether simulated whole-body movement shares a common process with spatial perspective taking, the present study manipulated a response method in which participants indicated their judgments about the position (left or right) of a target object in a task that resembled one used by Kessler and Thomson (2010). We assume that when participants intend to move in a clockwise or counterclockwise direction along the edge of a round table, they must put the left or right side of their bodies forward first, respectively (see Figure 1). Indeed, in our preliminary study of a real situation, we confirmed this assumption: A majority of 10 participants tended to move their left foot to start walking in the clockwise direction, but their right foot in the counterclockwise direction (for details, see the Appendix. If spatial perspective taking is analogous to such whole-body movements, corresponding motor simulation should facilitate the mental transformation process. Therefore, our hypothesis predicts that responses congruent with the direction of an imagined movement (e.g., moving the left foot forward during a clockwise transformation) would facilitate spatial perspective taking compared with incongruent responses (e.g., moving the left foot forward during a counterclockwise transformation).

The present study's task used 0, 40, 80, 120, and 160° angle conditions in clockwise and counterclockwise directions (see Figure 1). To focus on the top-down processing of spatial perspective taking, the viewpoints to be imagined were represented by a chair but not by an avatar because the avatar's existence triggers additional bottom-up processing (Kessler & Thomson, 2010). If our hypothesis is correct, the response congruency effect would lead to a result similar to the posture congruency effect observed in Experiment 2 in Kessler and Thomson (2010). That is, the congruency effect would occur only in high angle conditions (i.e., 120 and 160°) because low angle conditions might allow direct judg-

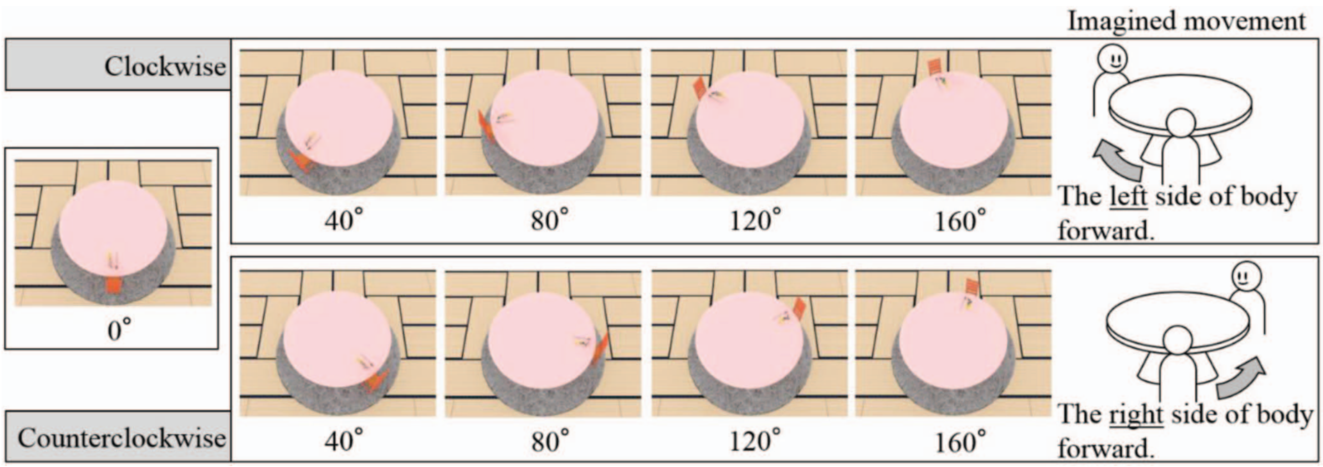

Figure 1. Stimuli used in the spatial perspective-taking task (Experiments 1, 2, 4, and 5). We assumed that participants first imagined moving the left or right side of their bodies forward depending on the stimuli presented. See the online article for the color version of this figure.

ments without perspective transformations (Kessler & Thomson, 2010).

Experiment 1

Experiment 1 examined whether performance on a spatial perspective-taking task is influenced by putting the left/right foot forward to respond. Our hypothesis predicts that an action congruent with the direction of an imagined movement would facilitate spatial perspective taking relative to an incongruent action, especially in high angle conditions (120 and 160°), in which spatial perspective taking is more involved than in low angle conditions (Kessler & Thomson, 2010).

Method

Participants. Participants in Experiment 1 were 24 undergraduate and graduate students (mean age = 21.4 years; 12 female and 12 male; 23 right-footed and 1 left-footed[2]). All had normal or corrected-to-normal vision, were naïve to the study's purpose, and received either prepaid cards for purchasing books or course credit for their participation. We determined this number of participants in advance, following Kessler and Thomson (2010) who chose the same sample size of 24 in all their experiments. According to post hoc analyses, this sample of 24 would give us more than .99 power to detect the main effect of congruency and the interaction of angle and congruency for RT data at the .05 significance level if the response congruency effect has as large effect sizes as the posture congruency effect in Kessler and Thomson's (2010) Experiment 3. For the same reason, we applied this sample size to Experiments 2, 4, and 5 as well. All experiments reported in this article were approved by the ethics board of the School of Human Sciences of Osaka University.

Stimuli and apparatus. Visual stimuli were created using the three-dimensional (3D) computer graphics software Blender 2.71 (Blender Foundation, Amsterdam). Stimuli showed a room with a circular table on which a flower (a chrysanthemum) and a sword were lying in front of a chair. The chair was positioned at 0, 40, 80, 120, or 160° angular disparity from the participants' viewpoint, clockwise or counterclockwise (see Figure 1). Our stimuli mimicked those used by Kessler and Thomson (2010). The circular table was viewed from an angle of 65° from horizontal. Although this kind of bird's eye view is somewhat unnatural in daily life, we adopted this angle for two reasons. First, we wanted to use stimuli comparable with those used by a number of previous studies on spatial perspective taking (e.g., Dalecki, Hoffmann, & Bock, 2012; Kessler & Rutherford, 2010; Kessler & Thomson, 2010; Michelon & Zacks, 2006; Surtees, Apperly, & Samson, 2013). Second, if the table had been viewed from a lower angle, the two target objects and their separation would have been foreshortened, so their appearance would have varied too much depending on their location. This might have contaminated results because people are notoriously poor at precisely estimating depth dimension from 2D pictures (e.g., Sugihara, 2015).

Stimuli were displayed on a 24.1-in-wide LCD monitor (NEC MultiSync LCD-PA241W; resolution of 1,920 × 1,200 pixels) at a viewing distance of about 80 cm. As shown in Figure 2A, participants stood, without their shoes, on a mat in front of a white line marked on the floor. A 120 mm (width) × 67 mm (length) dual-foot switch (USB 2FOOT SWITCH, Scythe Co., Ltd., Tokyo) was fixed on the floor about 2 cm in front of the participants' toes as a response device.

Procedure. Figure 3 illustrates the stimulus sequence. All participants completed the experiment individually in a laboratory. Each trial was initiated with a "Ready?" visual cue, which remained until participants stepped on either the right or left switch using one foot. During this time, participants could check the number of remaining trials by pressing the "T" key on a keyboard placed in front of the monitor. The participants' step initiated a 1-s blank screen; then a picture of the target object (flower or sword) appeared, with its noun (in Japanese kanji) for 1 s. Then, following

[2] In all experiments reported here, we determined participants' dominant hand and foot by asking "which is your dominant hand?" and "which foot do you use to kick a ball?", respectively.

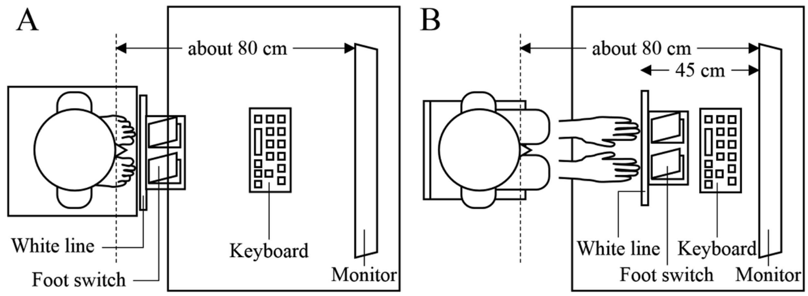

Figure 2. Overhead views of setups used in Experiments 1, 2, 3, and 5. (A) The foot condition. Without shoes, participants stood on a mat in front of a white line marked on the floor and responded by stepping on a foot switch. The position and tilt of a display were adjusted per participant so the viewing distance was about 80 cm. (B) The hand condition. Participants sat on a pipe chair with their hands placed in front of a white line on a table and responded by pushing a foot switch. A washcloth covered the foot switch for hygienic reasons, but it is not drawn here for the sake of simplicity.

a 1-s blank screen again, the experimental stimulus was presented. Participants imagined the viewpoint from the chair and then judged whether the target object would be on the chair's left or right side. They responded by stepping on the corresponding switch (left or right) with one foot as quickly and accurately as possible. The response foot (left or right) was manipulated across two blocks. During a trial, participants had to keep their eyes on the monitor. After the response, a 1-s blank screen appeared, and then the initial cue ("Ready?") was presented again for the next trial. Only in practice trials was visual feedback given on the blank screen when the response was incorrect. After every stepping response, participants moved the foot back to its original standing position.

The experiment consisted of two blocks of trials. Participants were instructed to keep using the same foot (left or right) to respond throughout each block, regardless of whether the response was left or right. The response foot (left or right) was switched between the two blocks, with the order counterbalanced across participants. Each block consisted of 108 trials in random order; each of nine angular disparities was repeated 12 times. The target object (flower or sword) and its position (left/right or right/left) were counterbalanced across trials. Hence, a correct response was left on half the trials and right on the other half. Before each block, participants completed 20 warm-up trials, in which a blue square was presented on either the left or right position on a gray background, and participants were required to step on the corresponding switch (left or right) and then complete 27 practice trials (randomly selected from main trials). At the end of the experiment, the experimenter asked participants for introspective reports (remarks, used strategies, and troubles faced during the experiment) via open questions.

In this study, participants were explicitly forbidden to infer the correct answer by symmetrically reversing the position from their own viewpoints (i.e., their own "left" = "right" at the table's opposite side), especially at high angles (i.e., 120 and 160°), because such a reversal strategy seems to require processing different from spatial transformation (Kessler & Wang, 2012; Wraga et al., 2000). Otherwise, the experimenter did not imply any specific strategy to be used, such as internal movement simulation or blink transformations (Wraga et al., 2000).

