December 2018Psychological Research
DOI: 10.1007/s00426-018-1131-3

Unidirectional Influence of Vision on Locomotion in Multimodal Spatial Representations Acquired from Navigation

Yu Du, Weimin Mou, & Lei Zhang

Department of Psychology, University of Alberta

Correspondence to:

Yu Du & Weimin Mou

Department of Psychology

University of Alberta

P-217, Biological Science Building

Edmonton, Alberta

Canada T6G 2E9

Email: du5@ualberta.ca and wmou@ualberta.ca

Abstract

Visual and idiothetic information are coupled in forming multimodal spatial representations during navigation (Tcheang, Bülthoff, & Burgess, 2011). We investigated whether idiothetic representations activate visual representations but not vice versa (unidirectional coupling) or whether these two representations activate each other (bidirectional coupling). In a virtual reality environment, participants actively rotated in place to face certain orientations to become adapted to a new vision-locomotion relationship (gain). In particular, the visual turning angle was equal to 0.7 times the physical turning angle. After adaptation, participants walked a path with a turn in darkness (idiothetic input only) or watched a video of the traversed path (visual input only). Then the participants pointed to the origin of the path. The participants who were presented with only idiothetic input showed that their pointing responses were influenced by the new gain (adaptation effect). By contrast, the participants who were presented with only visual input did not show any adaptation effect. These results suggest that idiothetic input contributed to spatial representations indirectly via the coupling, which resulted in the adaptation effect, whereas vision alone contributed to spatial representations directly, which did not result in the adaptation effect. Hence, the coupling between vision and locomotion is unidirectional.

Keywords: spatial representation; vision; idiothetic; amodal; multimodality

Imagine you are walking in a city. Usually you can see a change of your position in the environment with your eyes (e.g., relative to a specific visible building). Meanwhile, you can sense your body’s movement—your feet are down on the ground, your arms are swinging, and your body is moving forward. These two types of information are associated with two different systems. Visual information is processed by the visual system, whereas idiothetic information (i.e., inputs from internal sources including proprioceptive, vestibular systems and motor efference copy) is processed by the locomotion system (Loomis, Klatzky, Golledge, & Philbeck, 1999). Both systems construct spatial representations of the environment and then guide spatial navigation[1] (Tcheang, Bülthoff, & Burgess, 2011; see also Philbeck, Klatzky, Berhrman, Loomis, & Goodridge, 2001).

People use both systems in forming spatial memories (Avraamides, Sarrou, & Kelly, 2014; Mou, McNamara, & Zhang, 2013; Yamamoto & Shelton, 2005). For example, Avraamides, Sarrou, & Kelly (2014) had participants walked paths after they had visually previewed the room in one experiment. They found that participants used the information from walking paths (e.g., the walking direction) to determine the reference direction to encode objects’ locations. Yamamoto and Shelton (2005) demonstrated that participants who learned the object array by viewing the array and by walking a path through the locations of objects used both the visual information and their walking direction to establish two different reference systems to encode locations of objects in the same environment. These findings suggest influences from both idiothetic and visual cues in selection of reference systems for spatial memory.

---

[1] The distinction between visual and locomotion systems here is similar to the distinction between piloting and path integration in the literature (e.g., Chen et al., 2017) except that optic flow may be included in the visual system in the former but in path integration in the latter.

People also use the spatial representations in both systems to guide their navigation (Campos, Butler, & Bülthoff, 2014; Chen, McNamara, Kelly, & Wolbers, 2017; Ellard & Shaughnessy, 2003; Nardini, Jones, Bedford, & Braddick, 2008; Zhao & Warren, 2015). For example, in Tcheang et al. (2011, Experiment 1), participants after walking a two-leg path (outbound path) were asked to walk back to the origin of the path in a straight line (inbound path). In this triangle completion task, both visual and idiothetic cues were presented in the outbound path whereas only the idiothetic cues were presented in the inbound path. Tcheang et al. used an immersive virtual reality environment to change the relationship between visual and idiothetic inputs (gain) in the outbound path such that the visual rotation angle was 0.7 or 1.3 of the rotation angle indicated by idiothetic information. Therefore, these two cues indicated different turning angles (the gain was 0.7 or 1.3) for each outbound path. The resulting inbound walking direction was in the middle of the inbound walking directions predicted by individual cues. These findings suggest influences from both idiothetic and visual cues in estimating the origin of the outbound path. The resulting estimate might be a weighted average of the estimates based on either of the idiothetic and visual cues (i.e., cue combination). It is hypothesized that the weight of using cues in cue combination is determined by the reliability of different cues. The larger weight is assigned to the more reliable cue in cue combination (Cheng, Shettleworth, Huttenlocher, & Rieser, 2007).

Most strikingly, Tcheang et al. (2001) also demonstrated that there were both the visual and idiothetic influences even when there is only direct idiothetic information but no direct visual information during the triangle completion task. In their Experiment 2, participants first adapted to the gain of 0.7. Then they conducted the triangle completion task in darkness. Their spatial representations of the outbound path were inferred from their inbound walking direction.

The results showed that their representations of the outbound path, i.e., their estimation of the turning angle of the path, were affected by the gain, indicating that participants recalled the visual information from memory in forming the spatial representations of the outbound path, even when only idiothetic information was directly available in the outbound path as they walked in darkness. Tcheang et al. therefore proposed that the representations of these two systems are integrated into a multimodal spatial representation. Hence, even only one single modality input (e.g., idiothetic input) available during navigation can activate the multimodal spatial representation, demonstrating the influence of the modality that does not have the direct input.

Before Tcheang et al. (2011), the tight coupling relationship between vision and locomotion had been discovered in previous studies (for a review, see Rieser & Pick, 2007). In the original study (Rieser, Pick, Ashmead, & Garing, 1995), sighted participants adapted to a new relationship between visual and idiothetic inputs by walking on a treadmill towed forward by a tractor, creating a change in perceived walking speed (higher or lower than the actual one). In the subsequent test, they walked forward without vision to a previously-viewed target. The results indicated that their estimation of the body-target distance shown by their response distance was influenced by the new vision-locomotion relationship. Such an effect was also found in rotational locomotion with a similar experimental procedure (Pick, Rieser, Wagner, & Garing, 1999; Rieser et al., 1995). These findings suggest that vision is coupled with locomotion to calibrate locomotion, and changing the usual coupling between vision and locomotion would result in locomotor recalibration (Ziemer et al., 2013).

The important novel demonstration in Tcheang et al. (2011) is that the vision-locomotion coupling was activated even though the spatial task did not require the involvement of vision during the post-adaptation test as participants walked the path in the darkness. This novel demonstration is attributed to the triangle completion task used in their experiment. The triangle completion task does not require any direct visual input so the observed influence of visual information should be attributed to visual representations retrieved from memories. Moreover, the triangle completion task requires participants to infer the inbound path from the outbound path rather than directly replicate the outbound path. If they had used a task of reproducing the outbound path, participants would not necessarily have used the multimodal representation of the outbound path. Instead participants would have just used the idiothetic representation in the locomotion system to sufficiently reproduce the outbound path. Therefore, the triangle completion task provides opportunities to examine the multimodal spatial representation in memories.

Although the coupling of visual and idiothetic information leads to multimodal representations as indicated by Tcheang et al. (2011), it is not clear how vision and locomotion systems are coupled in forming multimodal spatial representations. We propose two models to conceptualize how vision and locomotion systems are coupled in forming multimodal spatial representations of a traversed path (see Figure 1).

Hypothesis 1: Unidirectional model

The visual system calibrates the locomotion system but not vice versa in forming spatial representations. Visual representations contribute to spatial representations only directly without activating the coupling relationship between vision and locomotion systems, whereas idiothetic representations indirectly contribute to spatial representations via the coupling[2]. Thus, the coupling is asymmetrical or unidirectional.

---

2 This model does not have a clear claim on whether idiothetic representations also directly contribute to spatial representations without activating the coupling relationship between vision and locomotion systems.

Hypothesis 2: Bidirectional model

The visual system and the locomotion system mutually calibrate each other. Both visual and idiothetic representations contribute to spatial representations directly. Meanwhile, both visual and idiothetic representations indirectly contribute to spatial representations by activating the coupling. Thus, the vision-locomotion coupling is bidirectional.

[Figure 1]

The current literature cannot clearly differentiate these two models. We speculate two approaches to differentiate these two models.

The first approach to differentiating these two models is based on the assumption that vision-locomotion coupling shares the same mechanism as cue interaction in forming spatial memory and in navigation. In particular, if vision is dominant over the idiothetic cues when both cues are available in cue interaction, vision is also dominant over the idiothetic cues in the vision-locomotion coupling. In other words, the visual system calibrates the locomotion system but the visual system is not calibrated by the locomotion system. Thus, the vision-locomotion coupling should be unidirectional (i.e., the visual system calibrate the locomotion system but not vice versa). As a consequence, the visual representation only directly contributes to the spatial representation whereas the idiothetic representation indirectly contributes to the spatial representation via the coupling relationship. In contrast, if both the visual and idiothetic cues are used when both cues are available in cue interaction, then the two systems mutually calibrate each other. Thus, the vision-locomotion coupling should be bidirectional. As a result, both visual and idiothetic representations contribute to the spatial representation directly and indirectly via the coupling relationship. Because both visual and locomotion systems were used in forming spatial memories (e.g., Avraamides et al., 2014) and in navigation (e.g., Tcheang et al., 2011, Experiment 1), the first approach suggests that the vision-locomotion coupling should be bidirectional.

The second approach to differentiate these two models assumes that the cue interaction and vision-locomotion coupling rely on different mechanisms. While the cue interaction (i.e., cue combination) might be determined by relative reliability of cues (Cheng et al., 2007), the vision-locomotion coupling might primarily rely on which system can directly perceive the environmental change as a consequence of locomotion.

The visual system, as an external source, can directly perceive the environmental change as a consequence of individuals' locomotion (Rieser, 1999). By contrast, idiothetic information, from an internal source, may only indirectly indicate the environmental change as a consequence of individuals' locomotion. According to the assumption that the fundamental modality of coupling is the one which can directly perceive the environmental change as a consequence of locomotion, vision is fundamental, whereas idiothetic information is specified in terms of visual information in the coupling between vision and locomotion systems. Because of the coupling, people who walk with eyes closed can still update their spatial relations relative to the environment (Rieser, 1989). Specifically, idiothetic information during locomotion can activate the coupling between vision and locomotion. Therefore, participants still can mentally *see* the change of their position in the environment as a consequence of their locomotion in the environment. The coupling emerges gradually through sufficient experience of individuals coordinating their actions with the environment; therefore the impaired performance of early-blind people in some spatial tasks requiring dynamic spatial updating may result from an inadequacy of early visual experience (Rieser & Pick, 2007).

There is also some evidence supporting the notion that the vision-locomotion coupling might primarily rely on which system can directly perceive the environmental change as a consequence of locomotion (e.g., Bruggeman, Zosh, & Warren, 2007; Hay, Pick, & Ikeda, 1965; Harris, 1963). Bruggeman, Zosh, and Warren (2007) reported that after participants adapted to a mismatch between the moving headings indicated by optic flow and their idiothetic cues, the idiothetic cues activated the new coupling of vision and locomotion systems whereas optic flow did not when participants walked with the original gain. In addition, neural evidence showed that although head direction cells are found firing several days before eye opening in the rat pups, only after eye opening do the signals become stable, indicating that visual inputs calibrate the idiothetic input in the head direction system (Tan, Bassett, O'Keefe, Cacucci, & Wills, 2015). Therefore, according to the second approach, the vision-locomotion coupling should be unidirectional.

