CrossMark

Reference frames in spatial updating when body-based cues are absent

Qiliang He[1] · Timothy P. McNamara[1] · Jonathan W. Kelly[2]

Abstract The current study investigated the reference frame used in spatial updating when idiothetic cues to self-motion were minimized (desktop virtual reality). In Experiment 1, participants learned a layout of eight objects from a single perspective (learning heading) in a virtual environment. After learning, they were placed in the same virtual environment and used a keyboard to navigate to two of the learned objects (visible) before pointing to a third object (invisible). We manipulated participants' starting orientation (initial heading) and final orientation (final heading) before pointing, to examine the reference frame used in this task. We found that participants used the initial heading and the learning heading to establish reference directions. In Experiment 2, the procedure was almost the same as in Experiment 1 except that participants pointed to objects relative to an imagined heading that differed from their final heading in the virtual environment. In this case, pointing performance was only affected by alignment with the learning heading. We concluded that the initial heading played an important role in spatial updating without idiothetic cues, but the representation established at this heading was transient and affected by the interruption of spatial updating; the learning heading, on the other hand, corresponded to an enduring representation which was used consistently.

Keywords Spatial cognition · Reference frame · Spatial updating · Idiothetic cues · Virtual reality

✉ Timothy P. McNamara
t.mcnamara@vanderbilt.edu

[1] Department of Psychology, Vanderbilt University, PMB 407817, 2301 Vanderbilt Place, Nashville, TN 37240-7817, USA

[2] Department of Psychology, Iowa State University, Ames, IA, USA

To navigate effectively, animals must update the spatial relations between their body and objects in the environment as they move. This process is referred to as spatial updating (Amorim, Glasauer, Corpinot, & Berthoz, 1997; Amorim & Stucchi, 1997; Farrell & Robertson, 1998; Rieser, 1989). Spatial updating can be achieved via external self-motion cues (vision, audition) or internal self-motion cues (vestibular, kinesthetic, efferent information; see Waller & Hodgson, 2013, for a review). Internal self-motion cues are often referred to as idiothetic cues. Regardless of the source of information used in spatial updating, people need a reference frame to represent the spatial relation between themselves and the target. A number of studies have investigated the nature of this reference frame when the full set of idiothetic cues were available during spatial updating (Hodgson & Waller, 2006; Kelly, Avraamides, & Loomis, 2007; Mou, McNamara, Valiquette, & Rump, 2004; Wang et al., 2006). In daily life, people often commute by car or train and therefore have a limited set of idiothetic cues during navigation, but few studies have examined the reference frame when idiothetic cues were limited or not available in spatial updating. In the current study, we investigated the reference system during spatial updating when idiothetic cues were not available in a virtual environment navigated by a keyboard. By bridging this gap in the literature, we can better understand the mechanisms of spatial updating and investigate how people adapt their spatial representations depending on the availability of the idiothetic cues.

Spatial updating can be an egocentric or an allocentric process. Egocentric spatial updating refers to the process whereby the navigator updates each object's location with respect to the body using a reference system centered on the body (and typically defined by the reference directions of front, back, right, or left; Wang, 2016). In contrast, allocentric spatial updating refers to the process whereby the navigator updates his or her position in the environment using a reference system

external to the body and anchored in the environment (e.g., using canonical directions of north, south, east, or west; Klatzky, 1998).

There is evidence of the use of both egocentric and allocentric reference systems in spatial updating when idiothetic cues are available. In relatively featureless environments, people may rely primarily on egocentric reference systems (Wang et al., 2006), but in more natural, feature-rich environments, both reference systems seem to be employed (e.g., Amorim et al., 1997; Hodgson & Waller, 2006; Holmes & Sholl, 2005; Kelly et al., 2007; Mou et al., 2004; Waller & Hodgson, 2006; Xiao, Mou, & McNamara, 2009). For example, Kelly et al. (2007) had participants learn a layout of objects from a fixed perspective and later had them recall the learned objects by pointing from several imagined perspectives. When recall occurred in the same room as did learning, recall was facilitated when the imagined perspective was aligned with (parallel to) the participant's facing direction during recall, considered evidence for an egocentric reference frame. Whether recall occurred in the learning room or an adjacent room, recall was facilitated when the imagined perspective was aligned with the learning view; this effect is considered evidence for an allocentric reference frame in long-term memory (e.g., Shelton & McNamara, 2001). Moreover, Mou et al. (2004) found that if the imagined perspective was aligned with the allocentric as well as the egocentric reference frame, performance was better than when the imagined perspective was aligned with only one of these reference frames. Taken together, these results indicate that egocentric and allocentric reference frames may be used simultaneously during spatial updating.

In the absence of idiothetic cues, past research shows that spatial updating is still possible with visual cues (Riecke, Heyde, & Bülthoff, 2005; Riecke, Veen, & Bülthoff, 2002; Ruddle, Volkova, & Bülthoff, 2011; Waller, Loomis, & Haun, 2004; but see Klatzky, Loomis, Beall, Chance, & Golledge, 1998). He, McNamara, and Kelly (2017) investigated the nature of the reference frame in a path integration task when idiothetic cues were limited. Participants navigated to three waypoints in a desktop virtual environment using the computer keyboard and then pointed to the first waypoint using a joystick (there was no learning or familiarization phase and only one waypoint was visible at any point in time). The results indicated that under these circumstances, participants used an allocentric reference frame in which the principal reference direction was defined by their initial perspective in the environment (the initial heading).

Spatial updating often occurs, however, in familiar environments (e.g., walking to one's bathroom at night in the dark). Besides the initial heading, another heading that people may use to establish the reference direction during spatial updating is the perspective from which they learn a layout of objects (the learning heading). Studies have shown that people organize their spatial memories in terms of a small number (1–2) of reference directions even when environments are experienced from multiple points of view (Kelly & McNamara, 2008; McNamara, 2003; Mou & McNamara, 2002; Shelton & McNamara, 2001; Valiquette, McNamara, & Smith, 2003; Waller, Montello, Richardson, & Hegarty, 2002), and generally prefer to use the learning heading to establish a reference direction to represent the object-to-object spatial relations and self-to-object spatial relations (Kelly et al., 2007; Mou et al., 2004).

Combining the aforementioned findings, we conjectured that both the initial heading and the learning heading could be used to establish reference directions in spatial updating in a familiar environment without idiothetic cues. In the current study, we manipulated the alignment among the initial heading, the learning heading, and the imagined heading (the heading participants needed to imagine they were facing in the virtual environment before responding) to examine the reference system.