Results and Discussion

For our analyses, we categorized trials into two conditions: congruency between a response foot (left or right) and the imagined movement direction (clockwise or counterclockwise). That is, clockwise trials were regarded as congruent in the left-foot block, but as incongruent in the right-foot block and vice versa for

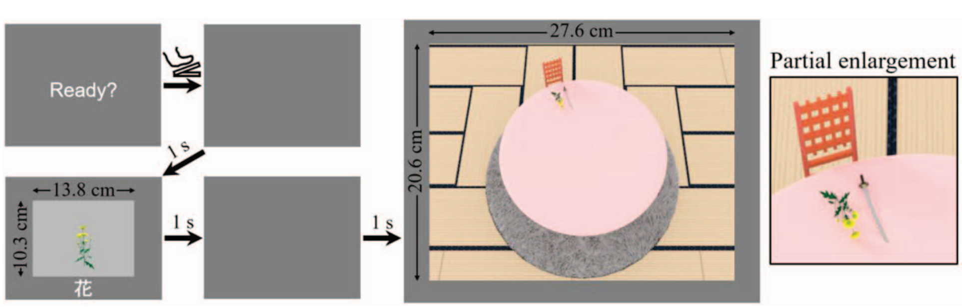

Figure 3. Procedure of the spatial perspective taking task in Experiments 1, 2, and 4. Participants memorized a target (flower or sword) and then judged its position (left or right) on the round table from the viewpoint of the chair. In this example, the correct answer is "right." The rightmost figure depicts objects on the table in a larger scale. See the online article for the color version of this figure.

counterclockwise trials. Because the 0° trials cannot be classified in terms of congruency, they were not included in comprehensive analyses but analyzed separately as necessary. Thus, there are two orthogonal experimental factors: congruency (congruent or incongruent) and angle (40, 80, 120, or 160°). We conducted repeated-measures analysis of variances (ANOVAs[3]) with these two factors on RT and error data. For RT analyses, we excluded error trials (2.3% of data) and trials that took longer than 2.41 s (= $M + 4\ SD$; 0.9% of data[4]) and then calculated the mean RTs per cell for each participant. The mean RTs and errors across participants are shown in Figure 4 and Figure 5, respectively.

The 2 × 4 ANOVA for RT data revealed significant main effects of angle ($F(3, 69) = 60.71$, $\eta_p^2 = .745$, $p < .001$) and congruency ($F(1, 23) = 13.10$, $\eta_p^2 = .463$, $p = .001$) and significant interaction of angle and congruency ($F(1, 69) = 6.17$, $\eta_p^2 = .241$, $p = .002$). Post hoc t tests[5] revealed a monotonic increase of RT with increasing angle, showing significant differences for any pair of two consecutive angles (40 vs. 80°, $t(23) = 2.08$, $d = 0.10$, $p = .049$; 80 vs. 120°, $t(23) = 5.39$, $d = 0.63$, $p < .001$; 120 vs. 160°, $t(23) = 4.31$, $d = 0.55$, $p < .001$). In addition, a separate paired t test confirmed a faster response at 0° than at 40° ($t(23) = 2.88$, $d = 0.13$, $p = .008$). Post hoc t tests also revealed that RTs in the congruent condition were shorter than in the incongruent condition at 120° ($t(23) = 2.91$, $d = 0.18$, $p = .023$) and 160° ($t(23) = 3.35$, $d = 0.26$, $p = .011$), but no congruency effects were detected at 40° ($t(23) = 0.38$, $d = 0.02$, $p = .710$) and 80° ($t(23) = 0.53$, $d = 0.03$, $p = .601$).

For error data, the 2 × 4 ANOVA revealed a significant main effect of angle ($F(3, 69) = 9.57$, $\eta_p^2 = .294$, $p < .001$), but neither of congruency ($F(1, 23) = 0.27$, $\eta_p^2 = .012$, $p = .607$) nor of angle by congruency ($F(1, 69) = 1.24$, $\eta_p^2 = .051$, $p = .294$). Post hoc t tests revealed that more errors occurred at 160° than at 40° ($t(23) = 3.81$, $d = 0.88$, $p = .004$) and 80° ($t(23) = 4.51$, $d = 0.89$, $p = .001$) and showed no other significant differences (all ps $> .070$). Given that very few errors occurred (2.3% overall), we consider RT data the major index of task performance.

Interpretation of the angle effect. Results showed a trend toward longer RTs and more errors with increasing angle, consistent with a number of previous studies on perspective change and viewer rotation (e.g., Carpenter & Proffitt, 2001; Creem, Wraga, et al., 2001; Easton & Sholl, 1995; Huttenlocher & Presson, 1973, 1979; Kessler & Rutherford, 2010; Kessler & Thomson, 2010; Michelon & Zacks, 2006; Presson, 1982; Rieser, 1989; Surtees et al., 2013; Wraga et al., 2000). This angle effect should be interpreted with caution as explained above. According to participants' introspective reports, 66.7% (16 of 24) spontaneously reported adopting a concrete perspective-taking strategy (e.g., "I imagined myself rotating around the table"; "I imaginatively moved to and sat on the depicted chair and then reached for a target object from the imagined position"). In other words, the majority consciously imagined placing themselves in a position from which they took a new perspective. The remaining 33.3% did not clearly describe what strategy they used. More importantly, none reported performing mental object rotation or using a reversal strategy. These introspections, suggestive of the angle effect, provided a rare glimpse into the mind because very few studies on spatial perspective taking have so far reported participants' introspections. Because the introspective data were merely an auxiliary measure, not our main concern, they were not conclusive. However, those introspections do suggest that perspective transformation is what most people naturally perform in the present task.

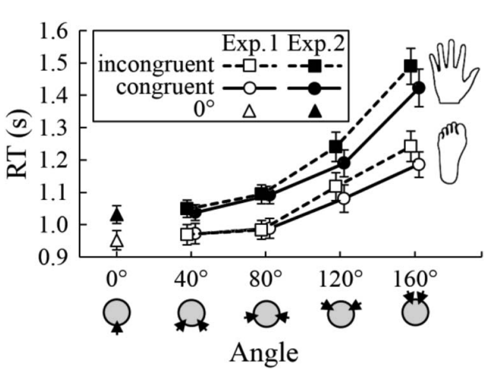

Figure 4. Means and *SE*s of RT data in Experiments 1 and 2.

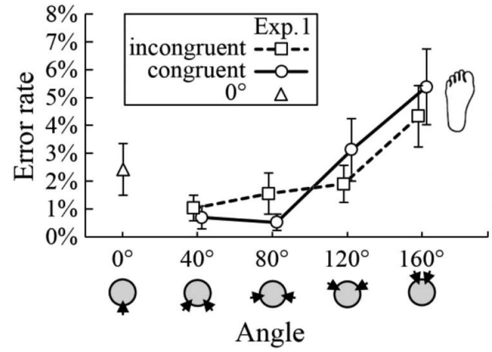

Figure 5. Means and *SE*s of error data in Experiment 1.

The response congruency effect. As we predicted, results showed that RTs at high angle conditions (120 and 160°) were shorter when a response method (putting a left or right foot forward) was congruent with an imagined movement (clockwise or counterclockwise) than when it was incongruent and that the response congruency effect was not detected at low angle conditions (40 and 80°). These results exhibited the same pattern as those of Kessler and Thomson's (2010) Experiment 2, which manipulated body postures. The response congruency effect may indicate that our participants internally engaged in whole-body movement simulation when they responded. Thus, the foot response consistent with simulation was facilitated, compared with the inconsistent response. This implies interdependence between spatial perspective taking and action related to whole-body movement, suggesting involvement of motor simulation in spatial perspective taking.

[3] For the repeated-measures ANOVAs conducted in this article, we reported p values corrected by Chi-Muller's ε (Chi, Gribbin, Lamers, Gregory, & Muller, 2012) without assuming sphericity.

[4] We used this criterion so that omission rates fell around 1% throughout our experiments. Nonetheless, application of another criterion of $M + 3\ SD$ did not affect results of significance tests.

[5] For any multiple comparisons in this article, we reported p values corrected by Holm's (1979) sequentially rejective Bonferroni procedure.

Since the present experiment could not include a baseline condition in which responses were neither congruent nor incongruent, whether spatial perspective taking was facilitated by congruent responses or hindered by incongruent responses remains unclear. On the other hand, Kessler and Thomson (2010) demonstrated both facilitation and interference effects caused by their posture manipulation, depending on whether the posture was congruent or incongruent. Thus, if the response congruency effect shares processes with the posture congruency effect, then the response congruency effect should also contain both facilitation and interference processes.

The occurrence of the congruency effect only in high angle conditions can be attributed to different processes at low and high angles because a position judgment at lower angles can be achieved by direct visual judgments from participants' perspectives and does not necessarily require spatial transformation (Kessler & Thomson, 2010). The difference in the congruency effect between high and low angles may reject another possible account, that is, the spatial stimulus-response (S-R) compatibility. This account predicts that a visual stimulus presented on the participant's right side can be processed faster by the right hand than by the left, even when stimuli's spatial layout is irrelevant to a given task (see Simon, 1990, for a review). If S-R compatibility occurred in our experiment, the congruency effect could be observed at all angle conditions because S-R compatibility occurs even in a very simple task (Simon, 1990). However, this was not the case in our experiment. Thus, the results of Experiment 1 should be interpreted as evidence that spatial perspective taking involves whole-body motor simulation. The possibility of spatial S-R compatibility is further investigated in Experiment 5.

Experiment 2

We demonstrated in Experiment 1 that, compared with the incongruent response, the congruent foot response facilitated spatial perspective taking. This raises the question of whether the congruency effect is specific to a foot response. Kessler and Thomson (2010) suggested that spatial perspective taking involves whole-body representations rather than those of specific body parts, like hands, in mental object rotation (Gardony et al., 2014; Wexler et al., 1998; Wohlschläger & Wohlschläger, 1998). If this is the case, spatial perspective taking might be influenced by any response method related to a whole-body movement, such as extending a left/right arm as well as putting a foot forward. Throughout most of the human species' biological evolution, before humans became bipedal, forelegs were essential to locomotion. Therefore, arm movement might influence spatial perspective taking as a proxy for foot movement when feet could not be used to respond. Thus, Experiment 2 examined whether the results of Experiment 1 can be replicated even when a hand, instead of a foot response, was used.

Method

Participants. Participants in Experiment 2 were 25 undergraduate and graduate students. One male was omitted from analysis because his mean RT was 3 *SD* longer than the mean RT across participants, perhaps because of a lack of the instruction to respond as quickly as possible. Therefore, analyses were based on data from 24 participants (mean age = 21.9 years; 12 female and 12 male; all right-handed). All participants had normal or corrected-to-normal vision, were naïve to the study's purpose, and received either prepaid cards for purchasing books or course credit for their participation. None had participated in the previous experiment.

Stimuli, apparatus, and procedure. The same stimuli and procedure described in Experiment 1 were used in Experiment 2, but the setup was modified for hand responses (Figure 2B). Participants sat on a pipe chair at an 80-cm viewing distance to the monitor and placed their hands in front of a white line marked on a table. The dual foot switch used in Experiment 1 was fixed on the table about 2 cm in front of participants' fingertips and covered with a washcloth for hygienic reasons. During the spatial perspective-taking task, participants responded by pressing the left or right switch with one hand. The response hand (left or right) was switched between two blocks, with the order counterbalanced across participants. After each response, participants replaced the responding hand in the original position.