Thus, the coupling relationship between vision and locomotion systems in the second experiment of Tcheang et al. (2011) could be bidirectional or unidirectional according to the two different theoretical assumptions mentioned in the two above approaches.

To our knowledge, the empirical evidence in the literature cannot distinguish between the bidirectional and unidirectional models. More specifically, the findings of Tcheang et al. (2011) could not distinguish between these two models. As illustrated in Figure 1, both models stipulate that idiothetic representations activate visual representations (see the blue arrow from idiothetic representation to visual representation). Therefore, when participants walked in darkness after adaptation, the idiothetic input and idiothetic representations would have used the vision-locomotion coupling to activate visual representations according to both models. As a result, the new vision-locomotion coupling that had been developed during the adaptation phase affected spatial representations even when participants walked in darkness.

To differentiate these two models, the current study investigated whether both visual input and idiothetic input alone can use the vision-locomotion coupling to form spatial representations. After adaptation to a new vision-locomotion relationship, participants walked in darkness (idiothetic input only) or watched a video of a traversed path from the first-person view (visual input only). Their spatial memory of a walking path was measured with a triangle completion task, in which after walking a path, participants pointed to the origin of the path (see Figure 2). Both models predict an adaptation effect in the condition with only idiothetic input as shown in Tcheang et al. (2011). However, these two models have different predictions for the condition with only visual input. According to the unidirectional model (Figure 1, left panel), the visual input and visual representation do not activate idiothetic representations via the coupling in forming spatial representations. Therefore the unidirectional model predicts no adaptation effect. By contrast, according to the bidirectional model, the visual input and visual representation also use the new vision-locomotion coupling to activate idiothetic representations (see the red arrow from the visual representation to the idiothetic representation on the right panel of Figure 1) in forming spatial representations. Therefore the bidirectional model predicts an adaptation effect.

Figure 2 illustrates the predictions of response pointing to the origin with a rotation gain of 0.7 in the current study. A rotation gain of 0.7 is a gain with which participants’ rotation angle indicated by the visual input is 0.7 times of the rotation angle indicated by the idiothetic input during participants’ turning. For the condition with only idiothetic input (i.e., walking, Figure 2A, 2B), both models predict that the idiothetic input would also activate visual representations in forming spatial representations (see Figure 1). In particular, with a gain smaller than 1, the turning angle indicated by visual representations is smaller than that indicated by idiothetic representations. As visual representations as well as idiothetic representations affect spatial representations, the turning angle (*A*) in spatial representations should be underestimated. With the gain of 0.7, as the turning angle (*A*) would be underestimated, the participant's estimated heading (H') would also be underestimated; therefore the response pointing direction (the angle between the direction from P to O' and the heading of H) would be larger than the correct direction (the angle between the direction from P to O and the heading of H).

For the condition with only visual input (i.e., transportation, Figure 2C), the bidirectional model predicts the visual input would also activate idiothetic representations in forming spatial representations (see Figure 1). Therefore, the visual input would also use the vision-locomotion coupling, causing an adaptation effect in the response pointing direction. In particular, with a gain smaller than 1 (i.e., the rotation angle indicated by the visual input is smaller than that indicated by the idiothetic input), the turning angle in the spatial representations of the path should be overestimated. With the gain of 0.7, as the turning angle (*A*) would be overestimated, the participant's estimated heading (H') would also be overestimated; therefore the response pointing direction would be smaller than the correct direction. However, the unidirectional model predicts that visual input would not activate idiothetic representations in forming spatial representations. Therefore, the responding direction would be the same as the correct direction.

[Figure 2]

It is important to note that to examine the adaptation effect in the condition with only visual input, we should ensure that participants' responses would not involve any idiothetic input sensitive to the new gain. In the study of Tcheang et al. (2011), participants turned their bodies to face and then walked back to the origin of the path. Therefore, the physical body rotation of the participants also involved idiothetic input sensitive to the new gain. Consequently, the actual turning angle in response was larger than that predicted by the spatial representation of the path (see Equation 1 of Tcheang et al., 2011). To remove the idiothetic input sensitive to the new gain during response, we asked participants to point to the origin with their fingers instead of asking them to turn their bodies to face the origin. We note that pointing also involves idiothetic information. However, such idiothetic information from pointing should not be sensitive to the new gain between vision and body turning because turning and pointing are functionally different from each other. Body turning causes co-variations in optic flow and idiothetic systems whereas pointing without head/body rotations does not cause such co-variations. As the gain between perception and action cannot be transferred between functionally different actions (Rieser et al., 1995), pointing with a finger without body rotation would not involve idiothetic input sensitive to the new gain between vision and body turning.

Table 1 illustrates the hypothetical differences between the responses with body turns and those with pointing when participants walk a course of two legs of equal length and the adaptation gain is 0.7. Examples with the turning angles between the legs being 63°, 90°, and 117° are also enclosed in the table.

[Table 1]

Experiment 1

The purpose of Experiment 1 was to replicate the findings of Tcheang et al. (2011) using the response of pointing instead of body turning before we differentiated the unidirectional model from the bidirectional model. It is not clear whether the adaptation effect can be replicated with the response of pointing instead of body turning. According to the theoretical analysis illustrated in Table 1, the adaptation effect based on pointing (*FP* in Table 1) appears much closer to the null adaptation effect (i.e., *FP* = 1) than that based on body turning (*FT* in Table 1). In particular, the theoretical difference ratio is overall 10% collapsed across turning angles (6%, 10%, and 14% for turning angles of 63º, 90º and 117º, respectively) in the former (i.e., (FP-C)/C) but is overall 58% collapsed across turning angles (52%, 57%, and 63% for turning angles of 63º, 90º and 117º respectively) in the latter (i.e., (FT-C)/C).

Moreover, although the theoretical difference ratio based on turning is overall 58%, the observed different ratio on turning in Experiment 2 of Tcheang et al. (2011) was only approximately 15% (8%, 1.14%, and 1.22% for turning angles of 63º, 90º and 117º respectively from their Figure 2C), indicating that the observed gain effect is much smaller than the theoratical one in the triangle completion task. Hence it is very likely that we could not observe any signficant adaption effect using the response of pointing that only produced the theoretical gain effect of 10%. If the null adaptation effect is obtained using the response of pointing in the idiothetic-input-only condition, then the null adaptation effect might also be expected in the visual-input-only condition regardless of spatial representations (bidirectional or unidirectional).

In Experiment 1, participants adapted to a new rotation gain of 0.7 (under-gain) in addition to the gain of 1 (normal-gain). These two gains were used just as in Tcheang et al. (2011, Experiment 2). In the testing phase, participants were asked to walk a path with two equally long legs in darkness as in Tcheang et al. (2011) and point to the origin’s direction. Both models predict that their pointing responses would be influenced by the new gain. In particular, Fractional Pointing (*FP*) in the under-gain condition would be larger than that in the normal-gain condition.

Method

Participants. A total of 33 university students participated in Experiment 1. After exclusion (see results section), 31 participants (18 female) were included in data analysis. All participants received course credit in an introductory psychology course for their participation.

In the second experiment of Tcheang et al. (2011), Cohen's *d* of the gain effect was about 1.27[3] (and the observed power was .99). As participants in the current study used pointing as the response instead of the turning response used in Tcheang et al. (2011), the effect size should decrease significantly. We planned to use about 32 participants instead of the 20 participants used in Tcheang et al. (2011) to address the reduced effect size. Assuming that the effect size decreased by half of what was obtained by Tcheang et al. (i.e., 0.64), the power at the .05 level with 32 participants would still be .80. Assuming that the effect size was 1, the power at the .05 level with 32 participants would be .99, comparable to the power corresponding to the Cohen's *d* of 1.27 and the 20 subjects in Tcheang et al. (2011).

Apparatus and virtual environment. The experiment was conducted in a 4 × 4 m physical room. The virtual environments were presented by an immersive virtual reality system that used Vizard software (WorldViz, Santa Barbara, CA), an nVisor SX60 head-mounted display (HMD, NVIS, Inc. Virginia), and an InterSense IS-900 motion tracking system (InterSense, Inc., Massachusetts). Screen resolution within the display was 1,280 × 1,024 pixels. Participants' viewing orientations and positions were tracked by the IS-900 motion tracking system. To record the participants' responses (i.e., their pointing directions), an InterSense IS-900 Wand (InterSense, Inc., Massachusetts) was used.

---

[3] Cohen's *d* of the gain effect was calculated based on $\sqrt{\frac{2F}{N}}$. In Tcheang et al. (2011), F value for the gain effect was 16.12 and N was 20.

A circular virtual environment with a grass-textured ground was used. The environment had a radius of 50 m and was bordered with a 10-m-high gray wall. The wall's center was set as fixed at the center of the physical room. Twelve large numbers (1 to 12) in an equal distance (30º) to each other were presented on the wall, which had the structure of a clock face (see Figure 3A). The direction of the number 12 was fixed and aligned with the axis of the physical room. We refer to this wall as clockwall.

Design and procedure. The primary independent variable is the gain that participants experienced in the adaptation phase. Each participant sequentially adapted to normal-gain and under-gain. The normal-gain was always introduced before the under-gain as the normal-gain was supposed to have no influence on the following under-gain. The adaptation phase of the normal-gain condition was also included to make these two gain conditions more comparable (i.e., there were the adaptation phase and the post-adaptation triangle completion task for both conditions). The adaptation phase for the normal gain condition seemed not necessary as the normal gain in the virtual environment should mimic the normal gain in our real life. However, adding the normal-gain adaptation would make the data interpretation more simply because the experiences of the virtual environment other than experience of the gain per se in the adaptation phase could also affect the triangle completion task after the adaptation.

At the beginning of each adaptation phase, participants saw the clockwall environment while they were standing at the center of the environment facing the number 12. The number 12 was always used as the starting number for all adaptation trials. On each trial, the participants were presented with turning directions (left or right) and one target number (1 to 11, randomly chosen), and then were asked to rotate to face the target number. Participants were asked to turn their body slowly while maintaining their position and keeping their head orientation aligned with their body orientation. For the normal-gain adaptation, which represents our normal perceptions, the visual turning angle of the participants was equal to their physical turning angle. That is, when they physically turned 90º, for example, turning right from 12 to 3, they could see the view of 90º from 12 (if 12 is at the direction of 0º). For the under-gain condition, the visual turning angle of the participants was equal to 0.7 times their physical turning angle. That is, when they physically turned 90º, they could see the view of 63º from 12.

Each adaptation phase included 16 trials and the trials were organized into four blocks of four trials each. The first block presented all twelve numbers all the time (i.e., numbers shown block 1). The second block only presented the number 12 all the time (i.e., numbers hidden block 1). The other numbers were shown only after a response was made by the participant. The third block was the same as the first block and referred to as numbers shown block 2 whereas the last block was the same as the second block and referred to as numbers hidden block 2.