Figure 1 outlines the experimental design in Experiment 1. Participants learned a layout of objects from a heading of 0° (the learning heading) in a virtual environment. After learning, they were placed in the same virtual environment and used a keyboard to navigate sequentially to two of the learned object locations. The starting orientation (initial heading) and location varied across experimental conditions. After navigating to the second object, participants occupied a position and an orientation (the final heading) which also varied across experimental conditions. Participants used a joystick to point to a third object from this final position and heading, and hence the final heading is referred to as the imagined heading (in Experiment 2, the final heading and the imagined heading differed). As a result of these manipulations, the initial heading could be aligned or 90° misaligned with the imagined heading, and the learning heading could be aligned or 90° misaligned with the imagined heading.

For brevity, the condition in which both the initial and learning headings were aligned with the imagined heading was named IL; the condition in which only the learning heading was aligned with the imagined heading was named L; the condition in which only the initial heading was aligned with the imagined heading was named I; and the condition in which both the initial and the learning headings were misaligned with the imagined heading was named M.

We assumed that spatial relations that are encoded in memory can be retrieved whereas those that are not encoded must be computed or inferred, which introduces errors and time to the decision processes (Klatzky, 1998; Shelton & McNamara, 2001). Better performance for an imagined heading relative to others indicates that the spatial relations are represented in memory with respect to a reference direction parallel to that imagined heading (Mou et al., 2004; Rump & McNamara,
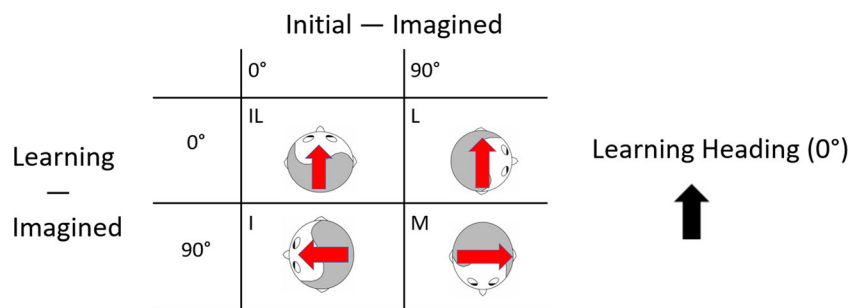
Fig. 1 Design of Experiment 1. The learning heading was 0° in all conditions. The initial heading corresponded to the participant's orientation at the beginning of the path. The imagined heading in Experiment 1 corresponded to the participant's orientation at the end of the path (= final heading). Differences in headings are absolute values. The design factorially manipulated (**a**) the difference between the initial heading and the imagined heading and (**b**) the difference between the learning heading and the imagined heading. The *letters* in each cell identify the experimental conditions (see main text for details). The *human figure* represents the orientation of the initial heading, and the *red arrow* represents the orientation of the imagined (final) heading. (Color figure online)

2013). Therefore, comparison of performance across experimental conditions can determine whether participants used the initial heading, the learning heading, or a combination of the two to establish the reference direction(s): For example, if performance in the I condition or in the L condition was better than performance in the M condition, it would suggest that participants used the initial heading or the learning heading, respectively, to establish reference directions; if performance in the IL condition was also better than performance in the I condition and the L condition, it would suggest that misalignment with one of two reference directions conferred a cost to pointing accuracy (e.g., Mou et al., 2004) or that the availability of two aligned reference directions enabled participants to construct a more accurate representation at the time of responding (perhaps in working memory).[1] On the other hand, if performance was equivalent across conditions, it would suggest that participants did not use the initial heading or the learning heading to establish a reference direction.

We considered that the reference directions established by the initial heading and the learning heading were components of allocentric reference systems because neither of these headings changed as participants changed their orientation in the virtual environment (an alternative interpretation is discussed in the General Discussion). However, we conjectured that the reference systems defined by the initial and learning headings differed in terms of stability (Allen & Haun, 2004): The reference system defined by the initial heading was assumed to be transient because this heading was not constant, and participants never viewed the layout of objects from this heading; the reference system defined by the learning heading, on the other hand, was assumed to be enduring because its orientation was constant and participants learned the layout from this perspective intensively. In Experiment 2, we tested this hypothesis by following almost the same procedure as in Experiment 1, but asking

participants to imagine a heading that was 90° misaligned with their final heading before responding (e.g., if their final heading was 0°, participants needed to imagine they were facing 90° or -90°). The four conditions of Experiment 1 were the same, but by introducing this mental rotation we changed the alignment of the imagined heading and the final heading. We hypothesized that this mental rotation would interrupt spatial updating and force people to switch to an enduring representation system (Waller & Hodgson, 2006), and, as a result, only the reference direction defined by the learning heading would be used during retrieval of spatial relations in Experiment 2. To anticipate our results, Experiment 1 showed that both the reference directions defined by the initial and learning headings were used during retrieval, whereas only the reference direction defined by the learning heading was used in Experiment 2.

The results of a preliminary experiment with 16 participants were used to establish sample sizes for the current project. This experiment included only the I, L, and M conditions, but otherwise the materials, procedure, and results were similar to Experiment 1 (we decided not to report this experiment fully because it was completely subsumed by Experiment 1). The observed power was .81 in pointing error. To ensure adequate power in the present project, sample sizes of 24 were used.

Experiment 1

Method

Participants Twenty-four students (12 women) from Vanderbilt University and the Nashville community participated in this experiment in return for extra credits in psychology courses or monetary compensation.

Materials and design The experiment was conducted on a 21.5-inch Apple iMac desktop computer. The virtual environment (see Fig. 2) consisted of eight virtual objects (dog, ball,

---

[1] This representation would also have to be transient, otherwise it would be available in I and L conditions as well.
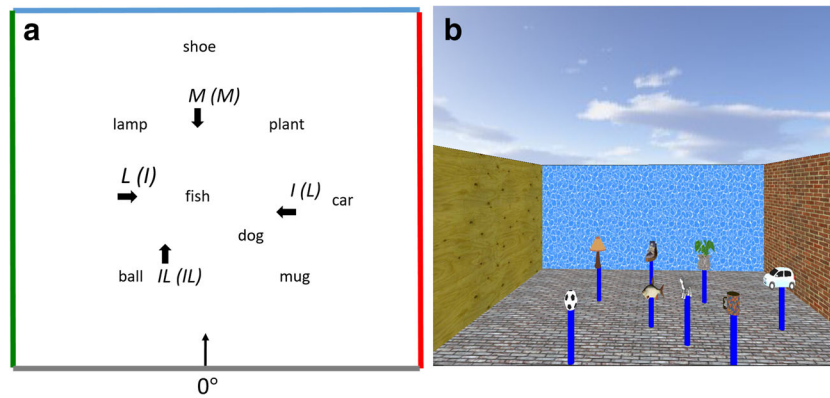
Fig. 2 **a** Plan view of the layout of objects. The *thin arrow* indicates the learning position and orientation in the learning phase. The *thick arrows* indicate the starting location and orientation for spatial updating trials in Experiments 1 and 2. The *letters* stand for the corresponding experimental condition (Experiment 2 in parentheses). An example trial in the *I* condition in Exp. 1 would be: I → plant → lamp, and point to car.