Results and Discussion

As in Experiment 1, we conducted repeated-measures ANOVAs with two factors (congruency and angle) on RT and error data. For RT analyses, we excluded error trials (2.3% of data) and trials that took longer than 2.93 s (= $M + 4$ *SD*; 0.9% of data) and then calculated the mean RTs per cell for each participant. The mean RTs and errors across participants are shown in Figure 4 and Figure 6, respectively.

The 2 × 4 ANOVA for RT data revealed significant main effects of angle ($F(3, 69) = 64.55$, $\eta_p^2 = .737$, $p < .001$) and congruency ($F(1, 23) = 9.67$, $\eta_p^2 = .297$, $p = .005$) and significant interaction of angle and congruency ($F(1, 69) = 4.24$, $\eta_p^2 = .156$, $p = .016$). Post hoc *t* tests revealed the monotonic increase of RT with increasing angle, showing significant differences for any pair of two consecutive angles (40 vs. 80°, $t(23) = 4.49$, $d = 0.38$, $p < .001$; 80 vs. 120°, $t(23) = 5.03$, $d = 0.68$, $p < .001$; 120 vs. 160°, $t(23) = 8.68$, $d = 0.97$, $p < .001$). In addition, a separate paired *t* test detected no difference between 0 and 40° ($t(23) = 1.09$, $d = 0.94$, $p = .285$). Post hoc *t* tests also revealed that RTs in the congruent condition were shorter than in the incongruent condition at 120° ($t(23) = 3.63$, $d = 0.24$, $p = .006$) and 160° ($t(23) = 2.60$, $d = 0.24$, $p = .048$), but no congruency effects were detected at

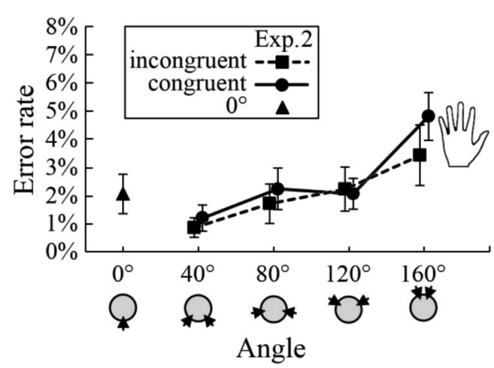

Figure 6. Means and *SE*s of error data in Experiment 2.

40° ($t(23) = 1.18$, $d = 0.09$, $p = .252$) and 80° ($t(23) = 0.20$, $d = 0.02$, $p = .845$).

For error data, the 2 × 4 ANOVA revealed a significant main effect of angle ($F(3, 69) = 6.45$, $\eta_p^2 = .219$, $p = .002$) but neither of congruency ($F(1, 23) = 1.23$, $\eta_p^2 = .051$, $p = .279$) nor of angle by congruency ($F(1, 69) = 0.71$, $\eta_p^2 = .030$, $p = .538$). Post hoc t tests revealed that more errors occurred at 160° than at 40° ($t(23) = 3.67$, $d = 0.85$, $p = .008$) and 80° ($t(23) = 2.99$, $d = 0.52$, $p = .033$), but no other significant differences (all ps $> .160$).

In summary, the same result pattern as in Experiment 1, using a foot response, was obtained in Experiment 2, using a hand response. According to participants' introspective reports, their main strategy was also similar to that in Experiment 1: 62.5% (15 of 24) reported that they used a concrete perspective-taking strategy, and none reported using a reversal strategy. Although a few participants (2 of 24; 8.3%) reported that they performed object rotation in some trials, this is not surprising because multiple solution strategies are commonly used for spatial problems (Schultz, 1991). Hence, no matter which body part (foot or hand) was used for responding, spatial perspective taking was facilitated or inhibited depending on congruency between a response method and the direction of the imagined movement. This suggests involvement not of a specific body part but a whole-body representation in spatial perspective taking.

Comparison Between Experiments 1 and 2

Although similar results were obtained in Experiments 1 and 2, whether the effects of foot and hand movement on spatial perspective taking share a common mechanism is still unknown. To examine this question, we directly compared the results of Experiments 1 and 2. Because Experiments 1 and 2 used the same experimental design (two levels of congruency and four levels of angle as within-participant factors), we can conduct a mixed-design ANOVA on 48 participants' RT data by adding a two-level between-participants factor of the responding body part.[6]

The 2 (responding body part) × 2 (congruency) × 4 (angle) mixed-design ANOVA revealed a significant main effect of the responding body part ($F(1, 46) = 7.57$, $\eta_p^2 = .141$, $p = .008$) and significant interaction of the responding body part and angle ($F(3, 138) = 7.51$, $\eta_p^2 = .140$, $p = .002$). No other two-way and three-way interactions of the responding body part were significant (responding body part and congruency, $F(1, 46) = 0.96$, $\eta_p^2 = .021$, $p = .331$; responding body part and congruency and angle, $F(3, 138) = 0.03$, $\eta_p^2 = .001$, $p = .981$). These results indicated that foot responses were faster than hand responses and that the amount of this foot advantage varied between angles, being largest at 160° (see Figure 4). In addition, a separate Welch's t test revealed a marginally significant foot advantage even at 0° ($t(46) = 1.96$, $d = 0.57$, $p = .056$).

To examine further the foot advantage and the congruency effect, we extracted only high angle conditions (120 and 160°), which may require processing distinct from low angle conditions (Kessler & Thomson, 2010), and then conducted a 2 (responding body part) × 2 (congruency) × 2 (angle) mixed-design ANOVA. The results showed that the main effects of all factors were significant (responding body part, $F(1, 46) = 8.27$, $\eta_p^2 = .152$, $p = .006$; congruency, $F(1, 46) = 28.44$, $\eta_p^2 = .382$, $p < .001$; angle, $F(1, 46) = 85.59$, $\eta_p^2 = .650$, $p < .001$). In addition, two-way interaction of the responding body part and angle was found to be significant ($F(1, 46) = 10.75$, $\eta_p^2 = .189$, $p = .002$), indicating that the foot advantage was more salient at 160° than at 120°. Furthermore, interactions of congruency with any one or two factors were not detected (congruency and responding body part, $F(1, 46) = 0.43$, $\eta_p^2 = .009$, $p = .515$; congruency and angle, $F(1, 46) = 1.21$, $\eta_p^2 = .026$, $p = .278$; congruency and responding body part and angle, $F(1, 46) = 0.01$, $\eta_p^2 < .001$, $p = .941$), implying that the amounts of the congruency effects were equivalent (53 ms on average) regardless of angle (120 or 160°) and responding body part (foot or hand).

Equivalence of the Congruency Effect

We first consider whether motor simulation of foot and hand movements modulates the process of spatial perspective taking in the same way. If the embodied nature of spatial perspective taking were more closely linked to one specific body part than another, the congruency effect would vary depending on the responding body part. However, comparison between experiments revealed that congruency effects in foot and hand conditions were indistinguishable. In other words, foot movement contributed to spatial perspective taking as much as hand movement, at least in the present study. Although not yet conclusive, this is compatible with our hypothesis that spatial perspective taking is mediated by simulated movement of not a specific body part (e.g., foot or hand) but a whole body.

Comparison between experiments also revealed that the RT difference between congruent and incongruent responses at 120° was as large as that at 160°, regardless of the responding body part. If simulation of a whole-body movement functioned throughout spatial perspective taking, the congruency effect would be larger at 160° than at 120° because of the additional demand of longer-distance movement, but this was not the case. Rather, our finding supports the notion that congruent movement leads to a "head-start" effect at the beginning of a perspective transformation, in accordance with Kessler and Thomson's (2010) explanation of the posture congruency effect.

Why Did the Foot Advantage Occur?

Surprisingly, our data showed that foot responses were faster than hand responses in all angle conditions and that they were especially salient at 160°. This phenomenon seems counterintuitive because "the hand is the human's favorite tool and the training effect for other extremities is limited because of physiological conditions" (Pfister et al., 2014, p. 4). Actually, Pfister et al. (2014) demonstrated that the mean RT for hands was shorter than for feet

[6] We can assume that samples in Experiments 1 and 2 were homogeneous for the following two reasons: (a) Because Experiments 1 and 2 were simultaneously planned, their 48 participants were recruited from the same class during the same period. In Japan, because of the strict entrance examination system and rigorous university rankings, students at Japanese universities are much more intellectually homogenous than students at Western universities. Therefore, we have no reason to suspect that a sample of 24 students differs from another sample. (b) Neither experiment showed reliable linear trends of the individual's mean RT for the spatial perspective taking task as a function of participation order (for Experiment 1, $r = -.281$, $p = .184$; for Experiment 2, $r = .342$, $p = .102$).

by strictly measuring simple RTs for a switch release. One possible reason for our contradictory finding is that the foot advantage was induced by a mechanism unique to spatial perspective taking. This unique mechanism, if any, might reflect that feet are more closely related to locomotion than hands. Although this explanation might seem incompatible with the involvement of a whole-body schema as described in the preceding section, the foot advantage is possibly induced by a process different from the response congruency effect. Another possibility is that the foot switch used in our experiments was particularly conducive to foot responses because of its design. In the next experiment, we examined whether the foot advantage was because of the use of the foot switch and whether it is unique to spatial perspective taking.

Experiment 3

Experiment 3 was designed to determine whether the foot advantage observed in Experiments 1 and 2 was unique to spatial perspective taking or if it could be ascribed to other simple reasons (e.g., properties of the response device used and/or general human abilities). For this purpose, Experiment 3 used a simple orientation judgment task, in which spatial perspective taking was unnecessary, but it was otherwise the same as that in Experiments 1 and 2.

Method

Participants. Participants in Experiment 3 were 16 undergraduate and graduate students (mean age = 21.9 years; 8 female and 8 male; all right-handed; 15 right-footed and 1 left-footed). All participants had normal or corrected-to-normal vision, were naïve to the study's purpose, and received prepaid cards for purchasing books for their participation. None had participated in previous experiments. In advance, we determined 16 as a sample size because a multiple of eight was needed for the three counterbalanced factors (gender, foot/hand order, and left/right order).

Stimuli, apparatus, and procedure. The table set stimuli used in Experiments 1 and 2 were replaced with two pictures of a flower and a sword presented side by side (see Figure 7). The trial sequence was the same as that used in the previous experiments: participants first memorized a target object (flower or sword) and then judged its position (left or right) in an arrangement. The target object (flower or sword) and its position (left/right or right/left) were counterbalanced across trials and presented in random order. In Experiment 3, all participants completed both the foot and hand conditions. Setups and response methods for both conditions were the same as those in Experiments 1 and 2, respectively (see Figure 2). Each condition contained left- and right-limb blocks, and each block consisted of 48 trials. Half the participants started with the foot condition, and the other half started with the hand condition. The order of the block (left → right or right → left) was also counterbalanced across participants. Before each block, participants completed eight practice trials, in which visual feedback was given for incorrect responses.

