Running Head: PATH INTEGRATION MODELS

Executing the Homebound Path is a Major Source of Error in Homing by Path Integration

Elizabeth R. Chrastil

William H. Warren

Brown University

Cognitive, Linguistic, & Psychological Sciences

University of California, Irvine

Neurobiology & Behavior

Author Note: Preparation of this manuscript was supported by National Science Foundation awards BCS-0214383 and BCS-0843940. The authors would also like the thank Mike Tarr and Scott Bell for reading early drafts of the manuscript, and members of the VENLab for assistance with conducting this experiment.

Correspondence concerning this article should be addressed to Elizabeth Chrastil

Current Address: chrastil@uci.edu

Abstract

Path integration—the constant updating of position and orientation in an environment—is an important component of spatial navigation, however, its mechanisms are poorly understood. The aims of this study are 1) to test the encoding-error model of path integration, which focuses solely on encoding as a potential source of error, and 2) to develop a model of path integration that best predicts path integration errors. We tested the encoding-error model by independently measuring participants' encoding errors in distance and angle reproduction tasks, and then using those reproduction errors to predict individual participants' errors in a triangle completion task. We sampled the distribution of encoding errors using Monte Carlo methods to predict the homebound path, and then compared the predictions to observed triangle completion behavior. The correlation between predicted errors and actual errors in the triangle completion task was extremely weak, whereas an alternative model using execution error alone was sufficient to describe the observed errors. A model incorporating both encoding and execution errors best described the triangle completion errors. These results suggest that errors in executing the response may contribute more to overall errors in path integration than do encoding errors, challenging the assumption that errors reflect encoding alone. Errors in triangle completion might not arise from failing to know where you are, but from an inability to get back home.

Keywords: navigation, self-motion, idiothetic, virtual environments, cognitive map

Public Significance Statement: This study challenges the long-standing assumption that homing errors in path integration stem from encoding the outbound path. Instead, this study demonstrates that the largest source of error is carrying out the homebound trajectory. These findings will help us understand more about the mechanisms underlying navigational systems.

Introduction

Human Path Integration

To navigate successfully through an environment, an animal needs to sense its own position and orientation with respect to places in that environment. *Path integration* is the continuous updating of position and orientation by integrating changes in position, velocity and acceleration based on idiothetic and visual motion information (H. Mittelstaedt & Mittelstaedt, 1982; M.-L. Mittelstaedt & Mittelstaedt, 1980). Some researchers have proposed that path integration is the basis for building up spatial knowledge of the environment, such as a metric *cognitive map* (Gallistel, 1990; McNaughton, Battaglia, Jensen, Moser, & Moser, 2006; Wang, 2015) or a labeled *cognitive graph* (Chrastil & Warren, 2014b; Warren, 2019; Warren, Rothman, Schnapp, & Ericson, 2017), by registering the distances and angles between places and landmarks. Little is known, however, about just how accurate and stable human path integration is, and whether it could provide a basis for deriving such spatial knowledge. Few satisfactory models have been proposed that account for the systematic errors seen in human path integration. The goal of this study is to test the contributions of potential sources of systematic error in human path integration in several alternative models.

Some animals, such as desert ants and nocturnal hamsters, have shown a remarkable ability to return to the nest or home location by means of path integration, known as *homing* (Etienne, Maurer, & Saucy, 1988; M.-L. Mittelstaedt & Mittelstaedt, 1980; Seguinot, Maurer, & Etienne, 1993; Wehner & Wehner, 1986; Wittlinger, Wehner, & Wolf, 2006). These animals may have developed accurate path integration because they live in environments without stable visible landmarks. Humans, on the other hand, are highly visual animals, and appear to rely on landmarks and other external cues for guidance, known as *landmark-based navigation*. Recent

evidence suggests that while humans have coarse path integration abilities, visual landmarks dominate shortcut and homing behavior (Foo, Duchon, Warren, & Tarr, 2007; Foo, Warren, Duchon, & Tarr, 2005; Zhao & Warren, 2015b). When landmarks are noticeably unreliable, however, people fall back on a strategy of path integration, although the latter does not appear to be an automatic "back up" system running in the background (Cheng, Shettleworth, Huttenlocher, & Rieser, 2007; Zhao & Warren, 2015a). Thus, humans must have some capacity to perform path integration.

The most common method for investigating human path integration is a homing task known as triangle completion (Fujita, Loomis, Klatzky, & Golledge, 1990; Kearns, Warren, Duchon, & Tarr, 2002; Klatzky, Loomis, Beall, Chance, & Golledge, 1998; Klatzky et al., 1990; Loomis et al., 1993; Peruch, May, & Wartenberg, 1997; Tcheang, Bülthoff, & Burgess, 2011). In this task, an experimenter guides a research participant down one leg of a triangle, then takes them through a turn, and finally guides them down a second leg of the triangle. At that point, the participant must determine both the distance and direction back to the starting point ('home') in order to complete the third leg of the triangle. Triangle completion studies have found systematic errors in path integration performance, such as treating outbound legs as if they are equal sides of an isosceles triangle, compressing responses by overturning small angles and underturning large angles, and overshooting short distances and undershooting long distances.

To help explain these systematic errors, Fujita, et al. (Fujita, Klatzky, Loomis, & Golledge, 1993; Loomis et al., 1993) broke down such path-completion tasks into five main components. The first three elements—sensing, creating a trace of the route, and forming a survey representation of the outbound path segments—make up the broader process of "encoding" the outbound path traveled by the animal. The fourth component is integrating the

outbound segments to determine the appropriate return trajectory back to the starting location. In other words, once the animal has experienced the outbound path, the integration step computes the necessary trajectory for a novel path to a particular goal, namely, the start position. In the fifth final component, the animal must execute that homebound trajectory. It is possible for systematic errors to accumulate during any of these processes.

The Encoding-Error Model of Path Integration

Fujita et al. (1993) proposed an encoding-error model that accounts for the systematic errors observed in path integration. Specifically, they posited that the major component of systematic error stems from encoding the outbound path. This model further theorizes that integrating the outbound segments to form the homebound trajectory—and executing that trajectory—play no role in overall path integration error. Thus, their model assumes that once the path integration system encodes the values of the outbound path it produces no other systematic errors.

The encoding-error model has four underlying assumptions: 1) the internal representation of the path satisfies the axioms of Euclidean geometry, 2) distances are encoded by just one function, so that equal distances are encoded the same way, 3) angles are also encoded by one function, and 4) there is no systematic error in either the integration of path segments or execution of the homeward trajectory (Fujita et al., 1993). Fujita et al. estimated the linear encoding functions in distance and angle reproduction tasks, then used those general functions to predict average path integration errors.

Although the initial test of the encoding-error model had some success, the assumption that there is no integration or execution error has proven problematic. Some studies have found little error in execution of a computed trajectory (Jurgens, Nasios, & Becker, 2003; Riecke, Van

Veen, & Bülthoff, 2002), but others have demonstrated significant bias in production of simple tasks (Bakker, Werkhoven, & Passenier, 1999, 2001; Chrastil & Warren, 2017; Israël, Sievering, & Koenig, 1995; Jetzschke, Ernst, Moscatelli, & Boeddeker, 2016; Klatzky et al., 1990), in violation of assumption 4. For example, Bakker, Werkhoven, & Passenier (1999, 2001) demonstrated that production errors for verbally-specified turns of 90°, 180°, or 270° ranged from approximately 5°–45° in conditions with visual, vestibular, and proprioceptive information, and from approximately 20°–120° in a purely visual task. Participants were verbally instructed which angle they should turn, then used a rotating turntable to turn through the specified angle. This task likely taps execution errors since these turn angles are orthogonal to each other and constitute body axes in an egocentric (viewer-centered) reference frame. However, even tasks that on the surface appear to measure pure execution error, such as turning 90°, might reflect some combination of encoding error and execution error, with accurate performance if the two elements are calibrated to compensate each other.