For the numbers shown blocks, after turning to face the target number (always visible), the participants rotated back to the number 12 and a new trial began. For the numbers hidden blocks, after participants turned to face the target number (in the estimated direction), all 12 numbers were shown to provide feedback for the participants. If the participants' estimated direction of the target number differed from the correct direction of the target number, they were asked to turn to the correct direction of that number before turning back to 12 for the next trial. The physical turning angle from 12 to the target number with the instructed turning direction was recorded as a response to check whether participants indeed adapted to the introduced gain. The numbers hidden blocks were used for two purposes: first, it would encourage participants to learn the gain because they had to turn to face the target number based on the learned gain in the numbers hidden blocks although they could turn to face the target number using the visible number in the number shown blocks. Second, we could measure the learned gain in the numbers hidden blocks to see whether participants really learned the gain (see results below).

The primary dependent variable is the observed gain in the testing phases that immediately followed each adaptation phase. The testing phases after the two adaptation phases were identical. In the testing phase, the testing environment was almost in darkness. The participants could only see the poles that were presented to guide them to walk. On each trial, a blue pole, a red pole, and a green pole were presented in sequence to indicate the origin, the turning point, and the end of the path (see Figure 3B). To make sure that participants represented their turning angle only from idiothetic information, we prevent participants from turning their heads to look for the next pole and learn the rotation angles from visual positions of the poles[4]. In particular, participants were required to only turn their body slowly in the direction indicated on the screen of HMD till they saw the next pole. When participants saw the pole, they stopped to walk towards it. Once participants reached the pole, it disappeared and the next pole appeared. At the end of the path, the participants were asked to point with their arms and fingers to the origin of the path. When their finger and forearm were not perfectly aligned, the participants were asked to verbally inform the experimenter which one could indicate their pointing. The experimenter aligned the wand with their pointing and clicked a button on the wand to record the data. The pointing direction of the wand relative to the current facing direction was recorded as the response pointing direction.

---

[4] If participants were allowed to turn their head to preview the direction of the pole before turning their body to face the pole, one may argue that participants visually saw the direction of the pole before they turned their body. Hence they learned the turning angle both from visual direction of the pole and idiothetic inputs during turning.

Each testing phase included six trials (Turning direction: left, right × Turning angle: 63º, 90º, 117º). The sequence of presenting the six trials was randomly determined for each participant. Similar to a traditional triangle completion task (e.g., Loomis et al., 1993), on each trial, participants walked a two-legged path in darkness and at the end of the path, pointed to the origin of the path. Each leg was 1.8 m long. The origin of each path was randomly chosen from a circle with a radius of 1.8 m whose center overlapped the center of the physical room. The turning point of each path was the center of the physical room. The turning angles (63º, 90º, or 117º) were chosen according to Tcheang et al. (2011).

[Figure 3]

Participants did the experiment individually. Each participant was led by an experimenter into the physical room with eyes blindfolded. After putting on the headset, the participants first saw a dark environment and received one training trial for the testing task (see the following sections). In particular, participants walked a path with a left turn of 110º, then pointed to the origin. This training trial was designed to ensure the participants understood the task in the testing phase correctly. After making a response, the participants were provided feedback for whether their response was approximately correct. Each participant then sequentially received the two conditions, each consisting of the adaptation and the testing phases. During the intervals between the phases within each condition and between the conditions, which lasted for approximately one minute each, participants were asked to close their eyes and take a rest while the experimenter switched the programs.

Results

Before looking at the observed data, we assumed that the participants could successfully adapt to the normal-gain and under-gain conditions in the adaptation phase. Therefore, if any response of one participant showed a violation to this assumption, that participant should be excluded. For the adaptation phase, the response gain was calculated for each trial as the physical response turning angle / the physical correct turning angle. For each adaptation condition, if the average response gain in the last block (without numbers) of one participant exceeded the range of mean ± 2.5SD, then this participant failed to adapt to the gain condition. According to this criterion, two participants were excluded. Thus, 31 participants entered data analysis.

In each adaptation phase, for the trials with all numbers shown, if absolute error exceeded 20º (as evidence that participants did not follow the instructions), then this trial was excluded. For the trials with numbers 1 to 11 hidden, all trials were kept except for the trials in which the participant did not follow the instructions (e.g., turning in the wrong direction). Figure 4 shows the average response gain (response turning angle/correct turning angles) in four blocks in both conditions. To examine whether the adaptation to the gain conditions was successful, we conducted a paired *t* test and found that there was a significant difference between the normal-gain and under-gain conditions in the last block of the adaptation phase, $t(30) = 13.65$, $p < .001$, Cohen’s $d = 3.47$. This result indicates that the participants successfully adapted to the gain conditions as predicted. In particular, the participants adapted to a new vision-locomotion coupling relationship in the under-gain condition.

[Figure 4]

In each testing phase, for each trial turning to the left, the response pointing direction was first transferred to a range from -180º to 180º in which 0º was their current facing direction (H in Figures 2 and 3), and then transferred to the absolute value (i.e., flipped). After flipping, we calculated the circular average of the response pointing direction for the two trials with the same turning angle and a different turning direction (e.g., turning to left 63º and right 63º).

Consequently, we obtained the Response Pointing direction (*RP*) for each combination of adaptation condition and turning angle for each participant.

Note that so far, the *RP* values ranged from 0º to 360º. Actually, a single *RP* direction could be represented by an infinite number of values. For example, 340º and -20º represent the same *RP* direction. As a consequence, it would lead to an infinite number of values of fractional pointing (*FP*, see Table 1), which indicates the adaption effect. To remove this ambiguity, for each combination of adaptation condition and turning angle, we normalized the *RP* for each participant so that the angular difference between the *RP* for each participant and the circular mean of the *RP* across participants was from -180º to 180º. This normalization assures that the normalized *RP* has the smallest deviation from the circular mean among all possible values of the same *RP* direction. For example, if the original *RP* for one participant is 340º and the circular mean is 90º, the *RP* for this participant will be changed to -20º. As a result, the absolute angular difference between the normalized *RP* for each participant and the circular mean was smaller than 180º. In this example, the absolute angular difference using the original *RP* is 250º whereas the absolute angular difference using the normalized *RP* is 110º.

We calculated the fractional pointing (*FP*) as the normalized *RP* / correct pointing direction (*C*) (see Figure 5). A 2 × 3 repeated ANOVA was conducted with the fractional pointing (*FP*) as the dependent variable and the adaptation condition (normal-gain, under-gain) and turning angle (63º, 90º, 117º) as the independent variables. The Greenhouse-Geisser correction for violation of sphericity was applied where applicable. None of the main effect or the interaction effect was significant [main effect of adaptation condition: $F(1, 30) = 0.61$, $MSE = 0.03$, $p = .442$, $\eta_p^2 = .02$; main effect of turning angle: $F(1.48, 44.28) = 2.77$, $MSE = 0.03$, $p = .088$, $\eta_p^2 = .09$; interaction: $F(2, 60) = 0.60$, $MSE = 0.01$, $p = .553$, $\eta_p^2 = .02$]. This result indicates that changes in the vision-locomotion coupling in the adaptation phase did not influence the participants' estimation of the turning angle in the subsequent walking trials.

[Figure 5]

Discussion

The results in Experiment 1 suggest that although participants became adapted to a rotation gain of 0.7 (i.e., the under-gain condition) in the adaptation phase (Figure 4), they did not show significant changes in their estimation of turning angle of the walking paths (Figure 5). These null adaptation effects were inconsistent with the findings in Tcheang et al. (2011). According to Table 1, we could see that even when participants completely adopted the new gain (0.7), in forming the spatial representation of the path during walking in darkness, the theoretical *FP* was only about overall 10% (6%, 10%, and 14% for turning angles of 63º, 90º and 117º, respectively) larger than the normal gain (i.e., *FP* = 1). Furthermore, Tcheang et al. (2011; see their Figure 2C) showed that the observed fractional turn (approximately overall 1.15 collapsed across all turning angles and specifically 1.08, 1.14, 1.22 for turning angles of 63º, 90º and 117º respectively) from) was actually much smaller than the theoretical values (overall 1.58 and specifically 1.52, 1.57, 1.63 for turning angles of 63º, 90º and 117º respectively) in Table 1. Thus, the observed *FP* in the current experiment could be much smaller than the theoretical values, which was about 10% only larger than that based on the normal gain (i.e., *FP* = 1). Therefore, the null adaptation effect might have been due to insensitivity of the observed *FP* to the change of the gain.

Mathematically, to produce a greater effect of adaptation on the observed *FP*, we can use a shorter second leg in the two-leg outbound path used in Experiment 1. Figure 6 plots the hypothetical *FP* as a function of the length of the second leg in a two-leg outbound path (assuming gain = 0.7). In Experiment 1, the second leg was 1.8 m long, the same as the first leg. As shown in Table 1, $C = 180 – A / 2$, $RP = 180 – 0.7A / 2$, and thus the hypothetical $FP = RP / C = 1 + (0.3 \times A) / (360 – A)$. The most extreme case of using a shorter second leg is to remove the second leg. In this case, $C = 180 – A$, $RP = 180 – 0.7A$, then the hypothetical $FP = RP / C = 1 + (0.3 \times A) / (180 – A)$, which is larger than the hypothetical *FP* when the legs are equally long. As illustrated in Figure 6 (left panel) and also in Table 2, when the second leg is 0 m, the hypothetical *FP* was 16%, 30% and 56% larger than the hypothetical *FP* with the normal gain (i.e., 1) for turning angles of 63°, 90° and 117°, respectively. To maximize the likelihood of replicating the adaptation effect in Tcheang et al. (2011), in Experiment 2 we had one group of participants who only walked the first leg, turned their bodies, and then pointed to the origin.

[Figure 6]

[Table 2]

With the expectation that the adaptation effect in Tcheang et al. (2011) would be observed by using only one leg, in Experiment 2 we differentiated the bidirectional model and the unidirectional model by adding a transportation group. This group of participants received the same adaptation phase as the walking group. However, in testing, they did not physically move but instead watched a video showing a path from a first-person view. If the visual representation contributed to spatial representation through the vision-locomotion coupling, then the participants' pointing responses in normal-gain and under-gain conditions would be different. More specifically, according to the bidirectional model, participants would overestimate their turning angle in the under-gain condition via the coupling (see Figure 2C). Therefore the *FP* would be smaller in the under-gain condition than in the normal-gain condition. Otherwise, according to the unidirectional model, there would be no difference in *FP* between the two conditions for the transportation group.

Figure 6 also plots the change of *FP* for the transportation group as a function of the length of the second leg according to the bidirectional model. When the second leg is 0 m, $C = 180 - A$, $RP = 180 - A / 0.7$, and the hypothetical $FP = RP / C$. As illustrated in Figure 6 (right panel) and in Table 2, when the second leg was 0 m, the hypothetical *FP* was 23%, 43% and 80% smaller than the hypothetical *FP* with the normal gain (i.e., 1) for the turning angles of 63°, 90° and 117°, respectively.

Experiment 2

Experiment 2 had two purposes. First, we intended to replicate the findings in Tcheang et al. (2011) when participants only walked one leg and made a rotation (walking group). Second, we intended to examine whether the vision-locomotion coupling is bidirectional or unidirectional. Therefore, a transportation group was added. This group received the same adaptation phases as the walking group, but in testing, they were presented with only visual information (i.e., participants physically did not move but watched a video showing a path). Both models (see Figure 1) predict that the *FP* would be larger in the under-gain condition than in the normal-gain condition for the walking group (see Figure 6). However, these two models presented a different prediction about the *FP* for the transportation group. According to the bidirectional model, the *FP* would be smaller in the under-gain condition than in the normal-gain condition, whereas according to the unidirectional model, there would be no difference in *FP* between these two conditions.