An example trial in the L condition in Exp. 1 would be: L → fish → shoe, and point to lamp. An example trial in the M condition in Exp. 1 would be: M → ball → mug, and point to fish. An example trial in the IL condition in Exp. 1 would be: IL → ball → lamp, and point to plant. **b** Participants' actual view in the learning phase. (Color figure online)

mug, fish, car, lamp, plant, and shoe) placed on identical 60-cm tall blue pillars.Objects were arranged in five columns, as shown in Fig. 2b. In addition, a square (7 m × 7 m × 3 m) virtual room surrounded the scene. The room floor was textured with a brick pattern. The four walls of the virtual room were textured with different colors and materials so that participants could use the texture of the wall to determine their initial heading at the beginning of a trial. The sky was textured with clouds in the learning phase, but was rendered uniformly blue in the test phase, so that participants could not use the sky to determine their position and orientation in the test phase. All participants learned the object locations from a fixed location and perspective (defined as 0°), which was 2 m away from the layout (see Fig. 2a). This viewing perspective ensured that participants could see all objects simultaneously.

We manipulated two headings in the current experiment: The initial heading, which was the heading participants faced at the beginning of a test trial in the virtual environment; and the imagined heading, which was the heading that participants were required to imagine they were facing before responding, and this imagined heading was always the same as the final heading participants occupied at the end of a test trial in the virtual environment. The learning heading was the fixed heading from which participants learned the layout in the learning phase (see Figs. 1 and 2).

To investigate the adopted reference frame, we used a 2 × 2 factorial design by manipulating the alignment between the initial heading and the imagined heading, and the alignment between the learning heading and the imagined heading as shown in Fig. 1. The orientations of the headings in each condition are listed in Table 1. Ten trials were constructed for each experimental condition, resulting in 40 total trials. These 40 trials were divided into 10 blocks of four trials each, with one trial from each condition in each block and presented randomly.

As stated previously, if participants used the learning heading or the initial heading to establish the reference direction, then performance in the L or the I condition, respectively, should be better than performance in the M condition. In addition, if participants were able to use aligned initial and learning headings to construct a more accurate representation at the time of responding or if misalignment with learning or initial headings produced processing costs, performance in the IL condition should be better than performance in the L or I condition.

To ensure that any significant differences observed between the aforementioned experimental conditions were not due to path complexity differences across conditions (Wan, Wang, & Crowell, 2013), we controlled the outbound path length (the shortest distance from the starting location to the first object, plus the shortest distance from the first to the second objects), outbound path turning angle (the shortest turning angle from the starting location to the first object, plus the shortest turning angle from the first to the second objects) and the correct pointing angle (the shortest angle from the second to the third object) across conditions.[2] These metrics are presented in Table 2.

Procedure

Learning phase The layout of eight objects was displayed (Fig. 2b) on a computer monitor, and the experimenter named each of the objects for the participants. After all of the objects were named, the participants were instructed to study the layout for 2 minutes. During learning, participants were told not to move from the study location. After learning, both the objects and pillars were hidden and one of the pillars, but not objects, would appear randomly. Participants named the

---

[2] The direction of turns (e.g., left-right-left or right-right-right) was not matched across conditions.

Table 1 Participants' headings (degrees) in the virtual environment across conditions by experiments

Experiment 1					Experiment 2			
	Learning	Initial	Final	Imagined		Learning	Initial	Final	Imagined
IL	0	0	0	0	IL	0	0	90	0
L	0	90	0	0	L	0	-90	-90	0
I	0	-90	-90	-90	I	0	90	0	90
M	0	180	90	90	M	0	180	0	-90

Note. The imagined heading was the same as the final heading in Experiment 1, but they were always misaligned by 90° in Experiment 2

corresponding object on that pillar. This learning sequence was repeated until the participant successfully named all the objects twice.

Test phase After learning the layout, participants performed the test trials in front of the same computer using keyboard and joystick. Participants started at the location corresponding to the trial condition (I, L, IL, or M). All objects and pillars were hidden, but room walls and the floor were present at the beginning so that participants could use the wall textures to identify their orientation in the virtual environment (see Fig. 2a). Participants could not change their orientation or position before they pulled the trigger on the joystick. After participants pulled the trigger, the room walls were removed, and one of the learned objects and the pillar beneath it appeared. Participants used the arrow keys on the keyboard to navigate to that object. Participants were instructed to first rotate the viewing perspective to face to the object, and then use the forward key to reach the object. The object disappeared upon arrival, and the second object would appear. Participants were instructed to release the forward key upon arrival and use the left or right key to look for the second object. Participants reached the second object in the same way. Upon arrival at the second object, everything disappeared, and a text message appeared at the center of screen, displaying the name of the third object to point to (e.g., "Please point to the shoe").

When participants saw the text message, they were told to imagine the environment from their final location (i.e., standing at the position and facing the orientation in the virtual environment they had been before the screen was blanked), and to use the joystick to point to the third object from that perspective. The pointing response was chosen in favor of a navigation response because the final heading was a key manipulation and we wanted to ensure that participants adopted and maintained their final heading during response. In addition, participants were told not to rotate the body during the test phase. If the joystick was deflected vertically or horizontally by more than 1 cm, the response would be recorded, and participants would be teleported to the next position and orientation corresponding to the experimental condition to start the next trial.

Before the test trials, participants performed three practice trials that were identical to the test trials, except that the objects in practice trials were randomly selected from the remembered layout.