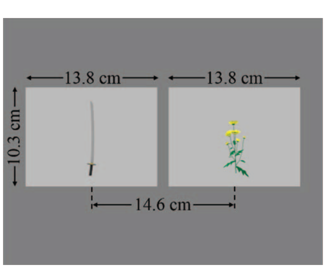

Figure 7. Stimuli presented in the simple orientation judgment task (Experiment 3). See the online article for the color version of this figure.

Results and Discussion

We excluded error trials (0.2% of data) and then calculated the mean RTs of foot and hand responses for each participant. Figure 8 presents the aggregated results. A paired *t* test showed that the mean RT for hand responses was significantly shorter than that for foot responses by 58 ms ($t(15) = 3.11$, $d = 0.54$, $p = .007$), contrary to the results from the spatial perspective-taking task in Experiments 1 and 2. This result suggests that the inherent process of spatial perspective taking induces the foot advantage.

Experiment 4

The results of Experiments 1 and 2 showed that spatial perspective taking was facilitated when participants responded using actions congruent with the direction of an imagined movement, compared with incongruent actions. Although this response congruency effect suggests the involvement of motor simulation of whole-body movement in spatial perspective taking, another interpretation is possible. The congruency effect could simply be attributed to which side of the body, left or right, participants used in responding, regardless of its relevance to a whole-body movement. This interpretation is based on the possibility that the left or right side of the body might function in the same way as body postures did in Kessler and Thomson's (2010) experiments. The left-or-right account predicts that the congruency effect would occur even when a response method is irrelevant to a whole-body movement, as long as a responding body part belongs to either the left or right side of the body. On the other hand, the motor simulation account we hypothesized predicts that a response method irrelevant to a whole-body movement would not cause the congruency effect. To examine which account is valid, Experiment 4 used the response of an index finger, a response movement that is most likely irrelevant to a whole-body movement.

Method

Participants. The participants in Experiment 4 were 24 undergraduate, graduate, and research students (mean age = 22.8 years; 12 female and 12 male; 22 right-handed and 2 left-handed). All had normal or corrected-to-normal vision, were naïve to the study's purpose, and received prepaid cards for purchasing books for their participation. None had participated in previous experiments. According to post hoc analyses of our Experiments 1 and 2, sample size 24 would give us more than .99 power to detect the main effect of congruency and the interaction of angle and congruency for RT data at the .05 significance level. Thus, this sample size is adequate to determine the congruency effect's presence or absence.

Stimuli, apparatus, and procedure. All stimuli, apparatus, and basic procedures in Experiment 4 were the same as those in

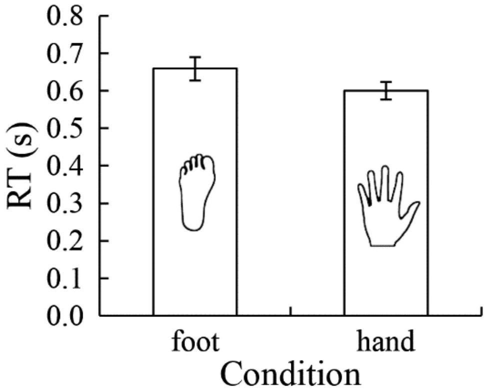

Figure 8. Means and *SE*s of RT for the simple orientation judgment task (Experiment 3).

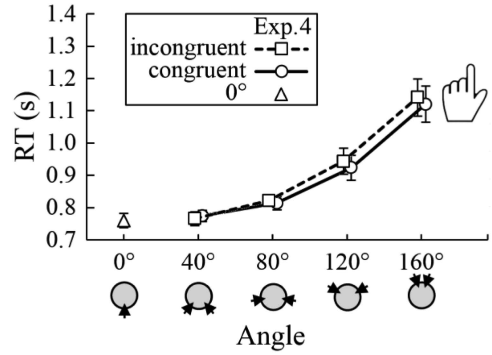

Figure 10. Means and *SE*s of RT data in Experiment 4.

Experiment 2, except that a finger response was used. As described in Figure 9, a keyboard used as a response device was placed on a table instead of the dual foot switch used in Experiment 2. Participants sat on a pipe chair at an 80-cm viewing distance to the monitor, placed one hand on the table with the index finger stretched and the thumb held by the other fingers, laid the index finger on the "down arrow (↓)" key, and kept the other hand on their laps. During the spatial perspective-taking task, participants responded by pressing the "left arrow (←)" key or "right arrow (→)" key, moving only the index finger. After each response, participants replaced the index finger in the original position. The response hand (left or right) was switched between two blocks, with the order counterbalanced across participants.

Results and Discussion

We conducted repeated-measures ANOVAs with two factors (congruency and angle) on RT and error data. For the RT analyses, we excluded error trials (3.2% of data) and trials that took longer than 2.58 s (= $M + 4\ SD$; 0.9% of data) and then calculated the mean RTs per cell for each participant. The mean RTs and errors across participants are shown in Figure 10 and Figure 11, respectively.

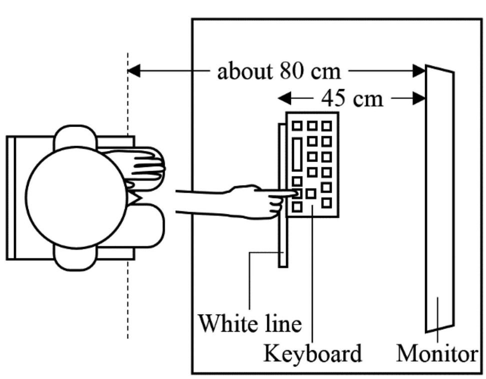

Figure 9. An overhead view of the setup in the finger condition (Experiment 4). Participants sat with the index finger of the left or right hand placed on the "down (↓)" key and responded by pushing the "left arrow (←)" key or "right arrow (→)" key. The other hand was placed in their laps.

The 2 × 4 ANOVA for RT data revealed a significant main effect of angle ($F(3, 69) = 59.69$, $\eta_p^2 = .722$, $p < .001$) but no main effect of congruency ($F(1, 23) = 1.76$, $\eta_p^2 = .071$, $p = .198$) and no interaction of angle and congruency ($F(1, 69) = 1.22$, $\eta_p^2 = .050$, $p = .309$). Post hoc *t* tests revealed a monotonic increase of RT with an increasing angle, showing significant differences for any pair of two consecutive angles (40 vs. 80°, $t(23) = 4.22$, $d = 0.45$, $p < .001$; 80 vs. 120°, $t(23) = 7.18$, $d = 0.84$, $p < .001$; 120 vs. 160°, $t(23) = 6.47$, $d = 0.89$, $p < .001$). In addition, a separate paired *t* test found no significant difference between 0 and 40° ($t(23) = 1.16$, $d = 0.10$, $p = .258$).

To clarify further whether relevance to whole-body movement was critical to the response congruency effect, we conducted a planned comparison of congruency-effect amounts in relevant (Experiments 1 and 2) versus irrelevant (Experiment 4) conditions at 120 and 160°. The 2 (relevance) × 2 (congruency) × 2 (angle) mixed-design ANOVA revealed a marginally significant interaction of relevance and congruency ($F(1, 70) = 3.97$, $\eta_p^2 = .054$, $p = .050$). This indicates that actions relevant to whole-body movement are necessary for the response congruency effect.

For error data, the 2 × 4 ANOVA revealed a significant main effect of angle ($F(3, 69) = 7.64$, $\eta_p^2 = .249$, $p = .001$) but no significant main effect of congruency ($F(1, 23) = 0.03$, $\eta_p^2 = .001$, $p = .867$) and no interaction of angle and congruency ($F(1, 69) = 0.11$, $\eta_p^2 = .005$, $p = .946$). Post hoc *t* tests revealed that more errors occurred at 160° than at 40° ($t(23) = 3.46$, $d = 0.71$, $p = .013$) and 80° ($t(23) = 3.05$, $d = 0.73$, $p = .028$) and at 120° than

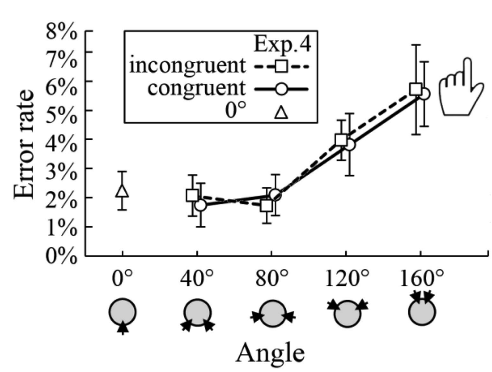

Figure 11. Means and *SE*s of error data in Experiment 4.

at 40° ($t(23) = 2.88$, $d = 0.50$, $p = .034$) but no other significant differences (all $ps > .095$).

In summary, finger responses did not lead to a congruency effect, unlike movement-related responses in Experiments 1 and 2. If participants in Experiment 4 used strategies other than embodied transformations because of a finger response, the congruency effect's absence might be attributed to a qualitative strategy shift. However, that is unlikely because, according to participants' introspective reports, the dominant strategy used by 62.5% (15 of 24) was still a concrete perspective-taking strategy (66.7% in Experiment 1; 62.5% in Experiment 2); a few participants (2 of 24; 8.3%) reported performing object rotation in some trials, just as in Experiments 1 (0.0%) and 2 (8.3%). None reported using a reversal strategy. Overall, these results support not the left-or-right account, but the motor simulation account as causing the congruency effect.

Experiment 5

Experiments 1 and 2 found the response congruency effect. As mentioned above, this effect might not be attributed to a kind of spatial S-R compatibility effect because the response congruency effect was not observed at lower angles (i.e., 40 and 80°). Nonetheless, Experiment 5 attempted to provide more direct evidence for ruling out this spatial S-R compatibility account for the response congruency effect. In this experiment, we manipulated the presentation position (left or right) of a stimulus itself, as well as the movement's direction (clockwise or counterclockwise). If the spatial S-R compatibility account were true, then RT would be shorter when the responding foot and the stimulus position were compatible (e.g., the left foot for a left stimulus) than when they were incompatible (e.g., the left foot for a right stimulus). Additionally, congruency between the rotational direction and the responding foot would have no or less effect on RTs. On the other hand, if the response congruency effect reflected the process of simulated whole-body movement during spatial perspective taking, then congruency between the movement's direction and the responding foot would contribute to the response congruency effect regardless of spatial S-R compatibility.

Method

Participants. Participants in Experiment 5 were 24 undergraduate and graduate students (mean age = 20.7 years; 12 female and 12 male; all right-footed). All had normal or corrected-to-normal vision, were naïve to the study's purpose, and received either prepaid cards for purchasing books or course credit for their participation. None had participated in previous experiments.

Stimuli, apparatus, and procedure. All stimuli, apparatus, and basic procedures in Experiment 5 were the same as those in Experiment 1, except for the following three differences. First, we created new stimuli by trimming both left and right edges of stimulus images in Experiments 1, 2, and 4, so new stimuli could be presented within the display's left or right half (see Figure 12). Second, stimuli's presentation position was randomly varied left to right from trial to trial. Stimuli were presented at the center of either the left or right display half. Third, we omitted the 0° condition to secure an adequate number of trials in limited experimental time. Thus, each block (for the left or right foot) consisted of 128 trials in random order: 8 angles × 2 presentation positions × 2 targets × 2 target positions × 2 repetitions.