Method

Participants. A total of 52 university students participated in Experiment 2 and were assigned into two groups (walking or transportation). After exclusion (see results section), 50 participants were included in the data analysis. All participants received course credit in an introductory psychology course for their participation. Each group included 25 participants with 12 males and 13 females.

Compared with Experiment 1, the effect size in Experiment 2 should be much larger due to removal of the second leg. Therefore, we planned to use about 24 participants for each group. The power at the .05 level with 24 participants and Cohen's *d* of 1.2 would be .99.

Apparatus and environment. They were the same as that used in Experiment 1.

Design and procedure. They were the same as those in Experiment 1 except the following two changes. First, the paths were changed (see Figure 3C). On each trial, the participants walked forward 1.8 m from the origin (indicated by a blue pole) to the turning point (indicated by a red pole) and turned to a certain direction (left or right) with a certain angle (63^o, 90^o, or 117^o) to face a green pole. However, they did not walk further towards the green pole. After the rotation, they were asked to point to the origin of the path. Critically, as in the previous experiment, participants were required to turn in the indicated direction (left or right) slowly till they saw the green pole and not to look for the green pole before they turned their body so that they estimated their rotation angle only based on the idiothetic information during turning. Second, the transportation group was added. The transportation group had the same adaptation phases as the walking group but had the different task in the testing phases.

The transportation group performed a video task in the testing phases. On each trial, the participants were asked to stand in place, watch a first-person-view video that showed a walking path from the walking group, and then point to the origin of the path with their arms and fingers.

They were not allowed to rotate their bodies or walk physically while watching the video. During the inter-trial interval, the participants were led by the experimenter to walk in darkness for approximately three seconds. In each of the videos, the red pole, the green pole, the circular wall without any number, and the grassland were shown. The videos were simulated from the recorded locations and orientations of participants in the walking group while they walked through the six paths. Each participant in the transportation group corresponded to one participant in the walking group (with the same gender) by using the recorded locations and orientations of the latter. Therefore the sequence of presenting the six paths was the same as that for the corresponding participant in the walking group. The transportation group also had one practice trial before the normal-gain adaptation phase, performing a video trial corresponding to the practice trial of the walking group.

Results and discussion

The same exclusion criteria were applied to both groups. Two participants were excluded from the transportation group because their average response gain in the last block of the adaptation phase exceeded the range of mean ± 2.5SD (i.e., they failed to adapt to the normal-gain or under-gain conditions). Finally, 25 participants of each group entered data analysis.

For the adaptation phase, Figure 7 shows the average response gain in four blocks in both conditions for both groups. To examine whether adaptation to the gain conditions was successful, a mixed ANOVA was conducted with the adaptation condition (normal-gain, under-gain) and testing group (walking, transportation) as independent variables for the last block of the adaptation phase. As in Experiment 1, we found a significant main effect of the adaptation condition [$F(1, 48) = 246.74$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = .84$]. Neither the main effect of the testing group nor the interaction was significant [main effect of testing group: $F(1, 48) = 1.04$,

$MSE = 0.01$, $p = .313$, $\eta_p^2 = .02$; interaction: $F(1, 48) = 0.46$, $MSE = 0.01$, $p = .501$, $\eta_p^2 < .01$]. The results indicate that, as predicted, both groups adapted to the new vision-locomotion relationship in the under-gain condition equally well.

[Figure 7]

For each testing phase, again, we calculated *FP* for each turning angle for both groups (see Figure 8).

As described above, the expected adaptation effects in the two groups should be different regardless of whether the coupling is bidirectional or unidirectional. In particular, the bidirectional model predicts opposite adaptation effects (i.e., *FP* would be larger in the under-gain condition than in the normal-gain condition for the walking group but smaller in the under-gain condition than in the normal-gain condition for the transportation group). The unidirectional model predicts that *FP* would be larger in the under-gain condition than in the normal-gain condition for the walking group but not change across the gain conditions for the transportation group. Thus, testing the adaption effect by combining the data of the two groups was not theoretically meaningful as it cannot differentiate these two models.

Consequently, two repeated-measure ANOVA were conducted separately for two groups with *FP* as the dependent variable and the adaptation condition (normal-gain, under-gain) and turning angle (63º, 90º, 117º) as the independent variables.

For the walking group, there was a significant main effect of the adaptation condition, $F(1, 24) = 14.43$, $MSE = 0.03$, $p < .001$, $\eta_p^2 = .38$. Neither the main effect of turning angle nor the interaction was significant [main effect of turning angle: $F(1.29, 31) = 0.26$, $MSE = 0.08$, $p = .674$, $\eta_p^2 = .01$; interaction: $F(2, 48) = 2.68$, $MSE = 0.02$, $p = .079$, $\eta_p^2 = .10$]. These results suggest that for the walking group, the participants underestimated the turning angle of the walking paths in the under-gain condition, which led to an increase in the observed *FP*. Thus, we replicated the findings in Tcheang et al. (2011) successfully.

By contrast, for the transportation group, the main effect of the adaptation condition was not significant [$F(1, 24) = 0.31$, $MSE = 0.23$, $p = .583$, $\eta_p^2 = .01$]. There was a significant main effect of turning angle, $F(1.34, 32.25) = 16.04$, $MSE = 0.24$, $p < .001$, $\eta_p^2 = .40$. The interaction between adaptation condition and turning angle was not significant, $F(2, 48) = 0.99$, $MSE = 0.10$, $p = .380$, $\eta_p^2 = .04$. The participants demonstrated no decrease in the calculated *FP*, which suggests that they did not overestimate their turning angle (i.e., they were not influenced by the adaptation gain). As the bidirectional model predicts that the *FP* in the normal-gain condition should be larger than that in the under-gain condition, the result of the null adaptation effect did not fit this hypothesis.

[Figure 8]

Since the *FP* values in the normal-gain condition for each turning angle were different for the transportation group, it is possible that using a simple average of the original *FP* values across three turning angles might not be sensitive enough to show the adaptation effect. Therefore, we calculated the difference ratios between normal-gain and under-gain conditions for each turning angle and then averaged the ratios for each participant (Table 3). According to our theoretical predictions (see Table 1 and Figure 6), for the walking group, the under-gain *FP* would be larger than the normal-gain *FP*. Therefore, the difference ratio was calculated as (under-gain *FP* - normal-gain *FP*) / normal-gain *FP*. For the transportation group, the normal-gain *FP* would be larger than the under-gain *FP*. Therefore, the difference ratio was calculated as (normal-gain *FP* - under-gain *FP*) / normal-gain *FP*.

[Table 3]

For both groups, if there was no adaptation effect, the difference ratio would be 0. One-sample *t* tests were conducted to compare the difference ratios collapsed across turning angles with the theoretical value of null adaptation effect (i.e., the difference ratio is 0) for the two groups separately. For the walking group, the difference ratio (i.e., mean = 0.14 or 14%) was significantly larger than 0, $t(24) = 4.15$, $p < .001$, Cohen's $d = 1.17$. Again, this indicates the adaptation effect. For the transportation group, the difference ratio (i.e., mean = -0.02) was not different from 0, $t(24) = 0.23$, $p = .820$, Cohen's $d = 0.07$, indicating no adaptation effect. The Cohen's *d* values were calculated as $t \times \sqrt{\frac{2}{N}}$, where N is the sample size (see Glass, McGaw, & Smith, 1981).

The null adaptation effect in the transportation group could still be due to lack of the power to detect the existing effect in the transportation conditions. There could be two reasons for the lack of the power in the transportation group. First, in general, the current experimental manipulation could only produce a small gain effect in the testing phase. Second, the task difficulty level in the transportation group might be higher than that in the walking group. Thus, the pointing responses were noisier in the transportation group than in the walking group. Therefore, a small gain effect was not detected due to the larger variance of the data.

Even in the walking condition, the observed gain effect was much smaller than the theoretical gain effect. As shown in Table 3, the observed gain difference ratio was 5%, 12%, and 25% for turning angles of 63º, 90º and 117º, respectively. As shown in Table 2, the theoretical gain effect (in terms of gain different ratio) was 16%, 30%, and 56%. The smaller observed gain effect than the theoretical gain effect was also reported by Tcheang et al. (2011). In their study, the observed gain effect was approximately 8%, 14%, and 22% (from their Figure 2C) and the theoretical gain effect should be 52%, 57%, and 63% (see FT in Table 1) for turning angles of 63º, 90º and 117º, respectively. Although the observed gain effect in the walking group was comparable to Tcheang et al. (2011), it might still not be large enough to detect especially when the participants' pointing response was noisier in the transportation condition due to the higher difficulty level, which was confirmed by the results. We compared the angular pointing errors in these two groups for the normal-gain conditions. We did not use the angular errors for the under-gain condition as they were also affected by gains in addition to the difficulty level. An ANOVA using absolute errors as the dependent variable and using group (walking vs. transportation) and turning angles as the independent variables showed that the only significant effect was the main effect of group, $F(1, 48) = 4.92$, $MSE = 1020.18$, $p < .031$, $\eta_p^2 = .09$. The mean absolute error collapsed across turning angles was 16º for the walking group (16º, 12º, and 19º for the turning angles of 63º, 90º, and 117º respectively) and 27º for the transportation group (30º, 23º, and 29º for the turning angles of 63º, 90º, and 117º respectively)[5]. This indicated that the pointing responses in the transportation group could be noisier due to the higher level difficulty.

Considering these two concerns, we note that it might not be valid to directly use the non-significant result to support the null hypothesis of no adaptation effect in the transportation group. To provide stronger evidence that the observed difference ratio is due to no adaptation effect, we also calculated Bayes Factors measuring the ratio of the likelihood of an adaptation effect to the likelihood of no adaptation effect in the transportation group (see Gallistel, 2009). The results indicate that the null effect is strongly supported (see Appendix for details).

---

[5] The mean absolute error collapsed across turning angles for Experiment 1 was 17º (17º, 18º, and 16º for the turning angles of 63º, 90º, and 117º respectively).

Experiment 2 indicates that when only idiothetic input was provided (the walking group), the participants' estimation was influenced by the gain, which suggests that idiothetic input can contribute to spatial representation via the vision-locomotion coupling. By contrast, when only visual input was provided (the transportation group), the participants' estimation was not influenced by the gain, which suggests that visual input cannot contribute to spatial representation via the coupling. Therefore, these results fit with the predictions of the unidirectional model.

General Discussion

The current study investigated whether the vision-locomotion coupling in multimodal spatial representations is bidirectional or unidirectional. After adapting to a new relationship between vision and locomotion, the participants who were presented with only idiothetic input about the outbound path showed that their spatial representation of the path was influenced by the new relationship. By contrast, the participants who were presented with only visual input about the outbound path did not show a significant influence from the new relationship. These findings were consistent with the predictions according to the unidirectional model, but not those according to the bidirectional model (see Figure 1).