More complex path integration tasks have also cast doubt on the idea of minimal execution error. For example, Wan et al. (Wan, Wang, & Crowell, 2013) found that path integration errors were related to the length of the correct homebound trajectory, indicating that execution of that trajectory could contribute to total path integration errors. Using similar triangles, Harootonian et al. (in press) found that participants tended to undershoot homebound distances as the triangles got bigger, but there was no change in the turn angle – which would be the same regardless of size. This result suggests that people could encode the triangles correctly, since the turn angle did not vary, but they had difficulty executing different homebound path lengths. Avraamides et al. (Avraamides, Klatzky, Loomis, & Golledge, 2004) found that verbal and pointing responses lead to different patterns of error during imagined spatial updating,

suggesting that response mode is an important factor in path integration. Finally, Chen et al. (Chen, He, Kelly, Fiete, & McNamara, 2015) showed that homebound paths are affected by environmental rescaling, indicating that these trajectories are not simply an executed motor plan, but rather depend on online information to create the homebound path. Together, these findings all point to a potential contribution of execution error during homing.

The assumption of a constant linear encoding function that fits all contexts (assumptions 2 and 3) has also been called into question. Context-free encoding implies that the same encoding function should apply when all of the leg lengths are short as when they are all relatively long. In violation of this assumption, Klatzky et al. (Klatzky, Beall, Loomis, Golledge, & Philbeck, 1999) found that a general linear encoding function was not sufficient for all contexts. Distance reproduction tasks generally demonstrate a compressed distance function, such that small distances are overestimated and large distances are underestimated (Harris, Jenkin, & Zikovitz, 2000; Israël et al., 2004b; Lappe, Jenkin, & Harris, 2007; May & Klatzky, 2000; M.-L. Mittelstaedt & Mittelstaedt, 2001; Redlick, Jenkin, & Harris, 2001; Schwartz, 1999; Sun, Campos, Young, Chan, & Ellard, 2004). This regression to the mean in these tasks is dependent on the contextual range of distances used (Petzschner & Glasauer, 2011). Other researchers have also found that distance reproduction varies depending on the gait type and speed of the outbound and response paths (Abdolvahab, Carello, Pinto, Turvey, & Frank, 2015; Chrastil & Warren, 2014a; M.-L. Mittelstaedt & Mittelstaedt, 2001; Turvey et al., 2009). Angle reproduction tasks have similarly shown underestimations of large angles and overestimation of small angles depending on the range (Becker, Jürgens, & Boss, 2000; Israël, Bronstein, Kanayama, Faldon, & Gresty, 1996; Ivanenko, Grasso, Israël, & Berthoz, 1997; Jurgens et al., 2003; Klatzky, Loomis, & Golledge, 1997; Marlinsky, 1999; Siegler, 2000; Siegler, Viaud-

Delmon, Israël, & Berthoz, 2000; Vidal & Bülthoff, 2010). In addition, angle reproduction is influenced by memory and the reference frame used during the task (Arthur, Philbeck, Kleene, & Chichka, 2012), further suggesting that the encoding function is not context-free.

Beyond possible violations of the model's assumptions, direct tests of the encoding-error model have been conducted, with mixed results. Péruch et al. (Peruch et al., 1997) found that the encoding-error model accounted for 89% of the variance in distance encoding and 93% of the variance in angle encoding, supporting the encoding-error account of path integration. May & Klatzky (May & Klatzky, 2000) also fit their data to the encoding-error model with a high correlation between expected and observed errors. Corollary assumptions that emerge from the model—that variability in each trajectory does not affect the others, no alignment of the paths into a common reference frame—have also been supported (Klatzky et al., 1999). On the other hand, Riecke et al. (Riecke et al., 2002) point out that participants in the Péruch et al. experiment undershot simple 180° turns by 16%, in violation of assumption 4. Riecke et al.'s application of the encoding-error model yielded results that violated axioms of trigonometry (assumption 1), such as negative values for encoded distances and angles. Furthermore, although participants in that study indicated they knew that both outbound legs of the triangle were equal, five of the twenty participants had mean final turns that could never complete any isosceles triangle, in violation of either assumption 1 or assumption 4. In addition, when applied to outbound paths of more than two legs, the authors of the encoding-error model themselves (Fujita et al., 1993) found that this model was not sufficient to explain the systematic errors.

Finally, the use of reproduction tasks to generate the encoding functions in path integration is also problematic. Reproduction tasks might confound encoding and execution errors: it is not possible to determine whether an undershoot in distance reproduction stems from

under-encoding the distance with accurate execution, accurate encoding with errors in distance execution, or some combination of the two. We previously found that magnitude of the response angle, not the magnitude of the encoded angle, predicted errors (Chrastil & Warren, 2017). This finding indicates that execution error—not encoding error—could make the largest contribution to systematic path integration errors. We also demonstrated that both encoding and execution errors contribute to total error in a distance reproduction task, and we provided a quantitative estimate of both types of errors (Chrastil & Warren, 2014a). Participants tended to be more accurate and less variable when the outbound and response modes matched (e.g. walking on both the outbound and response paths vs. walking out and then throwing a beanbag the same distance). These results suggest that the most accurate reproduction tasks are based on matching idiothetic information, rather than some extrinsic distance metric. Hence, reproduction tasks may be highly calibrated for accurate reproduction, so they might not reveal the true encoding function. Indeed, it is possible that encoding during triangle completion is fundamentally different from encoding during distance reproduction; in a reproduction task, the encoded information is matched during the response, whereas in triangle completion the encoded information is integrated together to generate a completely novel path.

In sum, the encoding-error model has received support for some of its assumptions and direct tests have found that the model accounts for a large portion of the variance. On the other hand, several of its assumptions do not hold up to scrutiny. Both the assumption of no execution error and the assumption of a consistent encoding function have been called into question. Even the assumption of Euclidean geometry may not hold in general (e.g. Warren, Rothman, Schnapp, & Ericson, 2017). We will now consider other models of path integration.

Other Path Integration Models

Klatzky et al. (1999) point out that the encoding-error model is a *configural model*, in which the entire outbound journey is stored in memory, and the entire configuration is used when the animal wishes to return home. In contrast, other models of path integration (Fujita et al., 1990; Merkle, Rost, & Alt, 2006b; Muller & Wehner, 1988) are moment-by-moment *homing vector models,* in which the animal continuously updates a vector back to its home location. Homing vector models are history-free, such that the animal could not return to any other location on the outbound path. Philbeck et al. (Philbeck, Klatzky, Behrmann, Loomis, & Goodridge, 2001) provide support for a homing vector model of path integration by demonstrating the importance of the origin for path integration. Participants received a brief view of the path layout at the start of each trial, and then walked without vision for the rest of the path. They were much more accurate at returning to the origin than to a rotationally equivalent position they had not previously visited. In a human neuroimaging study, several brain regions demonstrated increasing activation corresponding to Euclidean distance from the start location during movement in a circular trajectory, consistent with a homing vector model of path integration (Chrastil, Sherrill, Hasselmo, & Stern, 2015). Other recent evidence indicates that a homing vector could have separate position and heading estimates, with the implication that this homing vector could have an allocentric reference frame (Mou & Zhang, 2014; Zhang & Mou, 2017).