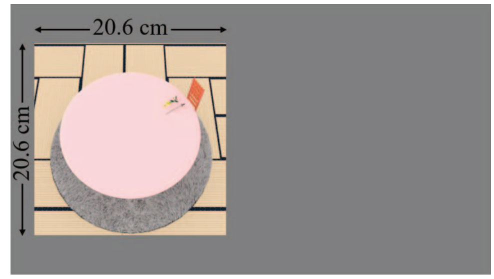

Figure 12. An example of a display showing a stimulus in Experiment 5. See the online article for the color version of this figure.

Results and Discussion

The basic analytical procedure was the same as those for Experiments 1, 2, and 4, except for addition of the new factor of spatial compatibility, defined by whether the presentation position (left or right) was compatible or incompatible with the responding foot (left or right). Thus, there were three orthogonal experimental factors: congruency (congruent or incongruent), spatial compatibility (compatible or incompatible), and angle (40, 80, 120, or 160°). We conducted repeated-measures ANOVAs with these three factors on RT and error data. For RT analyses, we excluded error trials (2.0% of data) and trials that took longer than 2.62 s (= $M + 4\ SD$; 1.1% of data) and then calculated mean RTs per cell for each participant. Mean RTs and errors across participants are shown in Figure 13 and Figure 14, respectively.

The 2 × 2 × 4 ANOVA for RT data revealed significant main effects of angle ($F(3, 69) = 127.82$, $\eta_p^2 = .848$, $p < .001$) and congruency ($F(1, 23) = 4.92$, $\eta_p^2 = .176$, $p = .037$). Post hoc t tests revealed monotonic increase of RT with increasing angle, showing significant differences for any pair of two consecutive angles (40 vs. 80°, $t(23) = 4.15$, $d = 0.14$, $p < .001$; 80 vs. 120°, $t(23) = 10.31$, $d = 0.55$, $p < .001$; 120 vs. 160°, $t(23) = 9.23$, $d = 0.84$, $p < .001$). More important, there was no main effect of spatial compatibility (1,139 and 1,134 ms for the compatible and incompatible conditions, respectively; $F(1, 23) = 1.93$, $\eta_p^2 = .078$, $p = .178$), suggesting that a spatial compatibility effect did not work in this case. An interesting find was that there was a significant interaction of congruency and spatial compatibility ($F(1, 23) = 10.60$, $\eta_p^2 = .316$, $p = .004$). To unfold this interaction, we conducted separate ANOVAs for compatible and incompatible conditions. Results revealed that the congruency effect was significant when the stimulus was presented on the opposite side of the responding foot ($F(1, 23) = 12.23$, $\eta_p^2 = .347$, $p = .002$), but not significant when the stimulus was presented on the same side of the responding foot ($F(1, 23) = 0.10$, $\eta_p^2 = .004$, $p = .759$). There were no other two-way and three-way interactions (congruency and angle, $F(3, 69) = 2.28$, $\eta_p^2 = .090$, $p = .098$; spatial compatibility and angle, $F(3, 69) = 0.45$, $\eta_p^2 = .019$, $p = .697$; congruency and spatial compatibility and angle, $F(3, 69) = 0.08$, $\eta_p^2 < .001$, $p = .945$).

For error data, the 2 × 2 × 4 ANOVA revealed a significant main effect of angle ($F(3, 69) = 11.57$, $\eta_p^2 = .335$, $p < .001$), but

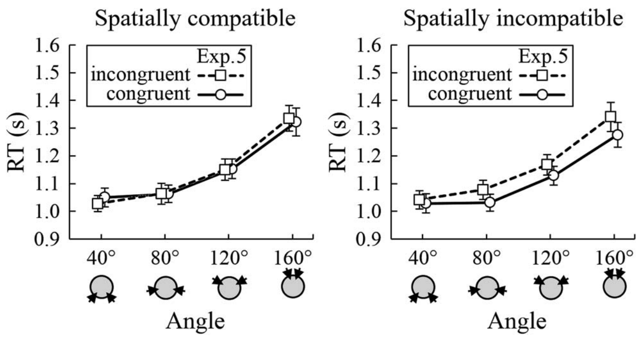

Figure 13. Means and *SE*s of RT data in Experiment 5.

neither of congruency ($F(1, 23) = 0.02$, $\eta_p^2 < .001$, $p = .902$), nor of compatibility ($F(1, 23) = 0.33$, $\eta_p^2 = .014$, $p = .571$). Post hoc t tests revealed that more errors occurred at 160° than at 40° ($t(23) = 5.03$, $d = 1.46$, $p < .001$) and 80° ($t(23) = 4.30$, $d = 1.18$, $p = .001$), at 120° than at 40° ($t(23) = 2.83$, $d = 0.83$, $p = .038$), and showed no other significant differences (all $ps > .050$). No interactions were significant (congruency and angle, $F(3, 69) = 1.25$, $\eta_p^2 = .051$, $p = .299$; spatial compatibility and angle, $F(3, 69) = 1.03$, $\eta_p^2 = .043$, $p = .382$; congruency and spatial compatibility, $F(1, 23) = 1.23$, $\eta_p^2 = .051$, $p = .279$; congruency and spatial compatibility and angle, $F(3, 69) = 1.25$, $\eta_p^2 < .051$, $p = .298$).

In summary, spatial S-R compatibility had no effect on performance in spatial perspective taking. In addition, the foot response congruent with the movement's direction shortened overall RT for spatial perspective taking only when the stimulus was presented on the opposite side of the responding foot. Although these results contradict the spatial S-R compatibility account, they also differ from our prediction of the motor simulation account in some ways. In this experiment, the response congruency effect was limited to the stimulus on the opposite side of the responding foot and was NOT limited to higher angle conditions. These unpredicted findings could probably be explained by considering trajectories from the participant's position to the chair position. Unlike Experiments 1, 2, and 4, the distance between the participant's position and the chair position in Experiment 5 depended on the rotational direction (clockwise or counterclockwise) even when the angle was the same. For example, when the stimulus was presented on the display's right side, the position of the clockwise (i.e., inward) 40° was closer to the participant than that of the counterclockwise (i.e., outward) 40°. This asymmetry of the imagined trajectory depending on the rotational direction may explain Experiment 5 results.

Suppose that you imagine moving to the right outward side of the right table. In this case, putting your right foot forward would make your body approach the table's left rather than right. Plus, rotating your body counterclockwise around the axis of your left leg would make your back turn to the right table. Thus, responses to the stimulus on the responding foot's same side would not necessarily be congruent with the imagined movement. On the other hand, putting your left foot forward would turn your whole body clockwise. Thus, when responding to the right table opposite to the responding foot, your action is congruent with movement to the left side of the table, but incongruent with movement to the right side. In this case, the movement strategy may be preferred even for objects at lower angles because of distance information that objects on the outward side are farther from you than objects on the inward side.

These results could also be interpreted as new counterevidence against the sensorimotor interference account. If putting the left or right foot forward mitigated interference between real and imagined perspectives, then the response congruency effect should occur regardless of the stimulus position because angular disparity between real and imagined perspectives was invariant regardless of whether the stimulus position was left or right. However, results of Experiment 5 showed the response congruency effect only for the stimulus presented on the opposite side of the responding foot. Therefore, findings in Experiment 5 support the motor simulation account for the response congruency effect, rejecting the sensorimotor interference account as well as the spatial S-R compatibility account.

General Discussion

Implication of the Response Congruency Effect

In accordance with Kessler and Thomson (2010), we hypothesized that spatial perspective taking is embodied as simulated whole-body movement. We found evidence that supported this hypothesis and also provided new suggestions about embodied processes of spatial perspective taking. Experiments 1 and 2 showed that spatial perspective taking at 120 and 160° was performed more efficiently when participants put forward a limb (left or right) congruent with the direction of an imagined movement (counterclockwise or clockwise) compared with incongruent movements. This finding conforms to Experiment 2 from Kessler and Thomson (2010), in which a participant's body posture was manipulated and the posture congruency effect was observed at 120 and 160°. In addition, a comparison of Experiments 1 and 2 showed that the response congruency effects in foot and hand conditions were indistinguishable, suggesting that simulated movement of a whole body, not of a specific body part, mediates the process of spatial perspective taking. This was further confirmed by Experiments 4 and 5. Experiment 4 used a response method irrelevant to a whole-body movement (i.e., index finger movements of either hand) and resulted in no congruency effects. Experiment 5 not only replicated the response congruency effect, but also rejected accounts from spatial S-R compatibility and sensorimotor interference.

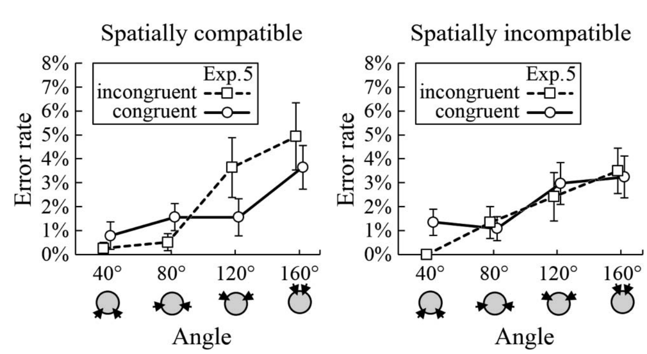

Figure 14. Means and *SE*s of error data in Experiment 5.

This response congruency effect suggests that a common neural basis underlies the execution of spatial perspective taking and motor simulation of a whole-body movement. This notion is evidenced by some previous research on brain activity during mental perspective transformations. For example, participants in Creem, Downs et al. (2001) performed in an functional magnetic resonance imaging (fMRI) environment a viewer rotation task similar to that used by Wraga et al. (2000); researchers found activation of the premotor area and other regions deemed to be involved in motor processing. Likewise, Wraga et al. (2005) used fMRI and, during a self-rotation task, observed that the left supplementary motor area was activated.[7] Additionally, using ERP mapping, Schwabe et al. (2009) reported activation of the posterior frontal cortex corresponding to the premotor area during a perspective transformation task.

However, some fMRI studies on perspective transformations showed no motor-related activations (e.g., Lambrey et al., 2012; Zacks et al., 2003). Such inconsistency might be attributed to two reasons. First, the use of a whole-body schema in spatial perspective taking does not seem obligatory but seems to be one possible strategy similar to motor strategies in mental object rotation (see Zacks, 2008, for a review). This notion is consistent with Creem, Downs et al.'s (2001) observation that some but not all participants showed premotor activation. The likelihood of using a movement strategy is probably affected by a given task's properties. For example, stimuli used by Lambrey et al. (2012) had as many as four objects on a table not aligned regularly; this seemed to impose somewhat-heavy cognitive demands on participants and prompted the use of different strategies than movement simulation. This issue is discussed in more detail below. Second, based on some limitations of fMRI measurements indicated by Kunz et al. (2009), participants' mobility is restricted in fMRI environments. In addition, a recent study showed that the supine posture itself, required by conventional fMRI studies, altered brain activities (Thibault, Lifshitz, & Raz, 2016). Such fMRI features might have nonnegligible effects on the strategy used in spatial perspective taking. To support this, some participants in the present study informally reported that they sometimes moved their faces or shoulders a bit during the task to ease their position judgment. This actual movement strategy is clearly impossible under the constrained fMRI condition.