Our findings are in accordance with the coupling view that vision and locomotion are coupled to calibrate actions (Rieser & Pick, 2007). When there was a mismatch between visual and idiothetic inputs through rotation (adaptation phase), the participants recalibrated their actions in the later walking task, as is consistent with the previous findings in translational locomotion (i.e., locomotion in straight lines) (Rieser et al., 1995; Ziemer et al., 2013) and triangle completion tasks (Tcheang et al., 2011).

However, we found that the coupling is asymmetrical or unidirectional, which extends the multimodal model theory of spatial representations (Tcheang et al., 2011). According to this theory, either vision or locomotion might contribute to a multimodal spatial representation. Therefore, it is not necessary to stipulate that visual input only directly contributes to spatial representations, whereas idiothetic input can also indirectly contribute to spatial representations through the vision-locomotion coupling. Thus, the multimodal model seems more consistent with the bidirectional model of the vision-locomotion coupling. It is difficult to use the bidirectional model to explain why the new coupling was involved when there was only idiothetic input, but not when there was only visual input, about the traversed path. Therefore, the multimodal model theory might need to be expanded to embody the unidirectional model.

According to the unidirectional model, participants in the walking group created idiothetic representations using idiothetic input and then created visual representations through the coupling between vision and locomotion (i.e., specifying the idiothetic information in terms of the visual representation). Then visual representations contributed to spatial representations. Therefore, the new relationship between vision and locomotion influenced spatial representations, producing the adaptation effect. By contrast, participants in the transportation group created visual representations using visual input. As the coupling was only from idiothetic representations to visual representations in forming spatial representations, the visual representations directly formed spatial representations and did not activate the coupling to form spatial representations. Therefore, there was no adaptation effect.

We note that the concept of *spatial representation* in the unidirectional model (Figure 1) is similar to the concept of *spatial image* in the amodal theory (Avraam, Hatzipanayioti, & Avraamides, 2018; Avraamides & Kelly, 2008; Giudice, Betty, & Loomis, 2011; Loomis, Klatzky, & Giudice, 2013; Shelton & McNamara, 2001; Yamamoto & Shelton, 2005, 2009). The amodal theory stipulates that different perceptual inputs (and language) construct functionally equivalent *spatial images*, representations that are amodal, and do not retain modality-specific features once created. Extending the amodal theory, the unidirectional model incorporates the unidirectional coupling relationship between vision and locomotion prior to forming the *amodal spatial images*. This coupling relationship specifies the routes of transferring the modality specific representations (i.e., visual representation and idiothetic representation) to the *amodal spatial representation*.

The current findings also have important implications about the relationship between cue interaction and vision-locomotion coupling. In particular, vision is fundamental in coupling although vision is not dominant over locomotion in cue interaction. There is a lot of evidence showing cue combination of the vision and locomotion systems in navigation (e.g., Ellard & Shaughnessy, 2003). More relevant, the first experiment of Tcheang et al. (2011) showed that participants combined their rotation angles indicated by vision and locomotion systems when both visual and idiothetic information about the path were presented. These cue combination findings suggest that vision is not dominant over locomotion in forming the spatial representations of the turning angles when both cues are available. Moreover, previous studies showed that path integration based on visual optical flow alone could be much more difficult than path integration based on idiothetic information alone (e.g., Kearns, Warren, Duchon, & Tarr, 2002; Klatzky, Loomis, Beall, Chance, & Golledge, 1998). Experiment 2 of the current study also showed worse pointing performance in the transportation group than in the walking group (i.e., absolute error was 16° for the walking group and 27° for the transportation group in the normal-gain condition). Thus, vision is indeed inferior to locomotion in indicating the direction of the path origin.

However, the current findings suggest that vision is fundamental in the unidirectional coupling between vision and locomotion systems. We speculate that while cue combination might primality rely on relative cue reliability (e.g., Chen et al., 2017), vision-locomotion coupling might rely more on direct or indirect perception of the environmental change as a consequence of locomotion (Rieser & Pick, 2007). Because the vision system can but the locomotion system cannot directly perceive the environmental change in locomotion, the vision system is fundamental in coupling and idiothetic information is specified by visual information during locomotion prior to forming spatial representations.

Although vision calibrates locomotion in the current study, we note that vision may not be the only possible perceptual input for recalibrating actions. Vision is just one source (and yet the most reliable one in many cases) of perceptual inputs that can perceive the environmental change as a consequence of locomotion. Under some circumstances, nonvisual external inputs may also contribute to locomotor recalibration. Usually, for most people, vision provides the most efficient and reliable feedback from the environment (Thinus-Blanc & Gaunet, 1997). Therefore the corresponding locomotor recalibration occurs naturally when visual feedback is available. However, people who are congenitally blind might use haptic or auditory information to perceive the consequences of their actions, and therefore develop the coupling between perception-action. Rieser, Guth, and Hill (1986) demonstrated that participants who were congenitally blind could develop a better spatial configural knowledge than those whose broad field of vision was damaged at an early age. Interestingly, recent studies showed that rodents' head directional cells were stabilized once infant rodents opened their eyes or when they could touch the boundary of the enclosure (Bassett, Wills, & Cacucci, 2018; Tan et al., 2015). These findings suggest that both vision and touch could specify the idiothetic signals during rotation of the body to face different directions. Future research can apply the current paradigm to other modality inputs in people with vision damage to examine whether there are similar coupling relationships between other perceptual modalities and locomotion.

One question of interest would be how humans develop multimodal spatial representation. Since many perceptual and motor abilities do not reach adult-level maturity at the very beginning of life, it is possible that young infants may not be able to form a multimodal spatial representation, although some evidence suggests that they can use information from a single modality in many cases (for a review, see Nardini & Cowie, 2012). Through interaction with the environment, infants gradually learn to coordinate their perception and actions to guide navigation. For example, from 7 to 12 months old, human infants show sensitivity to both visual and vestibular information, and the reliance levels on both modalities are closely related to their development of locomotion ability (Bremner, Hatton, Foster, & Mason, 2011). The crucial element in the development of coordinating perception and action may be the experience of interacting with the environment. Some animal studies also suggest that getting feedback from the environment initiates the functional tuning of spatial-related neurons. For example, although head direction cells are found firing several days before eye opening in the rat pups, only after eye opening do the signals become stable because visual inputs exert control over the head direction system (Tan et al., 2015).

However, multimodal representation seems to develop slowly. For example, even 10- to 12-year-old children coordinate perception and action at a coarse-grained level, not like adults at a fine-grained level, at least on some tasks (Chihak, Grechkin, Kearney, Cremer, & Plumert, 2014). Clearly, to further answer the question of how humans develop multimodal representation, future research is needed to link the existing behavioral and neural evidence.

In the current unidirectional model (Figure 1), idiothetic input could contribute to spatial representations both directly or indirectly via the vision-locomotion coupling. However, the current experiments did not investigate whether there is the direct contribution of idiothetic representations to spatial representations (the blue arrow from idiothetic representation to spatial representation). Previous studies showed that without vision, people can also develop spatial representations (e.g., Klatzky et al., 1998). One may argue that in the absence of vision people could still develop spatial representations via the vision-locomotion coupling, so the direct contribution may not be necessary. This argument may predict worse performance in spatial tasks when participants do not have any visual information of the experimental space assuming the arbitrarily created experimental space would less accurately constrain the possible spatial extents of walked paths. However, the empirical evidence to examine this argument is not clear. In some studies, performance in spatial tasks did not depend on whether participants had prior visual exposure to the environments (e.g., Philbeck et al., 2001; Yamamoto, Meléndez, & Menzies, 2014) whereas in other studies, performance was better when participants had a prior visual preview of the environment (Arthur, Philbeck, & Chichka, 2007; Philbeck & O’Leary, 2005; Warren, 1970). We chose to include the direct contribution in the current model as there is no clear evidence to exclude the direct contribution.

We acknowledge that in the transportation group, we did not (or could not) completely remove the influence of idiothetic information. Participants stood still while watching the video of the traversed path. Although the idiothetic information of standing still could not indicate the traversed path, it may bias the perception of the traversed path presented visually due to the incongruent idiothetic information (i.e., physical turning angle is 0º). In particular, participants might underestimate the turning angles indicated by the visual information, thus, the *FP* would increase (see Figure 2A and 2B). If we assume that a larger incongruence causes a larger bias, then the *FP* will increase with the turning angle. Indeed, the results showed that *FP* increased with the turning angle in the transportation group (see Figure 8). However, the possible bias from the idiothetic information of standing still should be independent of the manipulation of the gain (normal- vs. under-gain) because the gain manipulation should not change either the idiothetic information of standing still or visual turning angles. Therefore, this possible bias should not change the possible effect of the gain in the transportation group (normal- vs. under-gain) and then affect the dissociation between the unidirectional and bidirectional models by examining the gain effect.

We also speculate that while the new relationship between vision and locomotion was developed in the adaptation phase, the one established in daily life (i.e., gain = 1) was still stored in memories and retrieved in the testing phase. This speculation can explain that the observed adaption effects in the walking group of Experiment 2 were smaller than the predicted *FP* values. The mean of the observed difference ratios across different turning angles was 0.14 whereas the mean of the predicted difference ratios was 0.34. In the testing phase, both the gains developed in the adaptation phase (temporary) and in daily life (long-term) might have contributed to visual representations via the vision-locomotion coupling. This speculation can also explain why the observed adaption effects in the second experiment of Tcheang et al. (2011) were smaller than the predicted values (0.15 vs. 0.58). Although the contribution of the long-term gain might have reduced the adaption effect determined by the temporary gain, there is no clear reason to believe that the reduction was larger in the transportation group than in the walking group in Experiment

2 of the current study. Therefore the significant adaptation effect in the walking group but the null adaptation effect in the transportation group should not be attributed to the possible influence of the long-term gain.

Although the results of Experiment 1 could not differentiate the unidirectional model from the bidirectional model, they still had important theoretical and methodological implications. The results in Experiment 1 showed no adaptation effect when participants pointed to the origin instead of walking to the origin. This finding suggests that the adaptation effect reported in Tcheang et al. (2011) might have been attributed to the gain effect on the idiothetic representation of the turn that participants made to face the origin in the testing phase. Thus, the real adaptation effect on the spatial representation of the path might be much smaller than that indicated by the previous study. This conclusion is consistent with our analyses listed in Table 1. According to Table 1, the hypothetical *FP*, which was attributed only to the gain effect on the spatial representation of the path, was only about 10% larger than the normal gain (i.e., 1) whereas the hypothetical *FT*, which was attributed to both the gain effect on the idiothetic representation in response (i.e., turning to the origin before walking towards it) and the gain effect on the spatial representation of the path, was 58% larger than the normal gain. Therefore, any future study that will examine the adaptation effect on the spatial representation of a path should find a solution to increase the difference of the hypothetical *FP* from 1, like the solution used in Experiment 2 of the current study (i.e., decreasing the length of the second leg).

In conclusion, the current study not only replicated previous findings supporting the theory of multimodal spatial representations but also differentiated a unidirectional model from a bidirectional model of multimodal spatial representations. Consistent with the unidirectional model, the current findings demonstrate that while idiothetic information contributes indirectly to spatial representations via the vision-locomotion coupling, visual information only contributes directly to spatial representations.

Appendix

In Experiment 2, to provide stronger evidence that the observed difference ratio in the transportation group is due to no adaptation effect, we also calculated Bayes Factors measuring the ratio of the likelihood of an adaptation effect to the likelihood of no adaptation effect for the transportation group (see Gallistel, 2009).