Evidence against a homing vector model of human path integration comes from findings that error and time to initiate the homeward trajectory both increase with the increasing complexity of the outbound path (Klatzky et al., 1990; Loomis et al., 1993; Wan et al., 2013) and with changes in the configuration (May & Klatzky, 2000). A homing vector model should not be affected by increased complexity of the outbound path because only the vector back to the start

location is stored in memory, whereas a configural model would be so affected. In contrast, Wiener & Mallot (Wiener & Mallot, 2006) found no effect of path complexity on errors when the length of the outbound path was controlled. Similarly, Yamamoto et al. (Yamamoto, Meléndez, & Menzies, 2014) found that errors during blindfolded walking path integration were related to the outbound path length and turns, not to additional complexity in the paths. To potentially explain these conflicting results, recent research has demonstrated that humans are capable of both homing vector and configural strategies (He & McNamara, 2018; Wiener, Berthoz, & Wolbers, 2011).

Other models of path integration have focused on the integration component using different reference frames and coordinate systems (Benhamou, Sauve, & Bovet, 1990; Gallistel, 1990; Merkle et al., 2006b; H. Mittelstaedt & Mittelstaedt, 1982; Muller & Wehner, 1988; Wehner & Wehner, 1986) (see Benhamou & Séguinot, 1995; Maurer & Séguinot, 1995; Merkle et al., 2006 for reviews of path integration models). Many of these models have developed accurate and normative accounts of integration (e.g. Gallistel, 1990; Jander, 1957 referenced in Benhamou & Séguinot, 1995; Mittelstaedt & Mittelstaedt, 1982) as a tool for determining what information an animal must have in principle and how it must use that information, rather than explaining systematic errors (Maurer & Séguinot, 1995). These models are primarily concerned with describing the mathematical relationship between the encoded outbound path and the required response path back to the home location. A recent model of path integration in humans (Harootonian et al., in press) tested a configural model that used vector addition for the integration process. This model operates under the assumption that people over- or under-weight legs of the triangle due to uneven integration over time. The model fit triangle completion data better than the encoding-error model, but still focuses primarily on the encoding component.

More descriptive models have incorporated errors by introducing random noise (Benhamou et al., 1990) or a correction factor based on empirical data (Muller & Wehner, 1988). Other models have attempted to find ways to incorporate some of the systematic errors seen in distance reproduction (Merkle, Rost, & Alt, 2006a; Sommer & Wehner, 2004). In many ways, these models of systematic underestimations resemble the encoding-error model. For example, the systematic errors produced in the Benhamou model (Benhamou et al., 1990) stem not from the integration or execution of the homebound trajectory, but from errors in the estimations of the outbound distances and angles. Although Benhamou et al. (1990) introduced stochastic error and Fujita et al. (1993) used empirically derived error, they both assumed that once the initial error is introduced during encoding, no further error accrues during integration or execution of the homeward trajectory. However, a "leaky integrator" model, whereby the animal gradually "forgets" sections of the distance or angle traveled (Lappe & Frenz, 2009; Lappe et al., 2007), incorporates leakage during the execution phase of a distance task as well as encoding. In this model, the leaky integrator counts up during encoding and counts down during execution, which could provide a route to understanding how encoding and execution work together. We similarly aimed to incorporate error in multiple aspects of path integration in the present study.

Experimental Aims and Overview

The aims of the present study are to 1) directly test the encoding-error model of path integration and 2) to compare this model with alternatives that incorporate other types of error. While particular aspects of the encoding-error model have been tested previously, no study has directly tested the basic approach. The present study first uses distance and angle reproduction tasks to predict encoding errors in path integration, as in Fujita et al. (1993). Linearly combining these errors yields a response region in which errors are expected to lie if encoding is the only

source of systematic error. Errors that lie outside of this region can then be attributed to integration or execution errors.

For example, a person might encode Legs X and Y and the interior turn angle α (Figure 1a) as X', Y', and α', respectively (Figure 1b). The accurate return path for the encoded triangle is depicted in Figure 1b by the turn β and Leg Z. The encoding-error model predicts that the navigator turns through the angle β and walks the distance of Leg Z on the actual triangle (Figure 1a), without any integration or execution error. Note that β and Leg z are the same magnitudes in both figures, indicating the same execution. Thus, error-prone encoding of the outbound path followed by accurate integration and execution of the return path yields systematic and predictable errors (ellipse in Figure 1a). If the navigator makes systematic errors other than those predicted by the encoding-error model, then those errors must be attributed to the integration or execution components of path integration.

All experiments in the current study took place in an immersive virtual environment. In five sessions, participants performed the triangle completion task, followed by distance and angle reproduction tasks to generate empirical data for simulations. In the simulations, we sampled from the distance and angle reproduction errors using Monte Carlo methods, and linearly combined them using the law of cosines[1] to predict final homing positions. The final errors were then compared to the actual triangle completion data. We simulated both average data, as Fujita et al. (1993) did, but also used each individual person's encoding functions from reproduction tasks to model their own triangle completion data. Finally, we derived alternative models of path

[1] The law of cosines can be used to find missing elements of any triangle. Namely, if the first leg has length x, the second has length y, and the angle between those legs is α, then the length of the third leg, z, is given by $z = (x^2 + y^2 - 2xy*\cos\alpha)^{1/2}$. The turn angle β (between legs y and z) is then specified by $\beta = \cos^{-1}((y^2 + z^2 - x^2)/(2yz))$. In these simulations, z and β are the path length and turn angle, respectively, for the homebound trajectory.

integration that included execution errors.

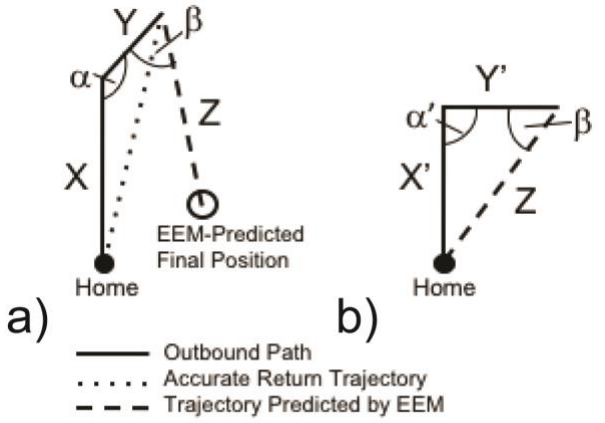

Figure 1. Predictions of the encoding-error model of path integration. **(A)** Actual walked triangle with outbound legs X and Y and turn angle α. **(B)** The encoded triangle, indicated by X', Y', and α'. If a navigator walks the homeward trajectory for the encoded triangle **(B)** on the actual triangle **(A)** without any integration or execution error, as predicted by the encoding-error model, systematic and predictable errors should be produced, as indicated by the ellipse in **(A)**. Note that the magnitude of β and the lengths of Leg Z are the same in both **(A)** and **(B)**, indicating the same execution.

Note that the triangle completion task was performed in both a hedge corridor and in an open field scene (Figure 2a). Many previous triangle completion studies have been conducted in an open environment or blindfolded, so we also sought to test whether performance in a hallway setting would generalize to an open environment.

Methods

Participants

Seven females and eight males participated in this study. Most were undergraduate or graduate students at Brown University, and all were paid for their time at the rate of $8/hour. One female and one male withdrew due to symptoms of simulator sickness. Ages of the remaining 13 participants ranged between 19-30, with a mean age of 25.73. All participants signed forms indicating their informed consent to be a part of the study in fulfillment of the requirements of the Brown University IRB.