Accordingly, neuroscientific data available so far must be interpreted with caution because they are as yet inadequate for determining whether movement simulation is actually involved in spatial perspective taking. The present behavioral study, however, has provided evidence supporting involvement of motor simulation in spatial perspective taking by examining effects of action related to whole-body movement. In addition, we suggested that movement simulation plays a significant role only at the beginning, not throughout spatial perspective taking (see above). Overall, our findings on the response congruency effect not only extended Kessler and Thomson's (2010) findings on the posture congruency effect, but also unveiled cognitive and motor processes of spatial perspective taking. Nonetheless, the present study is only the first step in investigating involvement of motor simulation, so our conclusion is still premature. Further studies from a broader perspective (including both behavioral and physiological viewpoints) are needed to draw a strong conclusion.

The response congruency effect also has implications for computational processes by which people know the direction or trajectory of simulated movement. There are at least two sources of information to determine the trajectory of simulated movement: One is the rotational angle of target objects and the other is a path between themselves and the target position on a stimulus image. This raises the question of whether people use information about the rotational angle only or about both the rotational angle and the path to calculate the trajectory of simulated movement. The present finding of the response congruency effect supports that both the sources were used because Experiment 5 demonstrated that the presentation position of stimuli modulated the response congruency effect despite the same rotational angles. In our paradigm, participants probably executed mental self translation and mental self rotation simultaneously by taking the smoothest and shortest path computed based on the prior information about the rotational angle and the path. However, what is the smoothest and shortest remains unclear. For example, does the layout of a scene (e.g., the presence of obstacles or the shape of a table) affect the trajectory of simulated movement? To clarify the nature of simulated whole-body movement, these issues should be addressed in future studies.

Implication of the Foot Advantage

A comparison of Experiments 1 and 2 showed that spatial perspective taking was processed more quickly when responses were made by a foot rather than by a hand. This is contrary to our common-sense notion, and the hand advantage was observed in simple orientation judgments (Experiment 3) and in a previous study (Pfister et al., 2014). Therefore, the foot advantage discovered here must be considered unique to spatial perspective taking and as evidence for movement simulation's contribution to spatial perspective taking.

While one is walking forward, visual input is continually updated, and an expanding optic flow occurs. Such a close link between locomotion and vision is well known. In this regard, some evidence indicates that walking alters visual perception and cognition. For example, Yabe et al. (Yabe & Taga, 2008; Yabe, Watanabe, & Taga, 2011) reported that a person walking on a treadmill perceived an ambiguous, apparent motion presented on the floor as moving backward, as if an optic flow actually existed, more frequently than did a person standing still on a treadmill. In another example, Kunz et al. (2009) demonstrated that the time for imagined walking without vision was closer to the time for real walking while participants were stepping in place than while they moved their arms circularly (irrelevant to walking) or merely standing still. Kunz et al. inferred that perceptual-motor conflict was eliminated by actual stepping, whereby a mental simulation of imagined walking became accurate. Consideration of these effects of foot movements on visuospatial representations, together with our hypothesis that spatial perspective taking involves simulated whole-body movement, leads to a prediction that spatial perspective taking would also likely be facilitated by concurrent foot movement.

[7] Although Wraga et al. (2005) supposed that activations of motor-related areas were not because of motor simulations but to demands of their high-level cognitive task, their finding is also compatible with the involvement of motor simulation.

To the best of our knowledge, no other phenomena are comparable with the foot advantage. Thus, we tentatively propose that the foot advantage in spatial perspective taking is because of the link between feet underpinning whole-body movement and visuospatial information. Investigations are underway to clarify the foot advantage's detailed mechanism and the conditions in which it occurs (partially reported in Muto, Matsushita, & Morikawa, 2016).

Although the foot advantage conforms to the notion that simulated whole-body movement underlies spatial perspective taking, its mechanism seems somewhat different from the response congruency effect. The major difference is that while the response congruency effect occurred only at high angle conditions (i.e., 120 and 160°), the foot advantage was seen at all angle conditions, including 0°, and it was most salient at 160°. At first glance, the foot advantage's ubiquity seems contradictory to the notion that low angle conditions required fewer perspective transformations than high angle conditions; thus, the response congruency effect was limited to the high angles (see above). Furthermore, while the response congruency effect that was independent of a responding body part (i.e., foot or hand) supports the involvement of a whole-body representation, the foot advantage clearly suggests a specific body part's role. Future studies should reveal the interconnection or independence between mechanisms of the response congruency effect and the foot advantage. Indeed, we have already undertaken such studies (e.g., Muto, Matsushita, & Morikawa, 2016).

Embodied Transformation as a Strategy

The present study has succeeded in demonstrating movement simulation's important role in spatial perspective taking. In this section, we discuss the extent to which our findings can be generalized to various situations, including real-life ones. Simulation of a whole-body movement is likely executed in limited situations instead of all situations. As described above, movement strategy seems more likely to be used when a given task emphasizes online rather than offline processing. This notion is supported by Gärling, Böök, Lindberg, and Arce's (1990) finding that estimations of elevation in a large-scale real environment based on a cognitive map can be accomplished without movement simulation such as "mental travel." To further understand strategy differences, we consider an alternative hypothesis postulated previously, that is, the sensorimotor interference account (e.g., Brockmole & Wang, 2003; May, 2004; Wang, 2005). According to the sensorimotor interference account, the angle effect of a spatial perspective-taking task stems not from cognitive loads of mental transformations but from the conflict between real and imagined perspectives. However, most previous findings regarded as evidence for this account can be interpreted without assuming sensorimotor interference. Rather, as described below, they exemplify strategy differences.

One of the most compelling pieces of evidence for the sensorimotor interference account is that allowing participants time to complete transformations in advance did not attenuate the angle effect (May, 2004; Wang, 2005). However, this is also accounted for by a strategy change to avoid large demands on working memory (Kessler & Thomson, 2010; see above). In another example, Brockmole and Wang (2003) found that imagined perspective change required less effort when participants changed perspective across environments (e.g., from facing west in the middle of a building to facing north in the middle of their office in the building) than when they changed perspective within a single environment (e.g., from facing north to facing east in the middle of their office). Although Brockmole and Wang (2003) attributed the benefit in across-versus-within conditions to reduced conflict between initial and updated perspectives, this finding can again be explained from another viewpoint. In the across condition, the initial perspective seems completely unnecessary for position judgment from the new perspective; thus, participants could directly recall the new perspective instead of changing perspective. Mou and McNamara (2002) demonstrated that humans' representations of spatial layouts can be abstractly encoded regardless of their actual visual experiences (i.e., representations from a never-seen-before viewpoint can be recalled). In summary, these previous findings related to the sensorimotor interference account can be interpreted as evidence of diverse strategies for spatial problems involving offline processing rather than as evidence of sensorimotor interference effects. The whole-body movement strategy is likely not necessarily suitable for these situations.

Even in the task we used, simulation of whole-body movement might not be obligatory, but one possible strategy. For example, reversal strategy (i.e., reversing the left/right position of objects) could also be used for high-angle conditions even though we eschewed this strategy in our experiments. Consistently, Kessler and Wang (2012) reported that female participants were more likely than male participants to use embodiment strategy for spatial perspective taking. Kessler and Wang (2012) inferred that this gender difference occurred because men adopted "rule-based" strategies such as the reversal strategy more often than women. To examine whether such a gender difference was also obtained in our results, we reanalyzed data from Experiments 1, 2, and 5 by including the gender factor, following Kessler and Wang (2012). The analysis included only conditions in which the congruency effect was detected (i.e., 120 and 160° angles of Experiments 1 and 2, and the spatially incompatible condition of Experiment 5). We subtracted mean RTs for congruent conditions from those for incongruent conditions per participant and treated RT difference as the index of the response congruency effect. A 2 (gender) × 3 (experiment) between-participants ANOVA revealed that the congruency effect of male participants (66 ms on average) was equivalent to that of female participants (62 ms on average; $F(1, 66) = 0.47$, $\eta_p^2 = .007$, $p = .497$). The interaction of gender and experiment ($F(2, 66) = 1.07$, $\eta_p^2 = .031$, $p = .348$) and the main effect of experiment ($F(2, 66) = 0.51$, $\eta_p^2 = .015$, $p = .602$) were also insignificant. This absence of gender difference was probably because of our eschewal of reversal strategy. Therefore, there seem to be multiple strategies for spatial perspective taking, like object-based mental rotation (Flusberg & Boroditsky, 2011; Kosslyn, Thompson, Wraga, & Alpert, 2001). Future research must determine conditions in which a certain strategy is more likely to be used.

Nonetheless, as Kessler and Wang (2012) stated, the embodiment strategy seems to be the natural, default method of spatial perspective taking because the vast majority of participants showed a posture or response congruency effect. Specifically, 81 of 96 participants (84%) in Kessler and Thomson (2010) showed a posture congruency effect indicated by a positive value of RT differences between incongruent and congruent conditions (Kessler & Wang, 2012). With the same criterion, a comparable proportion of our participants (81%, 58 of 72) exhibited the response congruency effect. A distinct feature of our task was the emphasis on online rather than offline processing (i.e., minimal demands on long-term memory); such a feature is common to everyday spatial problem solving (Freksa & Schultheis, 2014). Therefore, such embodied transformations are likely to be performed in real-life situations as well.

Evolutionary and Developmental Origins

We demonstrated that simulated whole-body movement subserves the online process of spatial perspective taking, unlike mental object rotation related to hand movements (e.g., Gardony et al., 2014; Wexler et al., 1998; Wohlschläger & Wohlschläger, 1998). This difference might reflect different evolutionary histories between perspective and object-based transformations and supports the multiple-systems framework (e.g., Zacks & Michelon, 2005; Zacks & Tversky, 2005). As discussed by Kessler and Thomson (2010), the embodied nature of spatial perspective taking can be considered a stepping stone from actual to imaginary movements. This notion is consistent with previous findings in comparative psychology, for example, that great apes are incapable of spatial perspective taking (Tomasello et al., 2005) but can physically move to a human's position to know what the human is looking at (Bräuer, Call, & Tomasello, 2005). However, currently available findings on nonhuman species are too indirect and few to draw such a conclusion.