In particular, we used the difference ratio observed in the walking group (i.e., 0.14 or 14%) as the real adaptation effect in the transportation group under the alternative hypothesis claiming an adaptation effect, assuming that the transportation group would show the same amount of adaptation effect as the walking group. The likelihood of any possible observed difference ratio under the null hypothesis (i.e., no adaption effect) can be measured by the probability density of the t value of the observed difference ratio ($t = \frac{-0.02\times\sqrt{N}}{SD}$, where $N = 25$, $SD = 0.36$ from the transportation group) in a t distribution ($df = 24$). The likelihood of any possible observed difference ratio under the alternative hypothesis can be measured by the probability density of the t value of the observed difference ratio in a noncentral t distribution (noncentral parameter λ is the t value of the theoretical adaptation effect, i.e., $\lambda = \frac{0.14\times\sqrt{N}}{SD}$, where $N = 25$, $SD = 0.36$ from the transportation group, $df = 24$). The probability density as a function of the observed difference ratio under the competing hypotheses is plotted in Figure A1. Results showed that the Bayes Factor (i.e., likelihood ratio) in favor of the null hypothesis was 10.61, providing strong support for null adaptation effect. The null effect is favored if Bayes Factor is larger than 3, and strongly favored if the Bayes Factor is larger than 10, whereas an adaptation effect is favored if the Bayes Factor is smaller than 1/3, and strongly favored if the Bayes Factor is smaller than 1/10 (Rouder, Speckman, Sun, Morey, & Iverson, 2009). If the Bayes Factor is between 1/3 and 3, neither is favored.

[Figure A1]

Similarly, we calculated Bayes Factors in favor of no adaptation effect for each turning angle (BF was 0.35, 4.32, and 12.95 for turning angles of 63º, 90º and 117º, respectively). The null effect was favored for the turning angle of 90º and strongly favored for the turning angle of 117º although the Bayes Factor could not distinguish between the null effect and an adaptation effect for the turning angle of 60º.

Note that when calculating the Bayes Factors, we used the observed gain effect (i.e., gain difference ratio) in the walking group as the real gain effect and used the observed variances of the gain effect in the transportation group as the variance. Therefore, Bayes Factors that favored the null effect in the transportation group were calculated as have already considered the influences from the two facts that the observed gain effect was much smaller than the theoretical one and that the pointing response in the transportation group was noisier.

Compliance with Ethical Standards

Funding: This research was funded by the Natural Sciences and Engineering Research Council of Canada to Weimin Mou.

Ethical approval: All procedures performed in the reported study involving human participants were in accordance with the ethical standards of the University of Alberta Research Ethics Boards and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

Conflict of Interest: Author Yu Du declares that she has no conflict of interest. Author Weimin Mou declares that he has no conflict of interest. Author Lei Zhang declares that she has no conflict of interest.

Acknowledgements

We thank Benson Ng, Janina Valencia, and Nim Binning for their contribution in data collection.

References

Arthur, J. C., Philbeck, J. W., & Chichka, D. (2007). Spatial memory enhances the precision of angular self-motion updating. *Experimental Brain Research*, *183*, 557–568. http://doi.org/10.1007/s00221-007-1075-0

Avraam, S., Hatzipanayioti, A., & Avraamides, M. N. (2018). Orientation-dependent spatial memories for scenes viewed on mobile devices. *Psychological Research*. http://doi.org/10.1007/s00426-018-1069-5

Avraamides, M. N., & Kelly, J. W. (2008). Multiple systems of spatial memory and action. *Cognitive Processing*, *9*, 93–106. http://doi.org/10.1007/s10339-007-0188-5

Avraamides, M. N., Sarrou, M., & Kelly, J. W. (2014). Cross-sensory reference frame transfer in spatial memory: The case of proprioceptive learning. *Memory & Cognition*, *42*, 496–507. http://doi.org/10.3758/s13421-013-0373-y

Bassett, J. P., Wills, T. J., & Cacucci, F. (2018). Self-organized attractor dynamics in the developing head direction circuit. *Current Biology*, *28*, 609–615. http://doi.org/10.1016/j.cub.2018.01.010

Bremner, J. G., Hatton, F., Foster, K. A., & Mason, U. (2011). The contribution of visual and vestibular information to spatial orientation by 6- to 14-month-old infants and adults. *Developmental Science*, *14*, 1033–1045. http://doi.org/10.1111/j.1467-7687.2011.01051.x

Bruggeman, H., Zosh, W., & Warren, W. H. (2007). Optic flow drives human visuo-locomotor adaptation. *Current Biology*, *17*, 2035–2040. http://doi.org/10.1016/j.cub.2007.10.059

Campos, J. L., Butler, J. S., & Bülthoff, H. H. (2014). Contributions of visual and proprioceptive information to travelled distance estimation during changing sensory congruencies. *Experimental Brain Research*, *232*(10), 3277–3289. http://doi.org/10.1007/s00221-014-

4011-0

Chen, X., McNamara, T. P., Kelly, J. W., & Wolbers, T. (2017). Cue combination in human spatial navigation. *Cognitive Psychology*, *95*, 105–144. http://doi.org/10.1016/j.cogpsych.2017.04.003

Cheng, K., Shettleworth, S. J., Huttenlocher, J., & Rieser, J. J. (2007). Bayesian integration of spatial information. *Psychological Bulletin*, *133*, 625–637. http://doi.org/10.1037/0033-2909.133.4.625

Chihak, B. J., Grechkin, T. Y., Kearney, J. K., Cremer, J. F., & Plumert, J. M. (2014). How children and adults learn to intercept moving gaps. *Journal of Experimental Child Psychology*, *122*, 134–152. http://doi.org/10.1016/j.jecp.2013.12.006

Ellard, C. G., & Shaughnessy, S. C. (2003). A comparison of visual and nonvisual sensory inputs to walked distance in a blind-walking task. *Perception*, *32*, 567–578. http://doi.org/10.1068/p5041

Gallistel, C. R. (2009). The importance of proving the null. *Psychological Review*, *116*(2), 439–453. http://doi.org/10.1037/a0015251

Giudice, N. A., Betty, M. R., & Loomis, J. M. (2011). Functional equivalence of spatial images from touch and vision: evidence from spatial updating in blind and sighted individuals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 621–634. http://doi.org/10.1037/a0022331

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-Analysis in Social Research*. Sage Publications.

Harris, C. S. (1963). Adaptation to displaced vision: visual, motor, or proprioceptive change? *Science*, *140*, 812–813. http://doi.org/10.1126/science.140.3568.812

Kearns, M. J., Warren, W. H., Duchon, A. P., & Tarr, M. J. (2002). Path integration from optic flow and body senses in a homing task. *Perception*, *31*, 349–374. http://doi.org/10.1068/p3311

Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., & Golledge, R. G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science*, *9*(4), 293–298. http://doi.org/10.1111/1467-9280.00058

Loomis, J. M., Klatzky, R. L., & Giudice, N. A. (2013). Representing 3D space in working memory: Spatial images from vision, hearing, touch, and language. In S. Lacey & R. Lawson (Eds.), *Multisensory Imagery* (1st ed., pp. 131–155). Springer-Verlag New York. http://doi.org/10.1007/978-1-4614-5879-1

Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., & Fry, P. A. (1993). Nonvisual navigation by blind and sighted: Assessment of path integration ability. *Journal of Experimental Psychology: General*, *122*(1), 73–91. http://doi.org/10.1037/0096-3445.122.1.73

Loomis, J. M., Klatzky, R. L., Golledge, R. G., & Philbeck, J. W. (1999). Human navigation by path integration. In R. G. Golledge (Ed.), *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes* (pp. 125–151). Baltimore: John Hopkins University Press.

Mou, W., McNamara, T. P., & Zhang, L. (2013). Global frames of reference organize configural knowledge of paths. *Cognition*, *129*, 180–193. http://doi.org/10.1016/j.cognition.2013.06.015

Nardini, M., & Cowie, D. (2012). The development of multisensory balance, locomotion, orientation and navigation. In A. J. Bremner, D. J. Lewkowicz, & C. Spencer (Eds.), *Multisensory Development*. Oxford University Press. http://doi.org/10.1093/acprof

Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. *Current Biology*, *18*, 689–693. http://doi.org/10.1016/j.cub.2008.04.021

Philbeck, J. W., Klatzky, R. L., Berhrman, M., Loomis, J. M., & Goodridge, J. (2001). Active control of locomotion facilities nonvisual navigation. *Journal of Experimental Psychology: Human Learning and Memory*, *27*(1), 141–153.

Philbeck, J. W., & O'Leary, S. (2005). Remembered landmarks enhance the precision of path integration. *Psicologica*, *26*, 7–24.

Pick, H. L., Rieser, J. J., Wagner, D., & Garing, A. E. (1999). The recalibration of rotational locomotion. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(5), 1179–1188.

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1157–1165. http://doi.org/10.1037/0278-7393.15.6.1157

Rieser, J. J. (1999). Dynamic spatial orientation and the coupling of representation and action. In R. G. Golledge (Ed.), *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes* (pp. 168–190). Baltimore: JHU Press.

Rieser, J. J., Guth, D. A., & Hill, E. W. (1986). Sensitivity to perspective structure while walking without vision. *Perception*, *15*, 173–188.

Rieser, J. J., & Pick, H. L. (2007). Using locomotion to update spatial orientation: What changes with learning and development? In J. M. Plumert & J. P. Spencer (Eds.), *The Emerging Spatial Mind* (pp. 77–103). New York: Oxford University Press.

Rieser, J. J., Pick, H. L., Ashmead, D. H., & Garing, A. E. (1995). Calibration of human locomotion and models of perceptual-motor organization. *Journal of Experimental*

Psychology: Human Perception and Performance, *21*, 480–497.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. http://doi.org/10.3758/PBR.16.2.225

Shelton, A. L., & McNamara, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology*, *43*, 274–310. http://doi.org/10.1006/cogp.2001.0758

Tan, H. M., Bassett, J. P., O'Keefe, J., Cacucci, F., & Wills, T. J. (2015). The development of the head direction system before eye opening in the rat. *Current Biology*, *25*, 479–483. http://doi.org/10.1016/j.cub.2014.12.030

Tcheang, L., Bülthoff, H. H., & Burgess, N. (2011). Visual influence on path integration in darkness indicates a multimodal representation of large-scale space. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(3), 1152–1157. http://doi.org/10.1073/pnas.1011843108

Thinus-Blanc, C., & Gaunet, F. (1997). Representation of space in blind persons: Vision as a spatial sense? *Psychological Bulletin*, *121*(1), 20–42. http://doi.org/10.1037//0033-2909.121.1.20

Warren, D. H. (1970). Intermodality interactions in spatial localization. *Cognitive Psychology*, *1*(2), 114–133. http://doi.org/10.1016/0010-0285(70)90008-3

Yamamoto, N., Meléndez, J. A., & Menzies, D. T. (2014). Homing by path integration when a locomotion trajectory crosses itself. *Perception*, *43*, 1049–1060. http://doi.org/10.1068/p7624

Yamamoto, N., & Shelton, A. L. (2005). Visual and proprioceptive representations in spatial memory. *Memory & Cognition*, *33*(1), 140–150. http://doi.org/10.3758/BF03195304

Yamamoto, N., & Shelton, A. L. (2009). Orientation dependence of spatial memory acquired from auditory experience. *Psychonomic Bulletin & Review*, *16*, 301–305. http://doi.org/10.3758/PBR.16.2.301

Zhao, M., & Warren, W. H. (2015). How you get there from here: Interaction of visual landmarks and path integration in human navigation. *Psychological Science*, *26*, 1–10.