Because we conducted several different types of tasks, sample size was determined based on previous experiments in our lab and other path integration studies (e.g. Klatzky et al., 1998; Loomis et al., 1993) that have had large effect sizes with similar sample sizes. Looking across all the analyses using a post-hoc power determination (Faul, Erdfelder, Lang, & Buchner, 2007) shows that our experiments had sufficient power to detect effects. For example, for a repeated-measures ANOVA with 13 participants and 5 measurements (as in the distance reproduction task), $\alpha = 0.05$, a conservative correlation between with-subject measures of 0.5, the maximum nonsphericity correction of 0.25, and the actual effect size $\eta_p^2 = 0.781$, the power to detect an effect is 1.000. All significant and marginal main effects and interactions from the ANOVAs, as well as all the correlation analyses, were found to have post-hoc power ranging from 0.548 to 1.000, with most of the effects on the higher end of the range. Any increase in correlation between measures increases the power substantially, which is a reasonable assumption given our within-subjects measures of the same types of tasks. For example, our lowest power of 0.548 becomes 0.906 if the correlation assumption is increased to 0.8 (while maintaining maximum nonsphericity). The power for angular measures is more difficult to determine with the circular statistical analysis, however, they had similar patterns as the distance measures and so would likely fall into the same range of power. We note that post-hoc power analyses are largely redundant with the outcomes of the data, therefore, we caution readers about the relatively small sample size of the study. There is a possibility that the study is underpowered, however, we note that most of the within-subjects effects are quite large. The subsequent modeling work uses participants' data to model their own outcomes, which mitigates some of the issues with smaller samples.

Equipment

The experiment took place in the VENLab, a 12 meter x 12 meter room using virtual displays. Images were presented to the participants using a Cybermind Visette 2 head-mounted display (HMD) with a 60^{o} horizontal x 46.8^{o} vertical field of view and resolution of 640 x 480 pixels. Participant movement was tracked using an InterSense IS900 tracking system with a 70ms latency. Participants made responses with a USB radio mouse. Images were generated on a graphics PC (Alienware, NVIDIA Quadro FX 3000 graphics card) using Vizard (WorldViz) to render the images. Cricket sounds were presented to the participants over headphones to create naturalistic noise to prevent participants from receiving information about their location or orientation in the room from auditory cues.

Environment

The environment consisted of hallways made of 3-meter high walls with a hedge texture, an opening for the blue sky above, and a gravel path below (Figure 2a, top). The corridor was 100-meters long, so as to give no noticeable change in visual angle at end of the hall when participants walked. The hallway was used for the straight-line translation segments of triangle completion, distance reproduction, and angle reproduction. During the segments of each task that required rotations in place, the hallway was replaced by a cylindrical hedge with a 1-meter radius surrounding the participant. This cylinder provided texture cues from optic flow while rotating without any additional landmark information about how far they had rotated. For triangle completion, we also conducted a condition in an open field environment. In the open field, the environment consisted of the textured ground plane and a green pole 100 meters away (Figure 2a, bottom). The ground surface was a gravel texture, and there were no other visual landmarks. The pole was designed to give participants some orientation so they could walk in a straight line but not provide any additional information from the change of size of the pole as they moved.

For all tasks, the location and orientation of the next trial was marked by an orange pole with arrows on ('start pole').

Experimental Tasks

Triangle completion. Triangle completion consisted of walking one outbound leg, turning, and then walking along a second outbound leg. At the end of the outbound path, participants were instructed to turn to face their starting location and click the mouse once they were facing the start location. They were then instructed to walk forward toward the starting location and to stop and click the mouse again when they reached that location. Two conditions of triangle completion were presented to participants: i) hallway and ii) open field. In the hallway condition, participants walked forward along the hallway and stopped when they heard a chime; the hallway then disappeared and was replaced by the cylindrical hedge. The participant then turned to the right or the left as directed by an auditory cue and stopped when they heard the chime again. Then the cylinder hedge disappeared and a new hallway opened up. The participant walked along this hallway until the chime sounded again and the cylindrical hedge appeared. The participant was instructed to turn to face the starting location and then click the mouse, whereupon a third hallway appeared, in the same orientation as the participant was facing when they clicked the mouse. The participant was then instructed to walk forward until they thought they reached the start location, then click the mouse again. All hedges then disappeared and the participant walked to the start pole for the next trial.

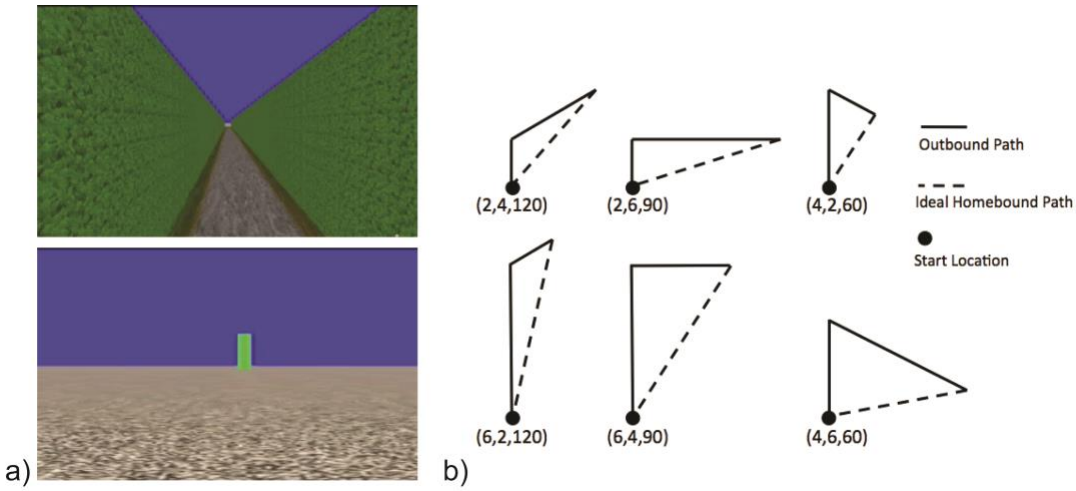

Figure 2. Virtual environments. **(A)** Views of the Hallway and Open Field conditions of triangle completion. The hallway was also used for distance and angle reproduction. **(B)** Triangle types used in the triangle completion experiments. In parentheses are the names of the triangle types, with the first term specifying the length of the first leg in meters, the second term indicating the length of the second leg in meters, and the third term designating the interior turn angle between the first and second legs in degrees.

The open field condition was included in order to compare the results of the hallway experiments to those of other studies. Instead of hallways, the participant walked forward towards a green pole 100 meters away (Figure 2a, bottom). When the chime sounded, the pole disappeared and the participant turned until they heard the chime again. Then, a second green pole appeared at the new orientation, and the participant walked forward towards this pole until the chime sounded again. Finally, the pole disappeared and the participants turned to face the start location, clicked the mouse, and walked forward in the open field until they reached the start location and then clicked the mouse again. The open field condition differed from the hallway primarily in amount of visual texture.

For both the open field and hallway conditions, as well as the two reproduction tasks, the chime sounded at a threshold 0.5 meters before the length of each leg and within ± 3° of the turn angle. The position of the chime threshold gave participants the chance to take a step after it sounded, while preventing them from overshooting the distance or angle and having to backtrack.

Triangle completion consisted of 48 trials, with an additional six practice trials at the beginning of the session. None of the triangles in the practice were part of the experimental block, although they were of similar scale. The lengths of the legs of the experimental triangles were 2, 4, or 6 meters, and the interior angle between legs 1 and 2 was 60°, 90°, or 120°; a subset of six triangles were tested out of the potential 27 combinations (Figure 2b). These triangle sizes were determined *a priori* both by the constraints of the research space and by the desire to have a variety of leg lengths and interior angles. Interior angles included right, obtuse, and acute angles, and sometimes the first leg was shorter than the second leg and vice versa. Both right-handed and left-handed versions of these six triangles were included; trials alternated between right and left triangles. The order of presentation was otherwise randomized for each participant. Each triangle was presented to a participant 8 times, four trials with left turns and four trials with right turns. There were 24 starting locations and four starting orientations in the room to prevent participants from receiving feedback on their performance and from using a constant frame of reference. These measures increased the likelihood that participants would treat each triangle as separate from the others.