To determine whether spatial perspective taking is unique to humans, the role of language should also be considered because judgment of spatial directions is closely linked to spatial terms (e.g., Franklin & Tversky, 1990; Imai, Nakanishi, Miyashita, Kidachi, & Ishizaki, 1999; Kessler & Rutherford, 2010). Kessler and Rutherford (2010) demonstrated that the posture congruency effect occurred whether judgment was made by key or verbal responses. However, even when the response modality was nonverbal, people could rely internally on linguistic processing for spatial perspective taking. Unfortunately, to the best of our knowledge, there has been no research on spatial perspective taking of humans without egocentric direction terms in their language. To reveal the evolutionary history of spatial perspective taking, future studies should also focus on linguistic/cultural factors.

Because spatial perspective taking involves motor processing, we also must focus on how the spatial perspective-taking ability develops in human children. Although mental object rotation ability is known to develop with action experience (e.g., Frick & Möhring, 2013), any developmental link between spatial perspective taking and motor skills is still unknown. Huttenlocher and Presson (1973) reported that 4th-grade children who had difficulty imagining the appearance of a hidden array from new perspectives showed better performances when they were allowed to move physically to the new perspective's position, suggesting that actual movement precedes imagined movement developmentally. In addition, the onset of self-produced locomotion (i.e., crawling and walking) helps children develop nonegocentric representations of locations (Needham & Libertus, 2011). Given these reports, the spatial perspective-taking ability might be related to walking experience. Consistent with this view, Creem, Wraga, et al. (2001) argued that the advantage of viewer rotation over array rotation on the ground plane is because of the daily experience of walking under gravity.

In summary, our findings on movement simulation's role in spatial perspective taking are informative in terms of its evolutionary and developmental origins. For example, the fact that a hand response produced as much congruency effect as a foot response suggests that our arms remain integrated into the human brain's locomotor system even several million years after our ancestors became bipedal. Spatial perspective taking should be explored from interdisciplinary perspectives to understand these issues more comprehensively.

References

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. http://dx.doi.org/10.1016/S0926-6410(96)00073-0

Bräuer, J., Call, J., & Tomasello, M. (2005). All great ape species follow gaze to distant locations and around barriers. *Journal of Comparative Psychology, 119,* 145–154. http://dx.doi.org/10.1037/0735-7036.119.2.145

Brockmole, J. R., & Wang, R. F. (2003). Changing perspective within and across environments. *Cognition, 87,* B59–B67. http://dx.doi.org/10.1016/s0010-0277(02)00231-7

Carpenter, M., & Proffitt, D. R. (2001). Comparing viewer and array mental rotations in different planes. *Memory & Cognition, 29,* 441–448. http://dx.doi.org/10.3758/BF03196395

Chi, Y. Y., Gribbin, M., Lamers, Y., Gregory, J. F., III, & Muller, K. E. (2012). Global hypothesis testing for high-dimensional repeated measures outcomes. *Statistics in Medicine, 31,* 724–742. http://dx.doi.org/10.1002/sim.4435

Creem, S. H., Downs, T. H., Wraga, M., Harrington, G. S., Proffitt, D. R., & Downs, J. H. (2001). An fMRI study of imagined self-rotation. *Cognitive, Affective & Behavioral Neuroscience, 1,* 239–249. http://dx.doi.org/10.3758/CABN.1.3.239

Creem, S. H., Wraga, M., & Proffitt, D. R. (2001). Imagining physically impossible self-rotation: Geometry is more important than gravity. *Cognition, 81,* 41–64. http://dx.doi.org/10.1016/S0010-0277(01)00118-4

Dalecki, M., Hoffmann, U., & Bock, O. (2012). Mental rotation of letters, body parts and complex scenes: Separate or common mechanisms? *Human Movement Science, 31,* 1151–1160. http://dx.doi.org/10.1016/j.humov.2011.12.001

Easton, R. D., & Sholl, M. J. (1995). Object-array structure, frames of reference, and retrieval of spatial knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 483–500. http://dx.doi.org/10.1037/0278-7393.21.2.483

Erle, T. M., & Topolinski, S. (2015). Spatial and empathic perspective-taking correlate on a dispositional level. *Social Cognition, 33,* 187–210. http://dx.doi.org/10.1521/soco.2015.33.3.187

Flusberg, S. J., & Boroditsky, L. (2011). Are things that are hard to physically move also hard to imagine moving? *Psychonomic Bulletin & Review, 18,* 158–164. http://dx.doi.org/10.3758/s13423-010-0024-2

Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General, 119,* 63–76. http://dx.doi.org/10.1037/0096-3445.119.1.63

Freksa, C., & Schultheis, H. (2014). Three ways of using space. In D. R. Montello, K. Grossner, & D. G. Janelle (Eds.), *Space in mind: Concepts for spatial learning and education* (pp. 31–48). Cambridge, MA: MIT Press.

Frick, A., & Möhring, W. (2013). Mental object rotation and motor development in 8- and 10-month-old infants. *Journal of Experimental Child Psychology, 115,* 708–720. http://dx.doi.org/10.1016/j.jecp.2013.04.001

Gardony, A. L., Taylor, H. A., & Brunyé, T. T. (2014). What does physical rotation reveal about mental rotation? *Psychological Science, 25,* 605–612. http://dx.doi.org/10.1177/0956797613503174

Gärling, T., Böök, A., Lindberg, E., & Arce, C. (1990). Is elevation encoded in cognitive maps? *Journal of Environmental Psychology, 10,* 341–351. http://dx.doi.org/10.1016/S0272-4944(05)80033-7

Hamilton, A. F., Brindley, R., & Frith, U. (2009). Visual perspective taking impairment in children with autistic spectrum disorder. *Cognition, 113,* 37–44. http://dx.doi.org/10.1016/j.cognition.2009.07.007

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence, 32,* 175–191. http://dx.doi.org/10.1016/j.intell.2003.12.001

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6,* 65–70.

Huttenlocher, J., & Presson, C. C. (1973). Mental rotation and the perspective problem. *Cognitive Psychology, 4,* 277–299. http://dx.doi.org/10.1016/0010-0285(73)90015-7

Huttenlocher, J., & Presson, C. C. (1979). The coding and transformation of spatial information. *Cognitive Psychology, 11,* 375–394. http://dx.doi.org/10.1016/0010-0285(79)90017-3

Imai, M., Nakanishi, T., Miyashita, H., Kidachi, Y., & Ishizaki, S. (1999). The meanings of FRONT/BACK/LEFT/RIGHT. *Cognitive Studies, 6,* 207–225.

Inagaki, H., Meguro, K., Shimada, M., Ishizaki, J., Okuzumi, H., & Yamadori, A. (2002). Discrepancy between mental rotation and perspective-taking abilities in normal aging assessed by Piaget's three-mountain task. *Journal of Clinical and Experimental Neuropsychology, 24,* 18–25. http://dx.doi.org/10.1076/jcen.24.1.18.969

Kessler, K., & Rutherford, H. (2010). The two forms of visuo-spatial perspective taking are differently embodied and subserve different spatial prepositions. *Frontiers in Psychology, 1,* 213. http://dx.doi.org/10.3389/fpsyg.2010.00213

Kessler, K., & Thomson, L. A. (2010). The embodied nature of spatial perspective taking: Embodied transformation versus sensorimotor interference. *Cognition, 114,* 72–88. http://dx.doi.org/10.1016/j.cognition.2009.08.015

Kessler, K., & Wang, H. (2012). Spatial perspective taking is an embodied process, but not for everyone in the same way: Differences predicted by sex and social skills score. *Spatial Cognition and Computation, 12,* 133–158. http://dx.doi.org/10.1080/13875868.2011.634533

Kosslyn, S. M., Thompson, W. L., Wraga, M., & Alpert, N. M. (2001). Imagining rotation by endogenous versus exogenous forces: Distinct neural mechanisms. *NeuroReport, 12,* 2519–2525. http://dx.doi.org/10.1097/00001756-200108080-00046

Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition, 29,* 745–756. http://dx.doi.org/10.3758/BF03200477

Kunz, B. R., Creem-Regehr, S. H., & Thompson, W. B. (2009). Evidence for motor simulation in imagined locomotion. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1458–1471. http://dx.doi.org/10.1037/a0015786

Lambrey, S., Doeller, C., Berthoz, A., & Burgess, N. (2012). Imagining being somewhere else: Neural basis of changing perspective in space. *Cerebral Cortex, 22,* 166–174. http://dx.doi.org/10.1093/cercor/bhr101

May, M. (2004). Imaginal perspective switches in remembered environments: Transformation versus interference accounts. *Cognitive Psychology, 48,* 163–206. http://dx.doi.org/10.1016/S0010-0285(03)00127-0

Michelon, P., & Zacks, J. M. (2006). Two kinds of visual perspective taking. *Perception & Psychophysics, 68,* 327–337. http://dx.doi.org/10.3758/BF03193680

Mou, W., & McNamara, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 162–170. http://dx.doi.org/10.1037/0278-7393.28.1.162

Muto, H., Matsushita, S., & Morikawa, K. (2016). Comparison between foot and hand responses for a spatial perspective-taking task. *International Journal of Psychology, 51,* 201. http://dx.doi.org/10.1002/ijop.12299

Needham, A., & Libertus, K. (2011). Embodiment in early development. *WIREs Cognitive Science, 2,* 117–123. http://dx.doi.org/10.1002/wcs.109

Pfister, M., Lue, J.-C. L., Stefanini, F. R., Falabella, P., Dustin, L., Koss, M. J., & Humayun, M. S. (2014). Comparison of reaction response time between hand and foot controlled devices in simulated microsurgical testing. *BioMed Research International, 2014,* 769296. http://dx.doi.org/10.1155/2014/769296

Presson, C. C. (1982). Strategies in spatial reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 243–251. http://dx.doi.org/10.1037/0278-7393.8.3.243

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1157–1165. http://dx.doi.org/10.1037/0278-7393.15.6.1157

Schultz, K. (1991). The contribution of solution strategy to spatial performance. *Canadian Journal of Psychology, 45,* 474–491. http://dx.doi.org/10.1037/h0084301

Schwabe, L., Lenggenhager, B., & Blanke, O. (2009). The timing of temporoparietal and frontal activations during mental own body transformations from different visuospatial perspectives. *Human Brain Mapping, 30,* 1801–1812. http://dx.doi.org/10.1002/hbm.20764

Schwartz, D. L., & Holton, D. L. (2000). Tool use and the effect of action on the imagination. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1655–1665. http://dx.doi.org/10.1037/0278-7393.26.6.1655

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171,* 701–703. http://dx.doi.org/10.1126/science.171.3972.701

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G. Reeve (Eds.), *Stimulus-response compatibility: An integrated perspective* (pp. 31–86). Amsterdam, the Netherlands: Elsevier. http://dx.doi.org/10.1016/S0166-4115(08)61218-2

Sugihara, K. (2015). Height reversal generated by rotation around a vertical axis. *Journal of Mathematical Psychology, 68–69,* 7–12. http://dx.doi.org/10.1016/j.jmp.2015.07.001

Surtees, A., Apperly, I., & Samson, D. (2013). Similarities and differences in visual and spatial perspective-taking processes. *Cognition, 129,* 426–438. http://dx.doi.org/10.1016/j.cognition.2013.06.008

Thibault, R. T., Lifshitz, M., & Raz, A. (2016). Body position alters human resting-state: Insights from multi-postural magnetoencephalography. *Brain Imaging and Behavior, 10,* 772–780. http://dx.doi.org/10.1007/s11682-015-9447-8

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28,* 675–691. http://dx.doi.org/10.1017/S0140525X05000129

Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition, 110,* 124–129. http://dx.doi.org/10.1016/j.cognition.2008.10.008

Waller, D. (2014). Embodiment as a framework for understanding environmental cognition. In D. R. Montello, K. Grossner, & D. G. Janelle (Eds.), *Space in mind: Concepts for spatial learning and education* (pp. 139–157). Cambridge, MA: MIT Press.