Ziemer, C. J., Branson, M. J., Chihak, B. J., Kearney, J. K., Cremer, J. F., & Plumert, J. M. (2013). Manipulating perception versus action in recalibration tasks. *Attention, Perception, & Psychophysics*, *75*, 1260–1274. http://doi.org/10.3758/s13414-013-0473-6

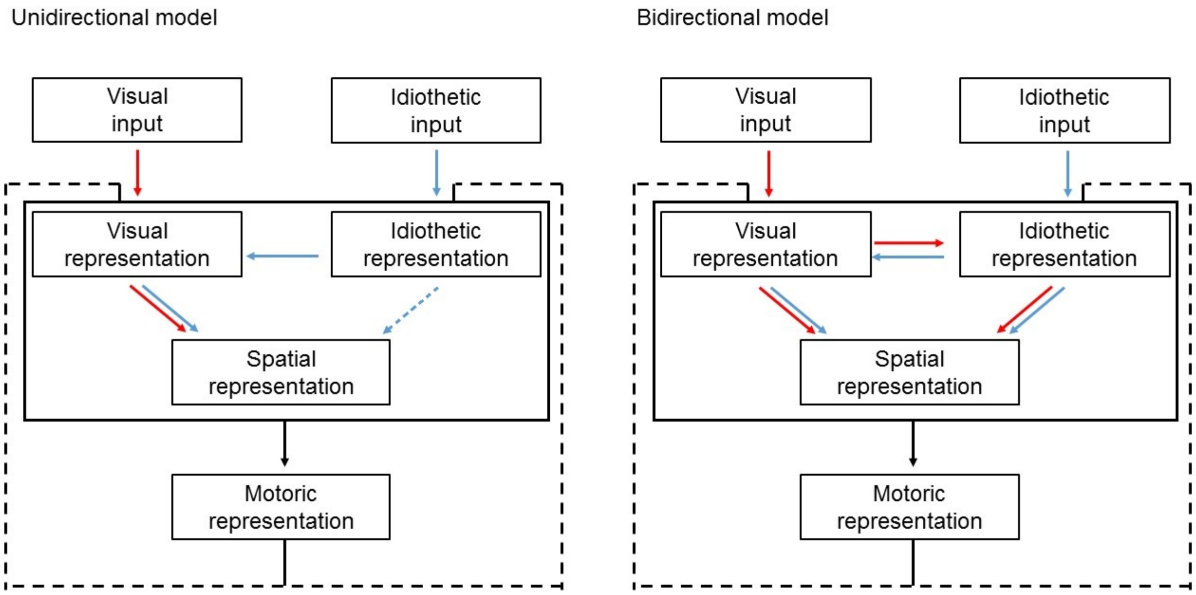

Figure 1. Two possible models of spatial information processing during navigation, extending the MMR Model in Tcheang et al. (2011, Figure 3B). An arrow indicates the information processing flow. The rectangle with bold solid lines indicates the multi-modal representation. Dash lines indicate the visual and idiothetic consequences of action. In the Unidirectional model, the dash arrow from the idiothetic representation to the spatial representation is used to indicate that this model does not have a clear claim on whether idiothetic representations also directly contribute to spatial representations.

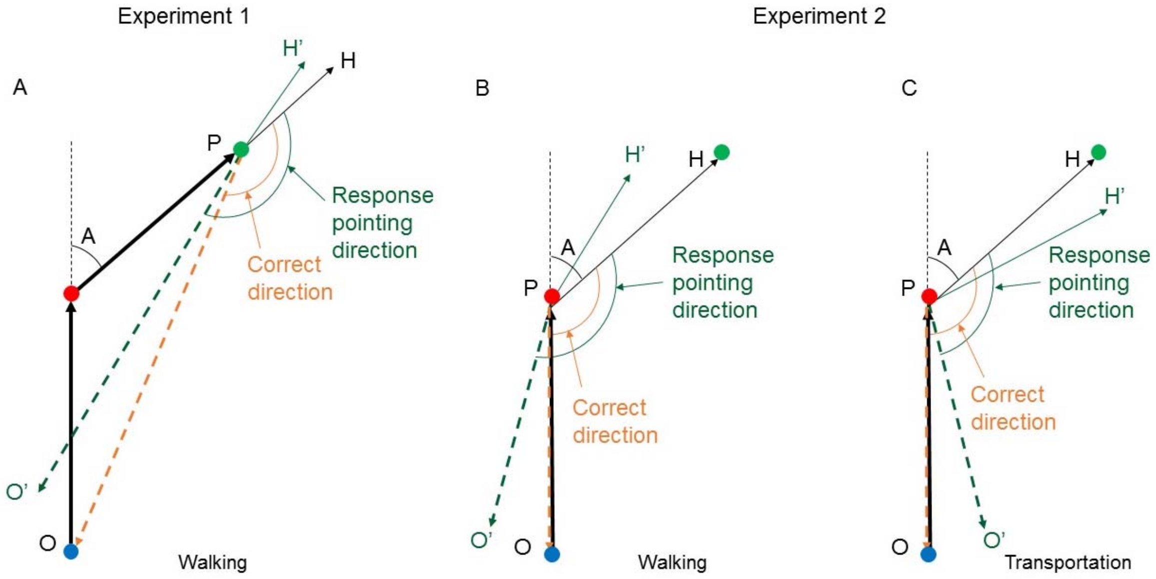

Figure 2. Predictions of response pointing with a rotation gain of 0.7 in a modified triangle completion task. The bold black arrows indicate the walking route. The blue dot denotes the origin (O) of the path. The red dot denotes the turning point of the path. The letter *A* denotes an actual turning angle in a path. *P* and *H* denote the participant's actual position and actual heading when making the response pointing. *H'* denotes the participant's estimated heading. In Panel A, the green dot denotes the end of the path. In Panel B and C, the green dot denotes the green pole that determines the final facing direction of the participants.

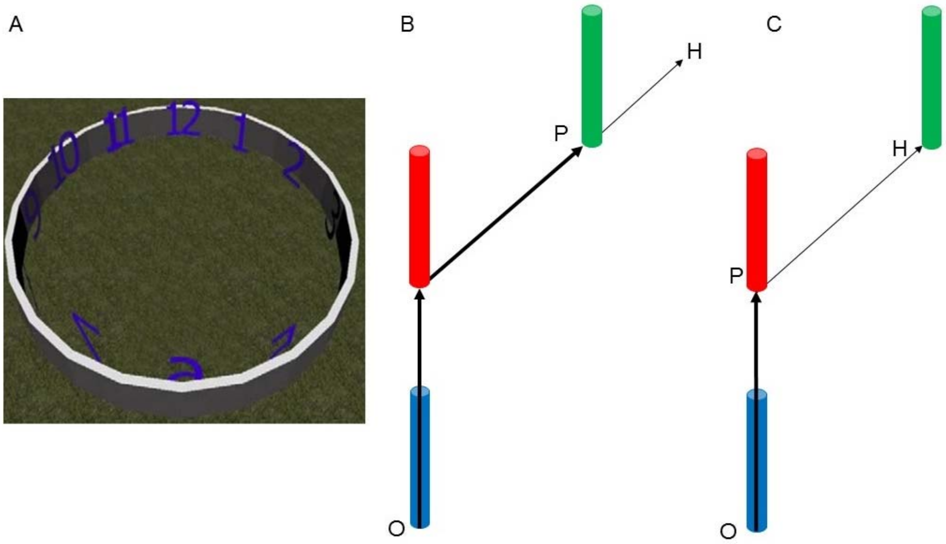

Figure 3. Experimental setup. Participants never saw the views during the experiment. Panel A represents the top view of the environment in the adaptation phase. Panel B and C are examples of the traversed path with the poles in Experiment 1 and 2, respectively. The blue pole indicates the origin (O) of a walking path (randomly determined). The bold black arrows indicate the walking routes. The red pole indicates the turning point of the path. *P* and *H* denote the participant’s actual position and actual heading.

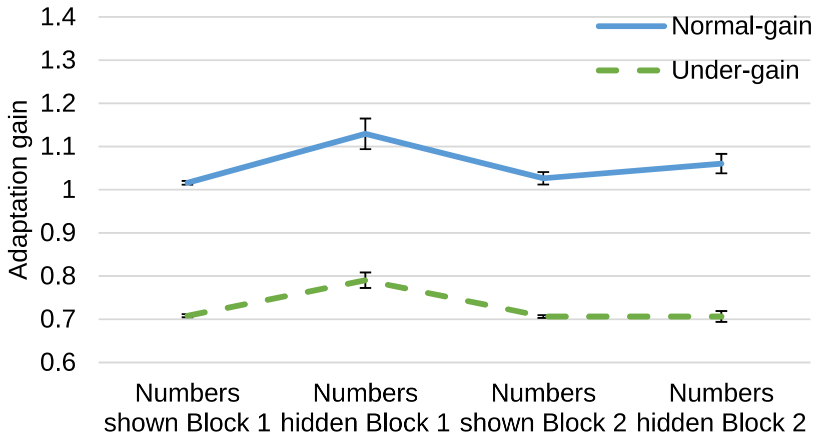

Figure 4. Average response gain in four blocks in adaptation phase in Experiment 1. Error bars represent SEM.

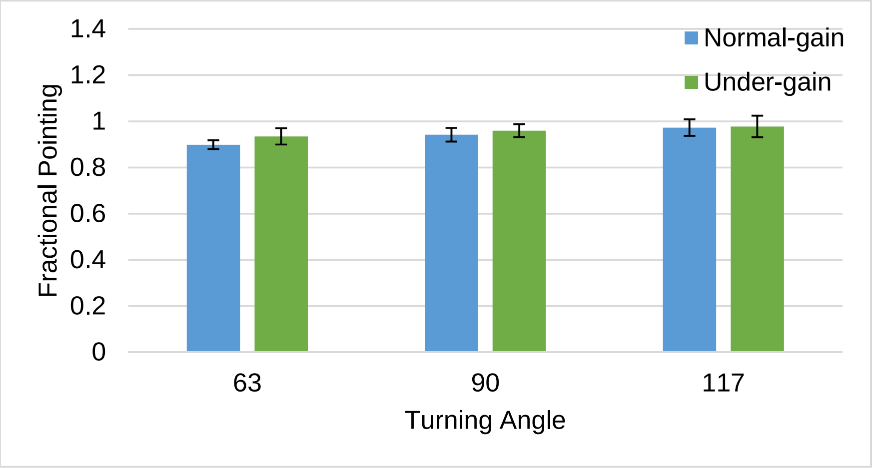

Figure 5. Average fractional pointing (*FP*) in testing phase in Experiment 1. Error bars represent SEM.