Distance reproduction. Participants walked down the virtual hallway until they heard a chime, at which point the hallway disappeared and the cylindrical hedge surrounded the participant. Participants then turned 90º to the right or left as directed by an auditory cue until they heard the chime again. When the chime sounded again, the cylindrical hedge disappeared and a new hallway appeared. Participants were then instructed to walk forward in this new hallway for the same distance they had walked on the original path, and then click the mouse.

In order to prevent participants from counting their steps, they performed an additional interference task. At the start of each trial, participants were given a seed number between 100-

500 over the headphones, chosen randomly from a random subset of 80 numbers. Participants then had to count aloud backward by threes from this number until they clicked the mouse. Two participants who were not native English speakers were allowed to count in their native language.

Distance reproduction consisted of 60 trials, with four additional practice trials at the beginning of the session. None of the distances in the practice trials were part of the experimental block. The magnitudes of the distances were the same as those of the hallways used in triangle completion plus two additional distances of 8 and 10 meters (2, 4, 6, 8, and 10 m). These magnitudes also match those we have used in previous work (Chrastil & Warren, 2014). Both right turns and left turns between the outbound and reproduced leg were included; trials alternated between right and left turns. The order of presentation was otherwise randomized for each participant. Each distance was presented to participants 12 times, with six left and six turns. There were 4 starting locations in the room, each with a different starting orientation, to prevent participants from receiving feedback on their performance.

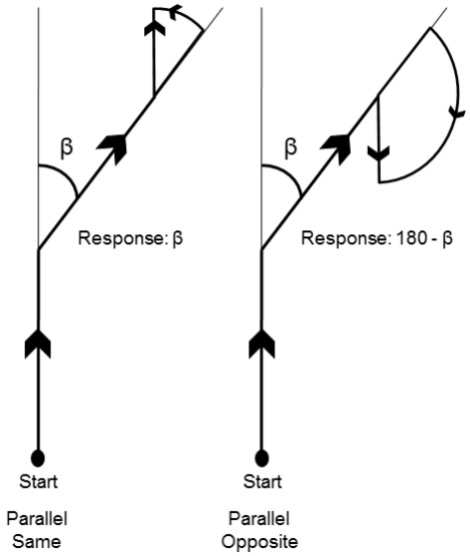

Figure 3. Left: The Same condition for angle reproduction. Participants walked in the direction of the arrows turning right, with the initial turn angle β. The required response is a left turn in the same magnitude of β. Right: The Opposite condition for angle reproduction. Participants walked in the

direction of arrows turning right, with the initial turn angle β. The required response is a right turn with a magnitude of 180-β (see Chrastil & Warren, 2017, for details).

Angle reproduction. The angle reproduction task has previously been described and the results have been previously reported in (Chrastil & Warren, 2017). Here, we only present the methods and results of the basic reproduction task and its relationship to the modeling work. Participants walked in the hallway until they heard the chime, which sounded after 5.86 meters. The participant was then surrounded by the cylindrical hedge, and turned right or left as directed by an auditory cue until they heard the chime again. A new hallway at the specified angle then replaced the cylindrical hedge. Participants walked down this hallway until they heard the chime again, after another 3.5 meters, and the cylindrical hedge appeared. In the Same condition, participants were then instructed to turn back to face in the same direction they had originally walked, but parallel to the original path (functionally equivalent to reproducing the first angle), then click the mouse, whereupon the cylindrical hedge disappeared and a new hallway opened up in the same orientation as the participant. Participants walked forward on this new path for 1.5 meters, as which point the trial ended and they were instructed to walk to the start location for the next trial.

In the Opposite condition, the procedure was similar except that participants were instructed to turn to face in the opposite direction, parallel to the original path (see Chrastil & Warren, 2017). This required them to turn through the supplement of the first angle: a parallel path in the opposite direction for a 30° right turn can be found by turning 150° to the right. These two versions of angle reproduction were crafted to probe the potential execution error in turning. If execution error increases as the turn angle increases, participants would show a greater error for large angles than for small angles, even if they encoded the turn angle accurately. These factors are confounded in the Same condition because smaller response turns are required by

smaller outbound turns and larger response turns are required by larger outbound turns. By including the Opposite condition—where participants make a smaller response for larger outbound turn angles—execution error can be compared between the two conditions.

Angle reproduction consisted of 60 trials, with four additional practice trials at the beginning of the session. None of the angles in the practice trials were part of the experimental block. Magnitudes of the turn angle were the same as those used in triangle completion plus two additional angles of 30° and 150° (30°, 60°, 90°, 120°, 150°; see Chrastil & Warren, 2017). Both right turns and left turns were included; trials alternated between right and left turns. The order of presentation was otherwise randomized for each participant. Each turn angle was presented to participants 12 times: six left and six right turns. There were 3 starting locations in the room, each with a different starting orientation, to prevent participants from receiving feedback on their performance and from using a constant frame of reference in the room.

Procedure

After informed consent was obtained, the inter-ocular distance for each participant was measured and entered into the graphics card, then the HMD was placed and adjusted on the participant's head. The participant also wore a backpack containing some cables, which weighed approximately three pounds and did not impede movement. To prevent participants from tripping over the cable connecting the HMD to the control box, an experimenter (the 'wrangler') continuously followed the participant keeping the cable out of the way at all times. A test environment and several practice trials for each task served as immersion time (5-10 minutes) for the virtual environments. Instructions for each task were presented over headphones in the HMD, which guided the participants through practice trials. The instructions for each task were then repeated before the start of the experiment. Experimental trials were presented in one block, with

frequent opportunity for breaks. The tasks took between 40-60 minutes to complete. Three individual sessions were stopped after 60 minutes and completed in a separate session due to time constraints. At the end of the second through fifth session, participants filled out a brief questionnaire asking for strategies used in the task and (if applicable) if one version of the task seemed easier than the other.

Participants completed five sessions for the experimental tasks, which were run in a semi-counterbalanced order. The first two sessions consisted of triangle completion, with one session of the open field condition and one session of the hallway condition counterbalanced across subjects. Sessions three, four, and five consisted of the two versions of angle reproduction (Same and Opposite) and one distance reproduction task. The order of these three tasks was counterbalanced across participants. The triangle completion sessions were performed first to prevent contamination from the angle and distance reproduction tasks influencing performance on triangle completion. Sessions were completed over the course of two to six weeks for each participant, with a break of at least 4 hours between sessions. Each participant generally came in for a session every four days.

Analysis

Analysis was conducted using JMP software (SAS) and SPSS for linear measures and custom Matlab (Mathworks) scripts for angular measures. For triangle completion, each participant's path on the homebound leg was evaluated in relation to an ideal path between the position where they clicked the mouse to begin the homeward path ('click location') and the actual start location. Three response measures were quantified from this path (Figure 4): a) *Position error* was calculated as the absolute distance between the final position of the participant and the start location. *Path length* is the distance traveled by the participant on the

homebound leg, with b) *Path length error* being the difference between the observed and ideal path length. A positive error indicates that the participant traveled too far, while a negative error indicates that the participant undershot the distance. c) *Initial turn angle error* is the difference between the participant's heading direction at the click location and the ideal heading direction from that point to the home location. These errors are positive if the participant overturned compared to the correct path and negative if the participant undershot the correct turn.