Results and discussion

Previous research suggested that gender differences may exist in path integration (Kelly, McNamara, Bodenheimer, Carr, & Rieser, 2009), so we included gender in the following analysis. Pointing error and latency were analyzed in 2 (gender) × 2 (alignment between the learning and imagined headings,

Table 2 The means and standard deviations (in parenthesis) of the outbound path length, outbound path turning angle and correct pointing angle across conditions by experiments

	Experiment 1			
	I	L	M	IL
Outbound path length (meter)	4.58 (0.64)	5.24 (0.91)	5.24 (0.91)	4.65 (0.81)
Outbound path turning angle (degree)	165.25 (72.01)	165.25 (72.01)	165.25 (72.01)	165.28 (60.71)
Correct pointing angle (degree)	106.15 (38.28)	105.35 (34.87)	106.15 (38.28)	105.35 (34.87)
	Experiment 2			
	I	L	M	IL
Outbound path length (meter)	5.24 (0.91)	4.58 (0.64)	5.79 (0.85)	4.56 (0.96)
Outbound path turning angle (degree)	165.25 (72.01)	165.25 (72.01)	180 (0)	168.07 (83.59)
Correct pointing angle (degree)	104.21 (45.52)	106.15 (43.19)	104.31 (40.31)	105.34 (54.56)

referred to as learning-imagined) × 2 (alignment between the initial and imagined headings, referred to as initial-imagined) mixed ANOVAs (see Fig. 3), with gender as the between-subjects factor and learning-imagined and initial-imagined as within-subjects factors. For pointing error (see Fig. 3a), the main effect of gender was not significant, $F(1, 22) = 3.92$, $MSE = 970.46$, $p = .06$, $\eta^2 = .30$, but the main effects of learning-imagined and initial-imagined were significant, $F(1, 22) = 9.44$, $MSE = 444.71$, $p = .006$, $\eta^2 = .15$; $F(1, 22) = 23.79$, $MSE = 60.62$, $p < .001$, $\eta^2 = .52$. In addition, all of the two-way interactions were significant: Learning-Imagined × Initial-Imagined, $F(1, 22) = 9.83$, $MSE = 69.12$, $p = .005$, $\eta^2 = .30$; Learning-Imagined × Gender, $F(1, 22) = 5.90$, $MSE = 444.71$, $p = .024$, $\eta^2 = .21$; Initial-Imagined × Gender, $F(1, 22) = 6.61$, $MSE = 60.27$, $p = .017$, $\eta^2 = .23$. The significance levels of the following t tests were Bonferroni adjusted (the p value must be less than or equal to .025 to be deemed significant).

Collapsing across gender, pairwise comparisons showed that pointing error was higher in the M condition than in the I and L conditions, $ts(23) > 3.52$, $ps < .002$, suggesting that participants used both the learning and the initial headings to establish reference directions in the current task. The IL condition did not differ from the I or the L condition, $t(23) = 1.32$, $p = .20$; $t(23) = 1.92$, $p = .07$, respectively. The significant interaction between learning-imagined and initial-imagined and the pattern of pointing error suggested that when the imagined heading was aligned with the learning heading, the alignment between the imagined heading and the initial heading did not play a role. On the other hand, when the imagined heading was misaligned with the learning heading, the alignment between the imagined heading and the initial heading affected the pointing error significantly. This interaction might be due to a floor effect such that the lowest average pointing error that could be achieved with the current pointing device is the pointing error of the L condition (~30°), and therefore the performance in the IL condition could not be better than the L condition. We discounted this possibility because two previous studies from our lab (He & McNamara, 2017; He et al., 2017) showed that the average pointing error could be as low as 20° with the current pointing device. Combined with the comparable performance among the I, L, and IL conditions, we concluded that participants used only one reference direction when both reference directions (defined by the initial and learning headings) could be utilized simultaneously.

Within gender, when the imagined heading was misaligned with the initial heading, women's learning heading effect (difference between L and M conditions: $Diff = -31.89$, $SE = 9.15$), $t(11) = 3.33$, $p = .007$, was larger than men's ($Diff = -4.75$, $SE = 4.30$), $t(11) = 1.05$, $p = .32$. When the imagined heading was misaligned with the learning heading, women's initial heading effect (difference between I and M conditions: $Diff = -20.05$, $SE = 3.04$), $t(11) = 6.31$, $p < .001$, was larger than men's ($Diff = -5.64$, $SE = 4.31$), $t(11) = 1.25$, $p = .23$. When the imagined heading was aligned with both headings (IL condition), performance was not significantly different from the I or L condition for women, $ts(11) < 2.35$, $ps > .04$, or men, $ts(11) < 0.08$, $ps > .94$.

For pointing latency (see Fig. 3b), only the main effect of learning-imagined was significant, $F(1, 22) = 8.04$, $MSE = 1.28$, $p = .01$, $\eta^2 = .26$, suggesting that participants responded
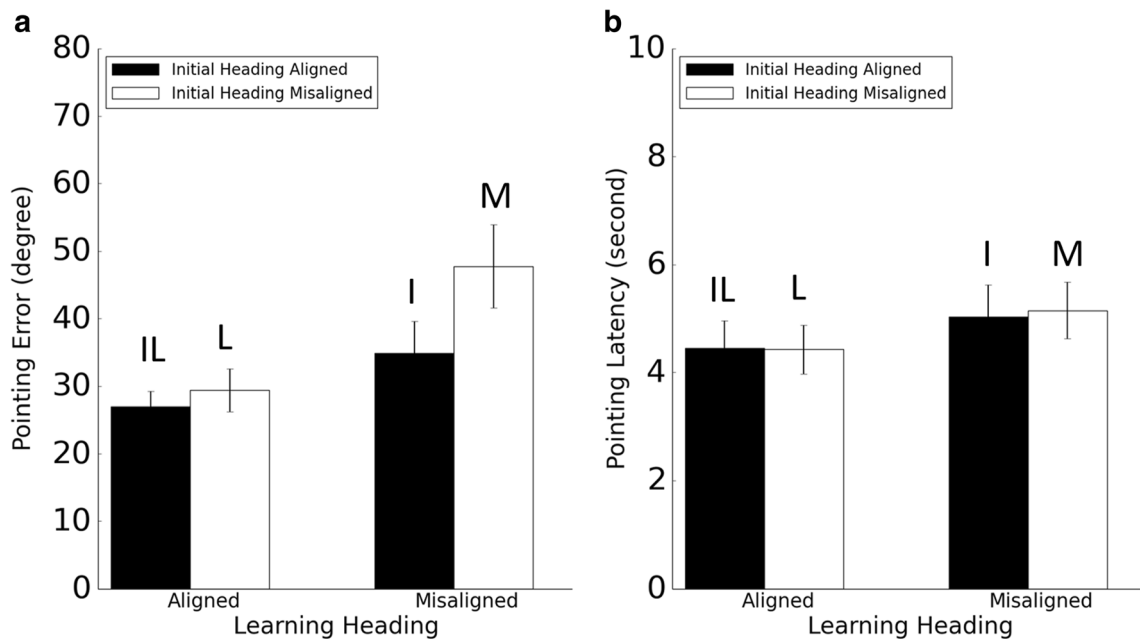

Fig. 3 Pointing error (**a**) and latency (**b**) across gender in Experiment 1. *Error bars* are ±1 *SEM* estimated from data within conditions. The *letters above the bars* identify the corresponding experimental conditions. *Alignment* and *misalignment* refer to the relation with the imagined heading

faster when the imagined heading was aligned with the learning heading.