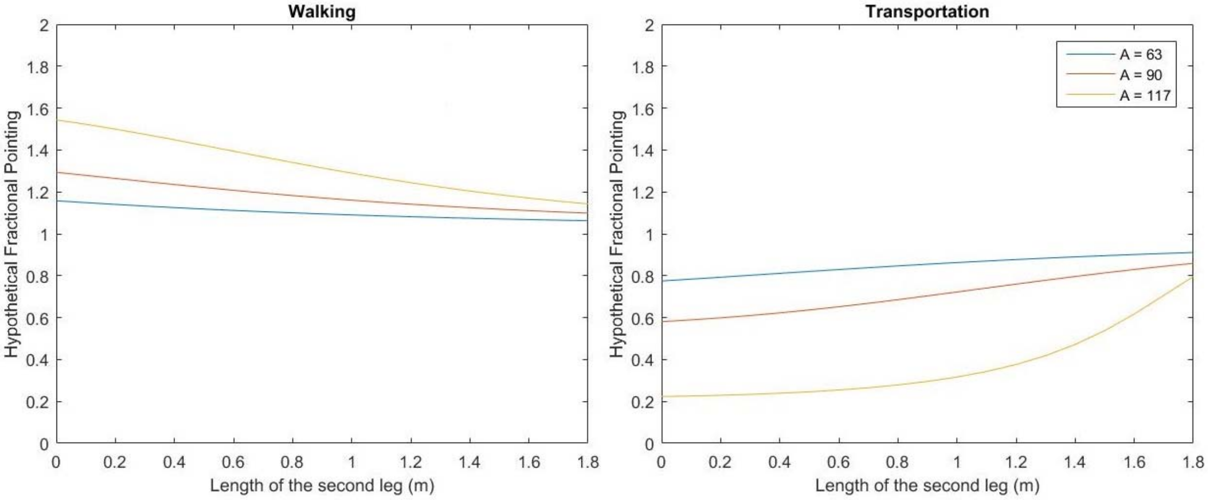

Figure 6. The hypothetical fractional pointing (*FP*) as a function of the length of the second leg for walking and transportation groups in the under-gain condition. *FP* = Response Pointing direction (*RP*) / Correct pointing direction (*C*). The first leg is assumed to be constant as 1.8 m and the observed gain in adaptation phase is assumed to be 0.7. The letter *A* is actual turning angle in the path. Note that $C = \mathrm{acos}(\frac{\mathrm{Lr}+\cos(\mathrm{A})}{\sqrt{1+Lr^2-2\times Lr\times Cos(180-A)}})$, Lr is the length ratio of the second leg to the first leg. To calculate RP, A is replaced by 0.7A for the walking condition and replaced by A/0.7 for the transportation condition.

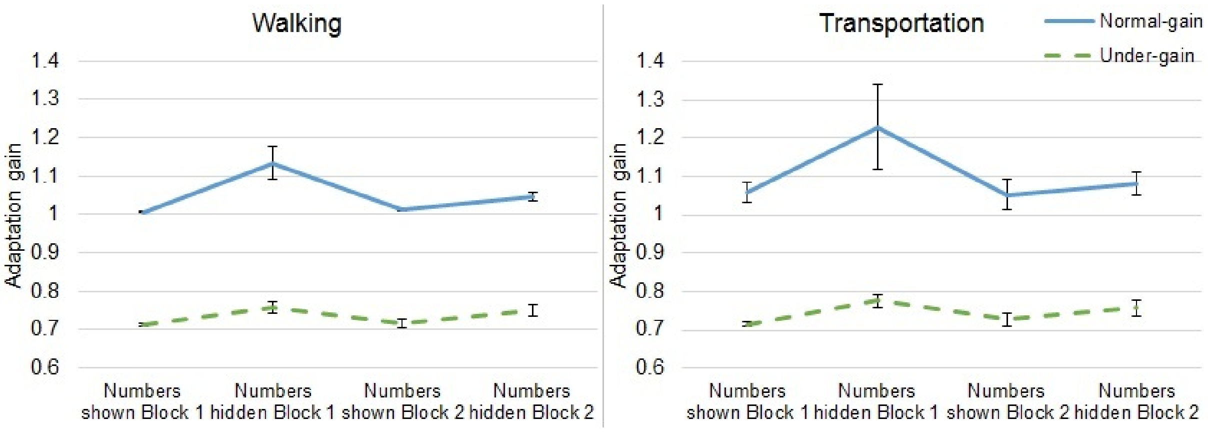

Figure 7. Average response gain in four blocks in adaptation phase in Experiment 2.

Error bars represent SEM.

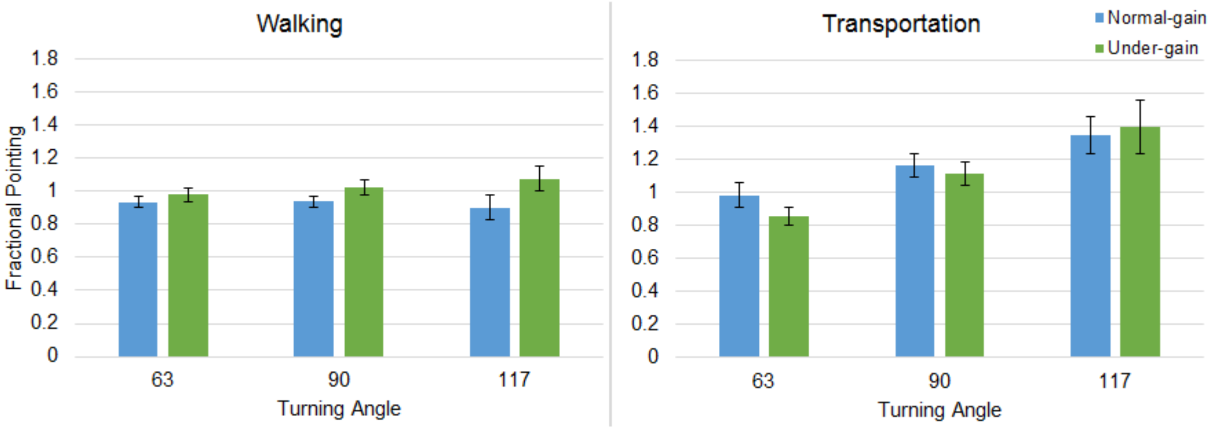

Figure 8. Average fractional pointing (*FP*) in testing phase in Experiment 2. Error bars represent SEM.

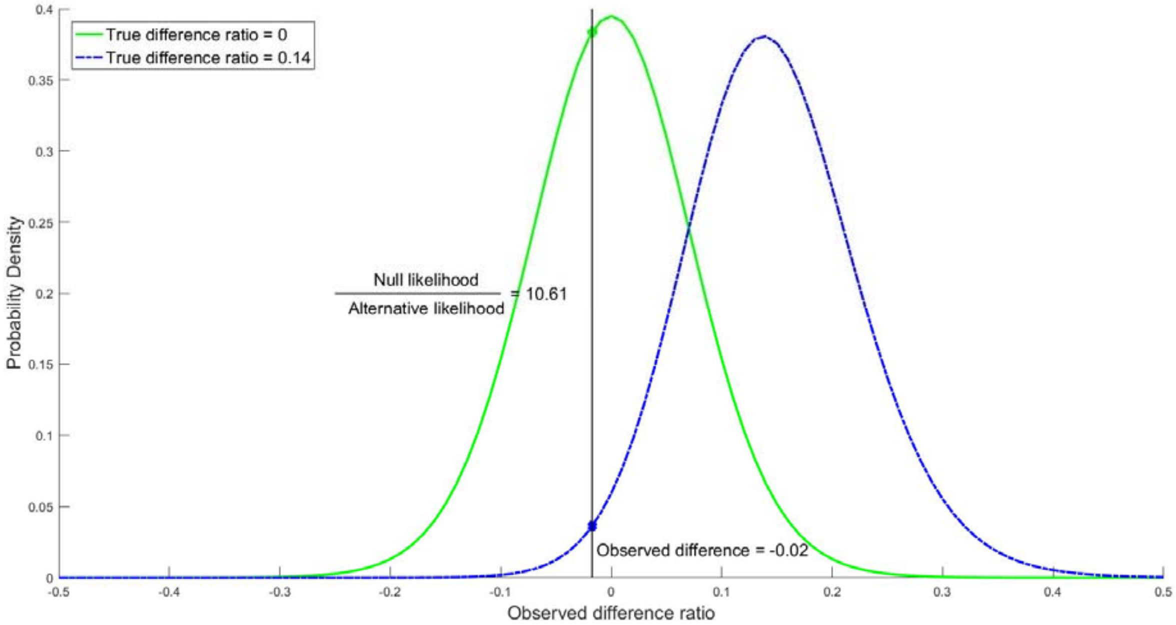

Figure A1. Probability density as a function of observed difference ratio under the null hypothesis that claims no adaptation effect (true difference ratio = 0, green solid curve) and under the alternative hypothesis that claims an adaptation effect (true difference ratio = 0.14, blue dashed curve).

Table 1. *Correct direction (C), response pointing direction (RP), response turning direction (RT), corresponding fractional pointing (FP) and fractional turn (FT) given a hypothetical turning angle (A) and the three turning angles used in Experiment 1 when the path has two equally-long legs and the rotation gain is 0.7.*

Actual turning angle (A)	**Correct direction (C)**	**Response Pointing direction (RP)**	**Response Turning direction (RT)**	**Fractional Pointing (FP)**	**Fractional Turn (FT)**
A	180 – *A* / 2	180 – 0.7*A* / 2	RP / 0.7	RP/C	RT/C
63	149	158	226	1.06	1.52
90	135	149	212	1.1	1.57
117	122	139	199	1.14	1.63

Note: The letter *A* denotes an actual turning angle in a walking path with two equally-long legs. The Response Turning direction (*RT*) = *RP* / 0.7 because presumably the participants were still carrying the adaptation effect during the response turn at the end of the path (Tcheang et al., 2011).

Table 2. *Correct direction (C), response pointing direction (RP), response turning direction (RT), corresponding fractional pointing (FP) and fractional turn (FT) given a hypothetical turning angle (A) and the three turning angles used in Experiment 2 when the path has only one leg and the rotation gain is 0.7.*

Actual turning angle (A)	**Correct direction (C)**	**Walking**		**Transportation**	
		Response Pointing direction (RP)	**Fractional Pointing (FP)**	**Response Pointing direction (RP)**	**Fractional Pointing (FP)**
A	180 – *A*	180 – 0.7*A*	RP/C	180 – *A* / 0.7	RP/C
63	117	136	1.16	90	0.77
90	90	117	1.30	51	0.57
117	63	98	1.56	13	0.20

Note: The letter *A* denotes an actual turning angle after the first leg. There was no second leg.

Table 3. *Mean and standard deviation (SD) of the difference ratios for each turning angle in Experiment 2.*

	Walking				**Transportation**			
	63º	**90º**	**117º**	**Collapsed**	**63º**	**90º**	**117º**	**Collapsed**
Mean	0.05	0.12	0.25	0.14	0.07	0.02	-0.14	-0.02
SD	0.18	0.30	0.33	0.17	0.24	0.29	0.81	0.36

Note: The columns “Collapsed” indicate the mean and standard deviation of the difference ratios across all three turning angles. For the walking group, the difference ratio was calculated as (under-gain *FP* - normal-gain *FP*) / normal-gain *FP*. For the transportation group, the difference ratio was calculated as (normal-gain *FP* - under-gain *FP*) / normal-gain *FP*.