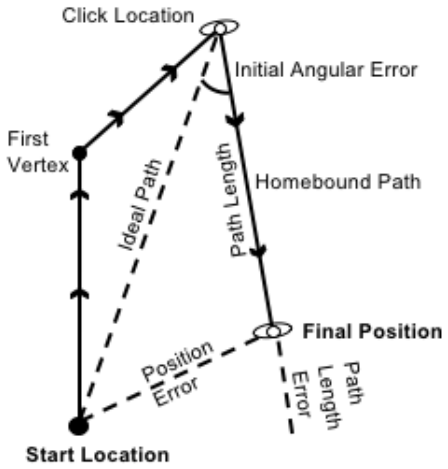

Figure 4. Response measures for triangle completion. See text for details.

Analyses of variance (ANOVAs) were performed on the means of all the response measures, as well as on their standard deviations (for linear measures) or angular deviations (for angular measures). For all comparisons, right and left turns were collapsed across trials. The linear measures of position error and path length error were analyzed with repeated-measures ANOVAs using a 6 (triangle type) X 2 (hallway or open field condition) design. For these measures, we report F and p values and indications of effect size (η_p^2, partial eta squared). The angular measure of initial turn angle error was analyzed using a multiple-sample Watson-Williams one-way test for circular data (Batschelet, 1981). Currently there are no higher-order ANOVAs or computations of effects sizes available for circular data. Thus, we report F and p

values for each comparison, and all two-way interactions were tested as Bonferroni-corrected separate one-way effects. Standard and angular deviations for all these measures were also analyzed. Linear regressions and correlation coefficients were also found for these measures to determine if there is a relationship between the initial distance/angle and the reproduced distance/angle or their errors.

For distance and angle reproduction, the only response measures were path length error and turn angle error, respectively. Distance error was calculated as the difference between the observed distance walked and the ideal response. A positive error indicated an overshoot of the distance while a negative error indicated an undershoot. Angle reproduction error was calculated as the difference between the direction the participant was facing after they made their response and the ideal heading direction. As distance error is a linear measure, a repeated-measures ANOVA using the five distances reproduced in this task was performed. Angular error was analyzed using a multi-sample Watson-Williams one-way ANOVA, collapsing across left and right turns. In addition to testing the factor of angle, the factor of condition (walk parallel in the same or opposite direction) was tested. Standard and angular deviations for all these measures were also analyzed. Linear regressions and correlation coefficients were also computed for path length and turn angle to determine whether there is a relationship between the ideal and actual responses.

Trials were excluded from analysis in cases of a software crash, a loss of tracking in the virtual environment, or the participant indicated that they had terminated the trial too early. In addition, trials from the distance and angle reproduction tasks were excluded if the participant walked too close to the physical wall of the room during their response. Portions of triangle completion data were also excluded for this reason, however, initial information such as initial

angle error was included in the analyses because this measure could be collected before the participant went out of bounds. In addition, if during triangle completion the participant drifted from the lines and was more than 1.5m from the actual location of the second vertex of the triangle (click location) when they clicked to return to home, only initial angle error were collected. This drift was more common in the Open Field condition due to the lack of hallways to keep people on course. In all, between 0.38% and 2.05% of the reproduction trials were excluded. 1.36% of the triangle completion trials were completely excluded, while 9.38% of triangle completion trials contained only initial angle information.

Results

Triangle Completion

Overall triangle completion performance is shown in Figure 5.

Linear measures. A 6 (triangle type) X 2 (hallway or open field condition) repeated-measures ANOVA was first conducted on the mean position errors. Mean absolute position errors showed a significant main effect of triangle type ($F_{5,60} = 2.607$, $p = 0.034$, $\eta_p^2 = 0.178$), suggesting that some triangles were more difficult for participants to return to the start location (Figure 6). There was no effect of open field/hall ($F_{1,12} = 0.005$, $p = 0.946$, $\eta_p^2 = 0.000$) and no interaction ($F_{5,60} = 1.465$, $p = 0.215$, $\eta_p^2 = 0.109$) for position error. A 6 x 2 ANOVA on the standard deviation of position error only showed a marginal effect of triangle type ($F_{5,60} = 2.048$, $p = 0.085$, $\eta_p^2 = 0.146$), and there was no main effect of open field/hall ($F_{1,12} = 0.974$, $p = 0.343$, $\eta_p^2 = 0.075$) and no interaction ($F_{5,60} = 0.941$, $p = 0.461$, $\eta_p^2 = 0.073$).

For path length, overall participants showed a compression of distance responses, as seen in previous work. For homeward trajectories that called for a shorter path length, participants generally walked too far, while they did not walk far enough for longer ideal paths (Figure 7a).

Path length did increase with the ideal path length (Hallway: $y = 0.391x + 2.761$, $r = 0.915$; Open Field: $y = 0.431x + 2.399$, $r = 0.868$). Path length errors differed between triangle types (Figure 7b). A 6 (triangle type) X 2 (hallway or open field condition) repeated-measures ANOVA found a significant effect of triangle type on path length errors ($F_{5,60} = 38.576$, $p < 0.001$, $\eta_p^2 = 0.763$). There was no main effect of open field/hall condition ($F_{1,12} = 2.683$, $p = 0.127$, $\eta_p^2 = 0.183$), and there was only a marginal interaction ($F_{5,60} = 2.103$, $p = 0.077$, $\eta_p^2 = 0.149$). Analysis of the path length errors standard deviations yielded no significant main effect for open field/hall ($F_{1,12} = 0.004$, $p = 0.954$, $\eta_p^2 = 0.000$), although the triangle type X open field/hall interaction was marginal ($F_{5,60} = 2.164$, $p = 0.070$, $\eta_p^2 = 0.153$), as was the main effect of triangle type ($F_{5,60} = 1.978$, $p = 0.095$, $\eta_p^2 = 0.141$).