In sum, the results from Experiment 1 showed that during spatial updating without idiothetic cues, participants used both the learning heading and the initial heading to establish reference directions but could not use them simultaneously. Within gender, we found that women relied on these two headings to establish reference directions, but this effect was not significant for men. Another interpretation of this gender difference is that men did rely on these two headings to establish reference directions, but when the imagined heading was not aligned with these headings (M condition), men were able to mentally rotate the layout of objects from the learning heading efficiently. We discuss the gender difference in more detail in the General Discussion.

Experiment 2

We assumed that the allocentric reference system established at the initial heading was transient, whereas the allocentric reference system established at the learning heading was enduring. To test this hypothesis, we used the same paradigm as in Experiment 1, but when participants reached the second object, they were required to imagine that they were facing 90° left or right to their final heading in the virtual environment and to point to the target object relative to this imagined heading. This mental rotation could interrupt spatial updating and encourage people to switch to an enduring representation (Waller & Hodgson, 2006). If the reference system established at the initial heading is a transient representation and the one established at the learning heading is an enduring representation, the significant difference between the L and M conditions should remain, whereas the difference between the I and M conditions should decrease or become insignificant.

Method

Participants Twenty-four students (12 women) from Vanderbilt University and the Nashville community participated in this experiment in return for extra credits in psychology courses or monetary compensation.

Materials and design The materials and design in Experiment 2 were similar to those in Experiment 1, except for the starting locations and orientations of the I and L conditions (see Fig. 2a and Table 1), as well as the path properties in each experimental condition (Table 2). The outbound path turning angles listed in Table 2 did not take into account the 90° mental rotation at the end of the outbound path, and the correct pointing angle was calculated based on the heading that participants were required to adopt after the mental rotation. All experimental conditions in the current experiment

were the same as in Experiment 1 after mental rotation (e.g., in the I condition, after the mental rotation, participants imagined a heading that was aligned with their initial heading, but not aligned with the learning heading).

Procedure The learning phase was identical to Experiment 1. The test phase was very similar to Experiment 2, except that when participants reached the second object, the text message would ask them to imagine facing 90° left or right relative to their final heading, and then point to a third object ("Please imagine you are facing 90° to your left of your heading in the virtual environment. Point to the dog"). In the L and I conditions, participants were required to imagine to face 90° to the right; in the IL and M conditions, participants were required to imagine to face 90° to the left.

Results and discussion

Pointing error and latency were analyzed in 2 (gender) × 2 (learning-imagined) × 2 (initial-imagined) mixed ANOVAs (see Fig. 4). For pointing error (Fig. 4a), the main effect of gender was significant, $F(1, 22) = 22.33$, $MSE = 862.80$, $p < .001$, $\eta^2 = .50$, as were the main effects of learning-imagined and initial-imagined, $F(1, 22) = 47.14$, $MSE = 267.40$, $p < .001$, $\eta^2 = .68$; $F(1, 22) = 10.77$, $MSE = 84.76$, $p = .003$, $\eta^2 = .32$. The interaction between learning-imagined and gender was significant, $F(1, 22) = 21.85$, $MSE = 444.71$, $p = .024$, $\eta^2 = .21$, but the interaction between initial-imagined and gender, $F(1, 22) = 2.67$, $MSE = 84.76$, $p = .11$, $\eta^2 = .10$, and the interaction between learning-imagined and initial-imagined, $F(1, 22) = 3.82$, $MSE = 113.74$, $p = .06$, $\eta^2 = .14$, were not significant. The three-way interaction was significant, $F(1, 22) = 7.62$ $MSE = 113.74$, $p = .01$, $\eta^2 = .25$. The significance levels of all the following t tests were Bonferroni adjusted ($\alpha = .025$).

Across gender, pointing error was significantly higher in the M condition than in the L condition, $t(23) = 3.98$, $p < .001$, but was lower in the M condition than in the I condition, $t(23) = 3.36$, $p = .003$. There was no significant difference between the IL and the L conditions, $t(23) = .59$, $p > .56$, but performance was better in the IL condition than in the I condition ($t(23) = 4.75$, $p < .001$).

Within gender, when the imagined heading was misaligned with the initial heading, women's learning heading effect (difference between L and M conditions: $Diff = -28.25$, $SE = 6.51$), $t(11) = 4.43$, $p < .001$, was larger than men's ($Diff = -9.06$, $SE = 5.72$), $t(11) = 1.58$, $p = .14$. When the imagined heading was misaligned with the learning heading, women's performance in the initial heading aligned condition (the I condition) was no better than chance level (89.04° vs. 90°), $t(11) = .17$, $p = .86$; seven out of 12 female participants' pointing error in the I condition exceeded 90°, which led to significantly worse performance than in the M condition (difference between I and M
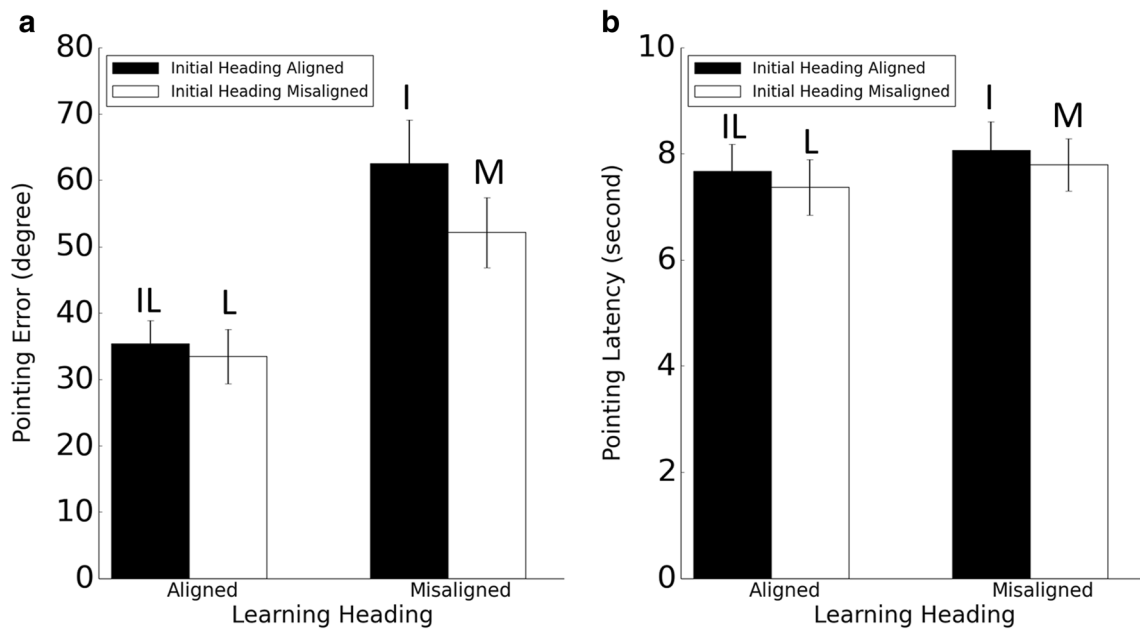
Fig. 4 Pointing error (**a**) and latency (**b**) across gender in Experiment 2. *Error bars* are ±1 *SEM* estimated from data within conditions. The *letters above the bars* stand for the corresponding experimental conditions. *Alignment* and *misalignment* refer to the relation with the imagined heading