Wang, R. F. (2005). Beyond imagination: Perspective change problems revisited. *Psicológica, 26,* 25–38.

Wexler, M., Kosslyn, S. M., & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition, 68,* 77–94. http://dx.doi.org/10.1016/S0010-0277(98)00032-8

Wohlschläger, A., & Wohlschläger, A. (1998). Mental and manual rotation. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 397–412. http://dx.doi.org/10.1037/0096-1523.24.2.397

Wolbers, T., & Hegarty, M. (2010). What determines our navigational abilities? *Trends in Cognitive Sciences, 14,* 138–146. http://dx.doi.org/10.1016/j.tics.2010.01.001

Wraga, M., Creem, S. H., & Proffitt, D. R. (2000). Updating displays after imagined object and viewer rotations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 151–168. http://dx.doi.org/10.1037/0278-7393.26.1.151

Wraga, M., Shephard, J. M., Church, J. A., Inati, S., & Kosslyn, S. M. (2005). Imagined rotations of self versus objects: An fMRI study. *Neuropsychologia, 43,* 1351–1361. http://dx.doi.org/10.1016/j.neuropsychologia.2004.11.028

Yabe, Y., & Taga, G. (2008). Treadmill locomotion captures visual perception of apparent motion. *Experimental Brain Research, 191,* 487–494. http://dx.doi.org/10.1007/s00221-008-1541-3

Yabe, Y., Watanabe, H., & Taga, G. (2011). Treadmill experience alters treadmill effects on perceived visual motion. *PLoS ONE, 6,* e21642. http://dx.doi.org/10.1371/journal.pone.0021642

Zacks, J. M. (2008). Neuroimaging studies of mental rotation: A meta-analysis and review. *Journal of Cognitive Neuroscience, 20,* 1–19. http://dx.doi.org/10.1162/jocn.2008.20013

Zacks, J. M., & Michelon, P. (2005). Transformations of visuospatial images. *Behavioral and Cognitive Neuroscience Reviews, 4,* 96–118. http://dx.doi.org/10.1177/1534582305281085

Zacks, J. M., & Tversky, B. (2005). Multiple systems for spatial imagery: Transformations of objects and bodies. *Spatial Cognition and Computation, 5,* 271–306. http://dx.doi.org/10.1207/s15427633scc0504_1

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. http://dx.doi.org/10.1162/089892903770007399

Appendix

Preliminary Experiment

In the beginning of the article, we assumed that for a clockwise movement, participants would put the left side of their bodies (e.g., left foot) forward first and for a counterclockwise movement, the right side (e.g., right foot) first (see Figure 1). To confirm this assumption and to corroborate the finding of Experiment 1, we conducted the following preliminary experiment in a real situation. In this experiment, participants were asked to physically move along the edge of a round table in a real situation. We manipulated a moving direction (clockwise or counterclockwise) and rotational angles (40, 80, 120, or 160°). Because participants' initial position was unclear in typical computerized spatial perspective taking tasks (e.g., Kessler & Thomson, 2010; Michelon & Zacks, 2006), we also manipulated the participants' initial positions (near or far). If our assumption is true, participants would tend to initially move their left foot in the clockwise condition and their right foot in the counterclockwise condition.

Participants were 10 undergraduate and graduate students (mean age = 24.0 years; 5 female and 5 male; 9 right-footed and 1 left-footed). All had normal or corrected-to-normal vision, were naïve to the study's purpose and received pre-paid cards for purchasing books. None had participated in the Experiments 1–5.

This experiment was conducted in a lecture room (610 × 706 cm). Figure A1 shows the configuration of the room. A round table (70 cm high and 180 cm in diameter) was positioned at the center of the room. A pipe chair was set at one of eight positions around the table according to the angle condition (40, 80, 120, or 160° in the clockwise or counterclockwise direction). The distance between the circumference of the table and the front side of the chair was 40 cm. Two white starting lines were drawn on the floor, 50 cm (near condition) or 100 cm (far condition) away from the table. Participants' movements were recorded by a fixed video camera right behind their initial positions. The experiment was guided by tones from two speakers. The experimenter controlled the procedure by using a personal computer behind participants.

(*Appendix continues*)

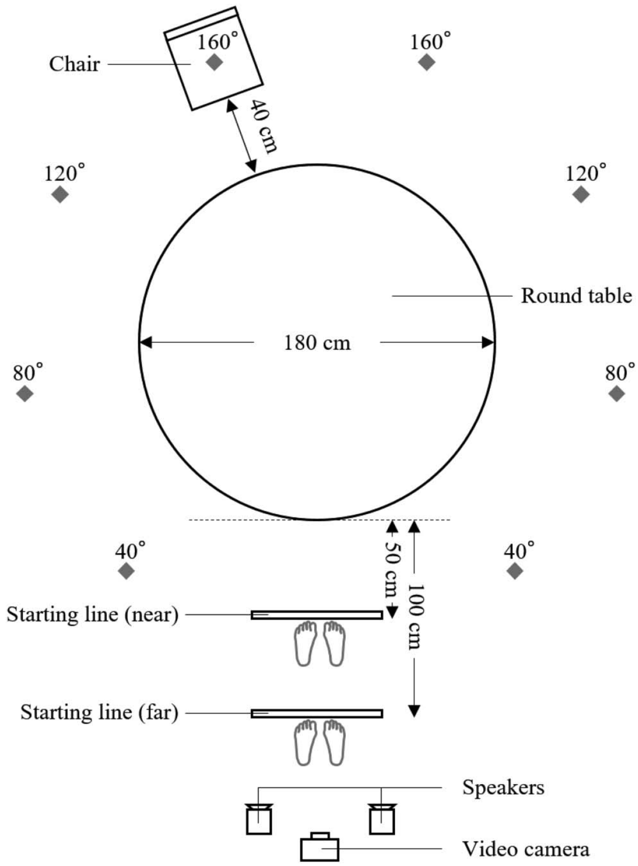

Figure A1. Overhead view of the apparatus used in the preliminary experiment. Participants stood in front of a starting line (near or far) and then walked down the shortest path to a chair and sat on it. Gray diamonds represent possible positions of a chair.

At the beginning of each trial, participants stood in front of one of the two starting lines (near or far) and closed their eyes. During this, the experimenter set a chair at one of the eight angle positions (40, 80, 120, or 160° in the clockwise or counterclockwise direction). Then, participants heard a 440-Hz tone and opened their eyes to confirm the position of a chair but remained standing still. Three seconds later, a 494-Hz tone was presented and participants had to quickly walk to and sit on the chair along the shortest path. After that, participants returned to the initial position and the next trial started.

Trials were blocked into two conditions of the initial positions (near or far) with the order counterbalanced across participants. Each block consisted of eight trials (for eight chair positions) in random order. Before the first block, participants completed two practice trials randomly selected from the first block to understand the experimental procedure.

By watching the recorded video, we judged whether each participant moved his or her left or right foot first away from the ground for each condition. Figure A2 shows rates of participants who moved their left (or right) foot first per condition. The results exhibited a clear pattern consistent with our assumption: Participants tended to move their left foot to start walking in the clockwise direction but right foot for the counterclockwise direction. To validate this, we conducted a 4 (angle; 40, 80, 120, or 160°) × 2 (direction; clockwise or counterclockwise) × 2 (initial position; near or far) repeated-measures ANOVA on first-moved foot (left foot = 0, right foot = 1). Consistent with our visual inspection, results showed that participants initially moved their right foot in the counterclockwise condition more frequently than in the clockwise condition ($F(1, 9) = 74.68$, $\eta_p^2 = .892$, $p < .001$).

Results also showed a significant main effect of angle ($F(3, 27) = 4.45$, $\eta_p^2 = .331$, $p = .036$), a significant interaction of angle and initial position ($F(3, 27) = 4.45$, $\eta_p^2 = .331$, $p = .036$), and a marginally significant interaction of direction and initial position ($F(1, 9) = 5.00$, $\eta_p^2 = .357$, $p = .052$). There were no main effect of initial position ($F(1, 9) = 0.13$, $\eta_p^2 = .014$, $p = .726$) and other two-way (angle and direction, $F(3, 27) = 1.54$, $\eta_p^2 = .146$, $p = .247$) and three-way interactions (angle, direction and initial position, $F(3, 27) = 1.54$, $\eta_p^2 = .146$, $p = .247$). These unpredicted significant patterns probably stemmed from an exceptional trend observed in the 40°-clockwise-far condition, in which as much as

(*Appendix continues*)

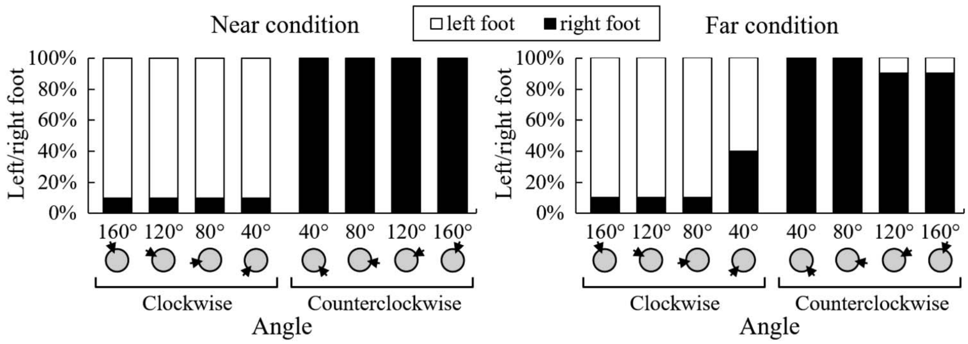

Figure A2. Rates of participants who moved their left (or right) foot first for each condition in the preliminary experiment. The rates of left and right feet are represented by white and black areas, respectively.

40% participants started with their right foot even for clockwise movement. In the far condition, pathways to the 40° positions were straight rather than curved (see Figure A1) and this might prompt participants to move their dominant foot (right foot for the majority) first in the same way as when they walked straight.

In summary, this experiment demonstrated our assumption that the first body movement depends on the movement direction. This finding is consistent with the response congruency effect found in Experiment 1, supporting the notion that spatial perspective taking involves whole-body motor simulation that corresponds to actual whole-body movement.

Received December 2, 2016
Revision received May 19, 2017
Accepted May 22, 2017 ■