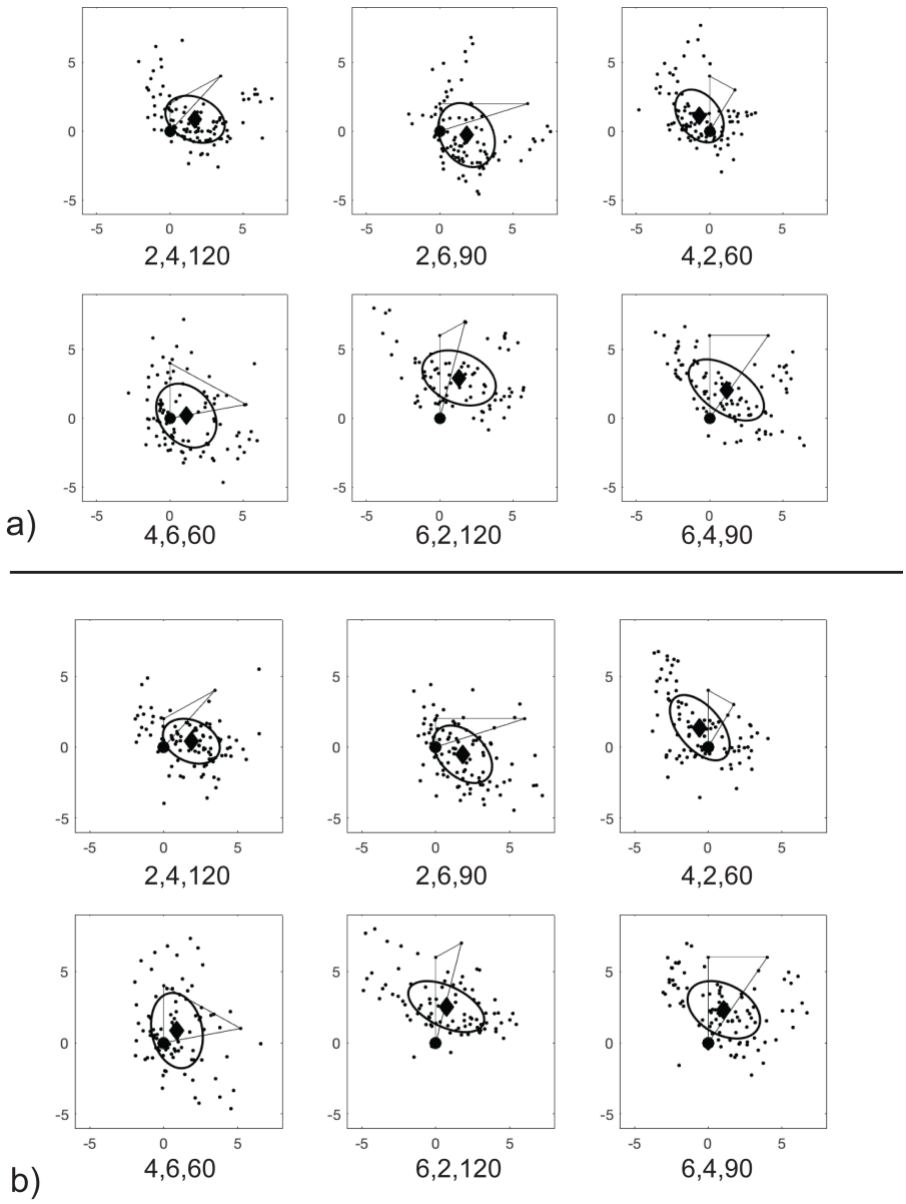

Figure 5. Overall pattern of results from triangle completion. **(A)** Open field condition. **(B)** Hallway condition. Individual dots represent individual trials by the 13 participants. The filled circle at the origin (0,0) is the start location. The filled diamond is the mean final location averaged over all the participants. Ellipses indicate 95% confidence intervals for the simulation. Left turns have been reversed and combined with right turns.

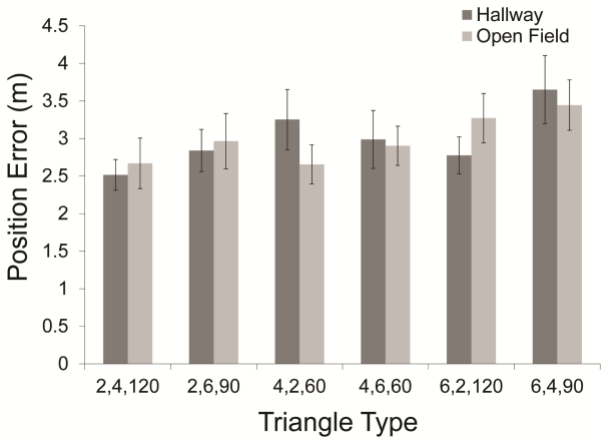

Figure 6. Position Error in the triangle completion tasks, measured as absolute distance between the final position and the starting location. Error bars indicate between-subjects standard error.

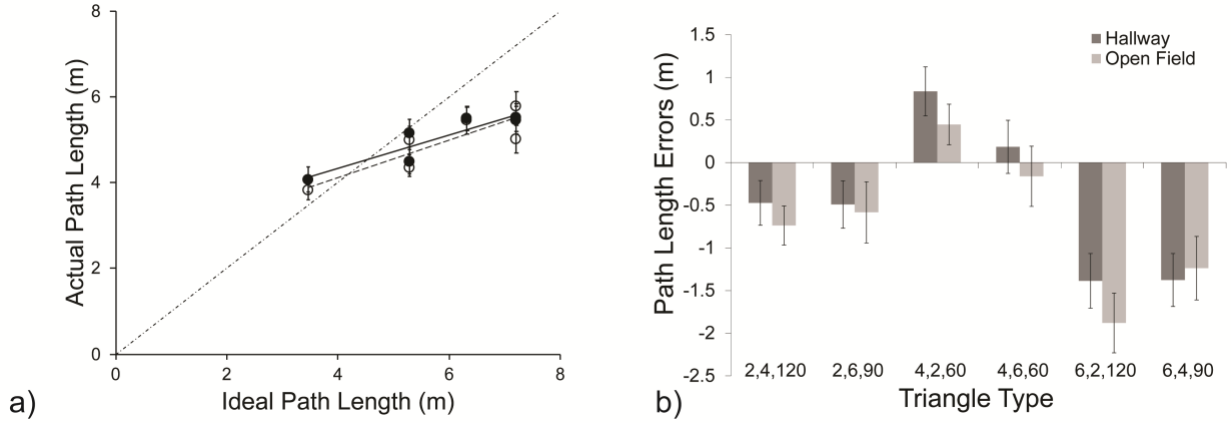

Figure 7. (A) Path length, compared to the ideal values for path length. Filled circles/solid line: Hallway condition. Open circles/dashed line: Open field condition. Diagonal line indicates correct performance. There were six triangle types; however, for two triangles the ideal path length turned out to be the same distance as two other triangles, resulting in only four ideal path length values on the x axis. **(B)** Path Length Errors in the triangle completion tasks. Negative errors indicate undershoots while positive errors indicate overshoots. Error bars indicate between-subjects standard error.

Angular measures. Large angular errors were observed relative to the magnitude of the correct angle. For both conditions, initial turn angle increased with the ideal turn angle, indicating that participants were able to able to discriminate between triangle types (Hallway: $y = 0.502x + 22.635$, $r = 0.964$; Open Field: $y = 0.609x + 19.931$, $r = 0.999$; Figure 8a). Turn angles were somewhat compressed, such that small turn angles were overestimated and larger turn

angles were underestimated. Turn angle errors revealed some slight differences between the open field and hallway conditions. In the hallway condition, a Watson-Williams test for circular data on the mean initial turn angle error ($F_{5,12} = 3.8426$, $p < 0.05$, $\eta_p^2 = 0.199$) showed significant effects of triangle type (Figure 8b). In the open field condition, on the other hand, there was no main effect of triangle type on mean initial turn angle error. However, an examination of individual pairwise comparisons between the open field and its matched hallway triangle type produced no significant differences.

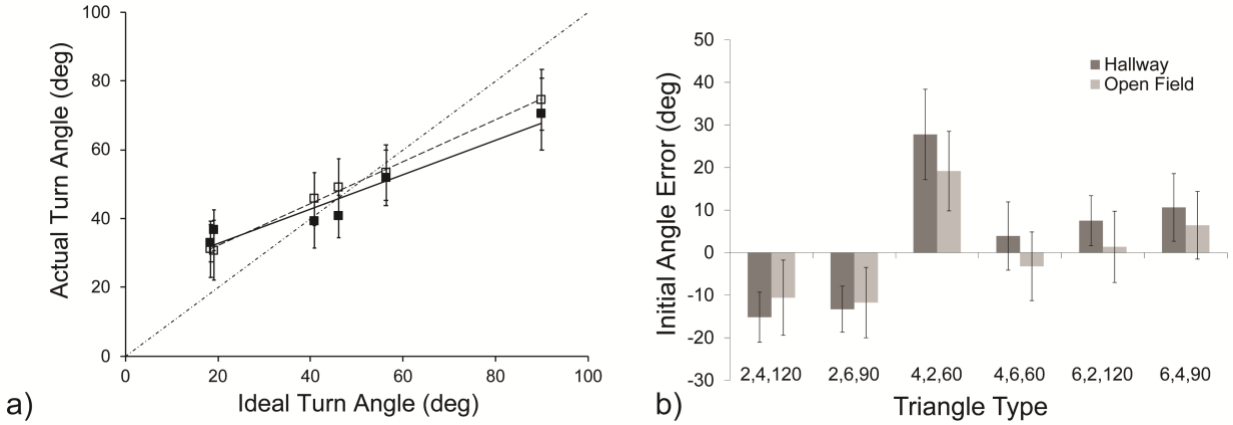

Figure 8. (A) Initial turn angle as a function of the ideal turn angle. Filled squares/solid line: Hallway condition. Open squares/dashed line: Open field condition. Diagonal line indicates correct performance. **(B)** Initial Turn Angle Errors. Negative values indicate an underturn, while positive values indicate an overturn. Error bars indicate between-subjects standard error.