conditions: $Diff° = 19.51$, $SE = 3.59$), $t(11) = 5.43$, $p < .001$. This trend was not significant in men (difference between I and M conditions: $Diff° = 1.33$, $SE = 3.50$), $t(11) = .38$, $p = .71$.

For pointing latency (see Fig. 4b), none of the main effects or interactions was significant.

In sum, the results from Experiment 2 confirmed our prediction that only the learning heading was used to establish the reference direction when the spatial updating process was interrupted. However, we did not expect that performance in the I condition would be significantly worse than in the M condition. The overall poor performance in the I condition was mainly caused by women's chance-level performance. We conjectured that women might have imagined a heading different from the instructed imagined heading in the I condition while responding. To test this hypothesis, we fit women's data in the I condition with the imagined heading as 0°, 90° (instructed imagined heading, pointing error = 89.04°), -90°, or 180°. We found that 0° fit the data the best, $F(3, 33) = 33.15$, $MSE = 426.72$, $p < .001$, $\eta^2 = .75$, producing a mean pointing error of 47.95°, and this level of performance was significantly better than if the instructed imagined heading were used as the imagined heading, $t(33) = 4.73$, $p < .001$. We also fit women's data in the other three conditions with the same four imagined headings, and found that 0° fit the data the best in all conditions, $Fs(3, 33) > 30$, $ps < .001$, except the M condition, in which if 0° were used as the imagined heading, performance (pointing error = 57.17°) was not significantly better than if the instructed imagined heading were used as the imagined heading (pointing error = 69.52°), $t(33) = 1.19$, $p = .24$. The combined results implied that women had difficulty

imagining a heading other than the learning heading of 0° when mental rotation was required.

General discussion

The current study investigated the reference system in spatial updating when idiothetic cues were not available. In two experiments, participants first learned a layout of eight objects from a fixed perspective in a virtual environment and were placed in the same virtual environment to navigate to two of the learned objects before pointing to a third object. The navigation was via keyboard, and therefore the idiothetic cues were reduced to a minimum. We manipulated the alignment between the imagined heading and the initial heading, and the alignment between the imagined heading and the learning heading, to reveal the reference system or systems used in the task. Results from Experiment 1 indicated that participants used the initial heading and the learning heading to establish reference directions but did not use both reference directions simultaneously when the imagined heading was aligned with both. Results from Experiment 2 showed that when participants needed to imagine a heading different from their final heading before responding, pointing performance was still affected by alignment with the learning heading but not by alignment with the initial heading.

The initial heading effect in Experiment 1 suggests that the perspective people first experience plays an important role in spatial updating without idiothetic cues. A similar initial heading effect can be found in Meilinger, Frankenstein, Watanabe,

Bülthoff, and Hölscher's (2014) study, in which the authors showed that participants used the initial orientation to establish the reference direction even when they were allowed to explore the environment in any direction. However, the initial heading effect in the current study is still surprising because, unlike other similar studies (Meilinger, Riecke & Bülthoff, 2014; Mou, McNamara, & Zhang 2013), participants never saw the layout of objects or any object in the presence of the environmental cues from the initial heading. and they experienced several initial headings, not only one. Yet participants still represented the spatial relations using a reference direction parallel to the initial heading. To our knowledge, this is the first evidence showing that even without explicit instructions to imagine an unseen perspective of a scene, people would do so and this imagery could facilitate information retrieval during spatial updating.

The imagined heading in the I condition was orthogonal to the learning heading, and previous research has shown that when the imagined heading is orthogonal to the dominant reference direction, performance can be as good as when it is aligned with the reference direction (Rump & McNamara, 2013; Shelton & McNamara, 1997, 2001). One interpretation of that effect is that orthogonal spatial transformations can be efficient (Street & Wang, 2014). The good performance in the I condition in Experiment 1 cannot be accounted for by efficient orthogonal transformations alone, because performance in this condition was significantly better than in the M condition, which also had an imagined heading orthogonal to the learning heading. Therefore, we conclude that separate reference systems were established at the learning heading and the initial heading in Experiment 1, not a single one at the learning heading.

In Experiment 2, we assume that a reference system was established at the initial heading but not used for pointing judgments, because the continuous spatial updating process was disrupted and participants switched to an enduring representation (Waller & Hodgson, 2006). Mental rotation in Experiment 2 was only required at the end of the outbound path, so participants should have had sufficient time to encode the information about the initial heading. Hence, the absence of the initial heading effect was probably not due to interference in the encoding phase. However, another possible explanation of the results of Experiment 2 is that participants did not encode the initial heading, because they learned that a mental rotation would be required at the end and decided that memorizing the initial heading was not useful. An experiment in which mental rotation is required on a random subset of trials could be useful to distinguish between these two explanations. Nonetheless, results from Experiment 2 suggested that the reference system defined by the initial heading was transient and not committed to long-term memory.

The learning heading effect, on the other hand, was consistent throughout the two experiments. The use of the learning heading to establish the reference direction to organize object-to-object and self-to-object spatial relations is a well-established finding (Kelly et al., 2007; Kelly & McNamara, 2008; Mou et al., 2004; Shelton & McNamara, 1997, 2001; Waller et al., 2002), and this representation is considered to be stored in long-term memory (Mou et al., 2004; Waller & Hodgson, 2006). The findings from the current study further demonstrated that the representation of the learning heading was enduring and was used in spatial updating even when idiothetic cues were not available.

The comparable performance in the IL, I, and L conditions in Experiment 1 implied that participants were not able to use two aligned reference directions in the IL condition to point more accurately, and that misalignment with one reference direction did not produce costs in pointing. We hypothesize that the two allocentric headings might have been used in the following manner: When participants saw the walls at the beginning of a trial, they imagined the layout of the learned objects from that perspective (the initial heading), and hence two representations of the layout were available, one defined by the initial heading and the other defined by the learning heading (formed at the time of learning). Once they reached the end of the outbound path, participants retrieved the layout of objects from the corresponding perspectives if the imagined heading was aligned with the initial or the learning heading. This proposed mechanism could explain why participants in the current study could only use one reference direction at a time, which contrasts with the findings from Mou et al.'s (2004) study, in which participants could use two reference frames simultaneously to generate a more accurate representation. Because both spatial representations in the current study were hypothesized to be allocentric, whereas one was egocentric and the other was allocentric in Mou et al.'s (2004) study, the combined findings imply that spatial representations of the same nature may compete for cognitive resources during retrieval.

An alternative explanation of our findings is that the learning heading or the initial heading or both were used to establish egocentric reference systems (e.g., Greenauer & Waller, 2008; Waller & Hodgson, 2006).[3] For example, it is possible that when participants memorized the layout of the objects, they represented self-to-object spatial relations in long-term memory, using a reference direction parallel to egocentric front. At the beginning of a test trial, participants might have formed a transient egocentric representation of the object array. If one assumes, as we have done, that pointing judgments based on represented spatial relations or familiar views are more efficient than pointing judgments based on inferred spatial relations or unfamiliar views, this model makes the same predictions as does ours in the present experiments (Wang, 2016). Although we cannot rule out this alternative

---

[3] We are grateful to an anonymous reviewer for this alternative explanation.

explanation, it does not contradict our conclusion that both the learning and the initial headings were used to establish reference directions and that those reference frames could not be combined to generate a more accurate representation. The two reference frames established at the learning and initial headings, respectively, are distinguished by their susceptibility to interruption: The enduring reference frame is used consistently, whereas the transient reference frame is discarded or suppressed when spatial updating is disrupted. Hodgson and Waller (2006) proposed that the transient spatial representation was more precise than the enduring representation, but, in our model, transiency does not correlate with precision.

Although there is some evidence showing that gender differences may exist in path integration (Kelly et al., 2008), we did not anticipate gender differences in the current study. Contrary to our anticipation, the gender difference was consistent throughout the two experiments, with men outperforming women when the imagined heading was not aligned with the learning or initial heading (M condition). One possible explanation of this effect is that men are more efficient than women in terms of orthogonal spatial transformation, so that men could more easily transform the learning heading to align with the imagined heading than women (i.e., mental rotation differences between genders; see Linn & Petersen. 1985, for a review). A similar male advantage was found when participants were asked to determine whether a map, presented at different orientations, correctly reproduced the spatial relations of the buildings participants had visited before (Iachini, Ruotolo, & Ruggiero, 2009). Another possibility is that men had more experience than women, on average, with first-person video games (Jansz & Tanis, 2007), and this experience facilitated their performance in our task, which shares many features with such games.

To conclude, we found that participants used the initial heading and learning heading to establish reference directions in spatial updating when idiothetic cues were not available. When participants were required to point to target objects relative to imagined headings that differed from their current headings in the virtual environment, continuous spatial updating was interrupted, and only the reference direction established at the learning heading was used. This pattern of results indicates that the learning heading corresponds to an enduring representation and the initial heading corresponds to a transient representation. When comparing the reference directions used in the absence or presence of idiothetic cues, we find that the learning heading is used in both scenarios, but people use the heading they eventually occupy (final heading) instead of the initial heading as a reference direction when idiothetic cues are available (Mou et al., 2004; Kelly et al., 2007). These differences in reference frame selection based on idiothetic cues may reflect a more general phenomenon, which is that our cognitive systems are flexible and can be dynamically adjusted based on situational demands (e.g., Fischer & Plessow, 2015; Graesser, Mills, & Zwaan, 1997).

References

Allen, G. L., & Haun, D. B. M. (2004). Proximity and precision in spatial memory. In G. L. Allen (Ed.), *Human spatial memory: Remembering where* (pp. 41–63). Mahwah, NJ: Erlbaum.

Amorim, M.-A., Glasauer, S., Corpinot, K., & Berthoz, A. (1997). Updating an object's orientation and location during nonvisual navigation: A comparison between two processing modes. *Perception & Psychophysics, 59*(3), 404–418. doi:10.3758/BF03211907

Amorim, M.-A., & Stucchi, N. (1997). Viewer- and object-centered mental explorations of an imagined environment are not equivalent. *Cognitive Brain Research, 5*(3), 229–239. doi:10.1016/S0926-6410(96)00073-0

Farrell, M. J., & Robertson, I. H. (1998). Mental rotation and automatic updating of body-centered spatial relationships. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(1), 227–233. doi:10.1037/0278-7393.24.1.227

Fischer, R., & Plessow, F. (2015). Efficient multitasking: parallel versus serial processing of multiple tasks. *Frontiers in Psychology*, 6. doi:10.3389/fpsyg.2015.01366.

Graesser, A. C., Mills, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48*(1), 163–189.

Greenauer, N., & Waller, D. (2008). Intrinsic array structure is neither necessary nor sufficient for nonegocentric coding of spatial layouts. *Psychonomic Bulletin & Review, 15*(5), 1015–1021.

He, Q., & McNamara, T. P. (2017). Spatial updating strategy affects the reference frame in path integration. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-017-1307-7

He, Q., McNamara, T. P., & Kelly, J. W. (2017). Environmental and idiothetic cues to reference frame selection in path integration. In T. Barkowsky et al. (Eds.), *Spatial cognition X*. Berlin: Springer.

Hodgson, E., & Waller, D. (2006). Lack of set size effects in spatial updating: Evidence for offline updating. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(4), 854–866. doi:10.1037/0278-7393.32.4.854

Holmes, M. C., & Sholl, M. J. (2005). Allocentric coding of object-to-object relations in overlearned and novel environments. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*(5), 1069–1087. doi:10.1037/0278-7393.31.5.1069

Iachini, T., Ruotolo, F., & Ruggiero, G. (2009). The effects of familiarity and gender on spatial representation. *Journal of Environmental Psychology, 29*(2), 227–234.

Kelly, J. W., Avraamides, M. N., & Loomis, J. M. (2007). Sensorimotor alignment effects in the learning environment and in novel environments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(6), 1092–1107. doi:10.1037/0278-7393.33.6.1092

Kelly, J. W., & McNamara, T. P. (2008). Spatial memories of virtual environments: How egocentric experience, intrinsic structure, and extrinsic structure interact. *Psychonomic Bulletin & Review, 15*(2), 322–327.

Kelly, J. W., McNamara, T. P., Bodenheimer, B., Carr, T. H., & Rieser, J. J. (2008). The shape of human navigation: How environmental

geometry is used in maintenance of spatial orientation. *Cognition, 109*(2), 281–286. doi:10.1016/j.cognition.2008.09.001

Kelly, J. W., McNamara, T. P., Bodenheimer, B., Carr, T. H., & Rieser, J. J. (2009). Individual differences in using geometric and featural cues to maintain spatial orientation: Cue quantity and cue ambiguity are more important than cue type. *Psychonomic Bulletin & Review, 16*(1), 176–181. doi:10.3758/PBR.16.1.176

Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: definitions, distinctions, and interconnections. In C. Freksa, C. Habel, & K. F. Wender (Eds.), *Spatial cognition: An interdisciplinary approach to representing and processing spatial knowledge, LNAI 1404* (pp. 1–17). Berlin: Springer-Verlag. doi:10.1007/3-540-69342-4_1

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. doi:10.1111/1467-9280.00058

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56*(6), 1479–1498.

Jansz, J., & Tanis, M. (2007). Appeal of playing online first person shooter games. *CyberPsychology & Behavior, 10*(1), 133–136.

McNamara, T. P. (2003). How are the locations of objects in the environment represented in memory? In C. Freksa, W. Brauer, C. Habel, & K. F. Wender (Eds.), *Spatial cognition III: Routes and navigation, human memory and learning, spatial representation and spatial reasoning, LNAI 2685* (pp. 174–191). Berlin: Springer-Verlag.

Meilinger, T., Frankenstein, J., Watanabe, K., Bülthoff, H. H., & Hölscher, C. (2014). Reference frames in learning from maps and navigation. *Psychological Research, 79*(6), 1000–1008. doi: 10.1007/s00426-014-0629-6

Meilinger, T., Riecke, B. E., & Bülthoff, H. H. (2014). Local and global reference frames for environmental spaces. *The Quarterly Journal of Experimental Psychology, 67*(3), 542–569. doi: 10.1080/17470218.2013.821145

Mou, W., & McNamara, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(1), 162–170. doi:10.1037/0278-7393.28.1.162

Mou, W., McNamara, T. P., Valiquette, C. M., & Rump, B. (2004). Allocentric and egocentric updating of spatial memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(1), 142–157. doi:10.1037/0278-7393.30.1.142

Mou, W., McNamara, T. P., & Zhang, L. (2013). Global frames of reference organize configural knowledge of paths. *Cognition, 129*(1), 180–193. doi:10.1016/j.cognition.2013.06.015

Riecke, B. E., Heyde, M. V. D., & Bülthoff, H. H. (2005). Visual cues can be sufficient for triggering automatic, reflexlike spatial updating. *ACM Transaction on Applied Perception, 2*(3), 183–215. doi:10.1145/1077399.1077401

Riecke, B. E., van Veen, H. A. H. C., & Bülthoff, H. H. (2002). Visual homing is possible without landmarks: A path integration study in virtual reality. *Presence Teleoperators and Virtual Environments, 11*(5), 443–473. doi:10.1162/105474602320935810

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. doi:10.1037/0278-7393.15.6.1157

Ruddle, R. A., Volkova, E., & Bülthoff, H. H. (2011). Walking improves your cognitive map in environments that are large-scale and large in

extent. *ACM Transactions on Computer-Human Interaction, 18*(2), 1–10:20. doi:10.1145/1970378.1970384

Rump, B., & McNamara, T. P. (2013). Representations of interobject spatial relations in long-term memory. *Memory & Cognition, 41*(2), 201–213. doi:10.3758/s13421-012-0257-6

Shelton, A. L., & McNamara, T. P. (1997). Multiple views of spatial memory. *Psychonomic Bulletin & Review, 4*(1), 102–106. doi:10.3758/BF03210780

Shelton, A. L., & McNamara, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology, 43*(4), 274–310. doi:10.1006/cogp.2001.0758

Street, W. N., & Wang, R. F. (2014). Differentiating spatial memory from spatial transformations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(2), 602–608. doi:10.1037/a0035279

Valiquette, C. M., McNamara, T. P., & Smith, K. (2003). Locomotion, incidental learning, and the selection of spatial reference systems. *Memory & Cognition, 31*(3), 479–489. doi:10.3758/BF03194405

Waller, D., & Hodgson, E. (2006). Transient and enduring spatial representations under disorientation and self-rotation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(4), 867–882. doi:10.1037/0278-7393.32.4.867

Waller, D., & Hodgson, E. (2013). Sensory contributions to spatial knowledge of real and virtual environments. In F. Steinicke, Y. Visell, J. Campos, & A. Lécuyer (Eds.), *Human walking in virtual environments* (pp. 3–26). New York: Springer. doi:10.1007/978-1-4419-8432-6_1

Waller, D., Loomis, J. M., & Haun, D. B. M. (2004). Body-based senses enhance knowledge of directions in large-scale environments. *Psychonomic Bulletin & Review, 11*(1), 157–163. doi:10.3758/BF03206476

Waller, D., Montello, D. R., Richardson, A. E., & Hegarty, M. (2002). Orientation specificity and spatial updating of memories for layouts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(6), 1051–1063. doi:10.1037/0278-7393.28.6.1051

Wan, X., Wang, R. F., & Crowell, J. A. (2013). Effects of basic path properties on human path integration. *Spatial Cognition and Computation, 13*(1), 79–101

Wang, R. F. (2016). Building a cognitive map by assembling multiple path integration systems. *Psychonomic Bulletin & Review, 23*(3), 692–702. doi:10.3758/s13423-015-0952-y

Wang, R. F., Crowell, J. A., Simons, D. J., Irwin, D. E., Kramer, A. F., Ambinder, M. S.,…Hsieh, B. B. (2006). Spatial updating relies on an egocentric representation of space: Effects of the number of objects. *Psychonomic Bulletin & Review, 13*(2), 281–286. doi:10.3758/BF03193844

Xiao, C., Mou, W., & McNamara, T. P. (2009). Use of self-to-object and object-to-object spatial relations in locomotion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(5), 1137–1147. doi:10.1037/a0016273

Open practices

All data have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/cw2zx/?view_only=3cb01f7698cb430298360e0559d0f520.