Overall, the hallway condition had a slightly more divergent pattern of errors, leading to an overall main effect of triangle type, while the somewhat smaller range of errors in the open field condition was not strong enough to create a main effect. For any given triangle type, there was no difference between the open field and hallway condition, but when examining the overall range of errors, the hallway condition had reliably more pronounced errors. This result likely stems from the fact that participants were less able to distinguish between angles in the hallway condition, leading to more similar turn angles. The angle errors were consequently larger in the hallway condition because the participants were not making the full range of turns. However, this

effect of open field/hallway is fairly small, considering that none of the individual contrasts showed an effect. The angular deviations for initial turn error showed no main effects of triangle type for either the open field or the hallway condition.

Distance Reproduction

Mean reproduced distance as a function of initial distance appears in Figure 9a. Linear regression of reproduced distance on initial distance revealed that the slope was less than 1 with a positive intercept ($y = 0.75x + 1.28$, $r = 0.975$), meaning that participants overestimated shorter distances and underestimated longer distances. The distance reproduction regression equation is similar to the distance encoding function used by Fujita et al. (1993: $y = 0.60x + 1.20$), although with a somewhat steeper slope. There was a significant main effect of distance on reproduction errors ($F_{4,48} = 42.763$, $p < 0.0001$, $\eta_p^2 = 0.781$). Post-hoc pairwise comparisons revealed significantly different errors between the following pairs (all $p < 0.05$, Bonferroni corrected): 2m-10m, 4m-8m, 4m-10m, 6m-8m, 6m-10m, and 8m-10m. These results confirm that long distances tend to be undershot and short distances tend to be overshot. The standard deviations of distance errors also showed a main effect of distance ($F_{4,48} = 19.693$, $p < 0.0001$, $\eta_p^2 = 0.621$). Post-hoc tests showed that the 2m distance had significantly less variability than all other distances (all $p < 0.001$, Bonferroni corrected).

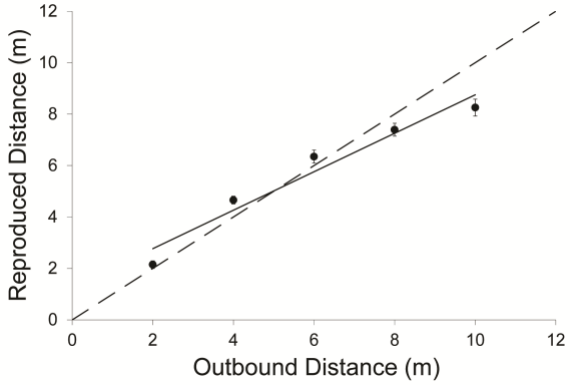

Figure 9. Responses in the Distance Reproduction Task compared to the actual distance walked. Dashed

line indicates veridical performance. Error bars indicate between-subjects standard error.

Angle reproduction. Results for angle reproduction have been reported elsewhere (Chrastil & Warren, 2017) and those previously-reported data are summarized here. Mean reproduced angle is plotted as a function of initial angle in Figure 10a. In the Same condition, the reproduced angle increased linearly with the initial angle ($y = 0.63x + 50.7$, $r = 0.999$), whereas in the Opposite condition the reproduced angle corresponded to the supplement of the initial angle, and so decreased ($y = -0.52x + 152.9$, $r = 0.993$). The regression equation is similar to the encoding function for angle used by Fujita et al. (1993: $y = 0.48x + 50$), although with a somewhat steeper slope, suggesting that this method of angle reproduction is comparable to theirs. Participants tended to underturn when a large response was required, and overturn when a small response was required. Responses of approximately 120°–135° were made relatively accurately.

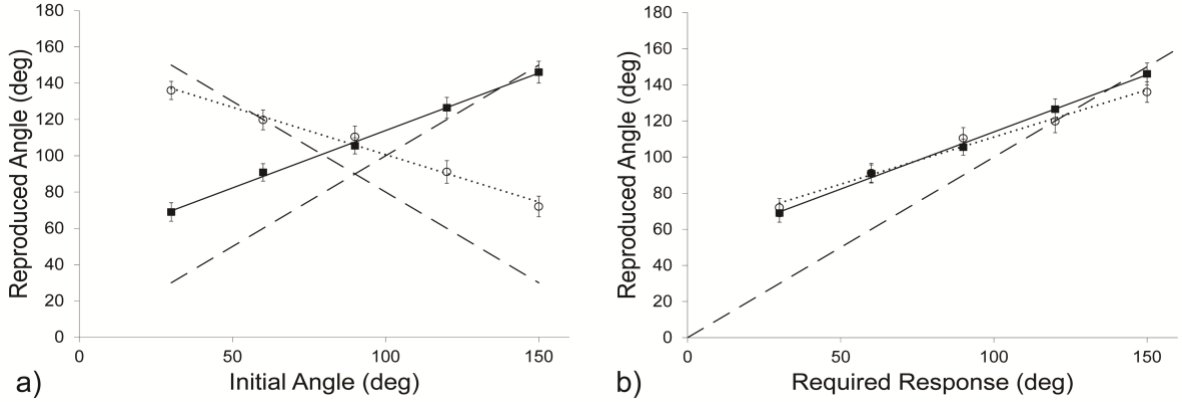

Figure 10. Results from the Angle Reproduction Task. Filled squares and solid lines indicate data from the Same condition. Open circles and dotted lines indicate data from the Opposite condition. The dashed lines show veridical performance. Error bars indicate between-subjects standard error. **(A)** Actual turns from the Angle Reproduction Task. **(B)** Reproduction as a function of the required response in angle reproduction tasks.

For mean angular errors, in both the Same and Opposite the Watson-Williams test found a main effect of turn angle (Same: $F_{4,12} = 7.3003$, $p < 0.01$, $\eta_p^2 = 0.319$; Opposite: $F_{4,12} = 5.8438$, $p < 0.01$, $\eta_p^2 = 0.266$). When reorganized by the required response angle (e.g. required response

of 30° for an initial angle Same 30° or for Opposite 150°), rather than the initial angle, the reproduction errors are quite similar (Figure 10b). Paired Watson-Williams tests showed no significant differences between these pairs, even before Bonferroni correction (see Chrastil & Warren, 2017, for more details). In other words, when responding to an initial turn of 30°, the participants made different errors in the Same condition (requiring a 30° turn) and in the Opposite condition (requiring a 150° turn). On the other hand, errors were equivalent when responding to a 30° turn angle in the Same condition and a 150° turn angle in the Opposite (both requiring a response turn of 30°). These results suggest substantial execution errors, such that people overshoot small turns and undershoot large turns. In our previous analysis of angular errors (Chrastil & Warren, 2017), we also found main effect of encoding angle and an angle x condition interaction, consistent with a contribution of encoding error. Overall, these findings suggest a large contribution of execution error, but that encoding error plays a role in angle reproduction as well.

Questionnaires

In triangle completion, participants reported attempting strategies such as counting steps and trying to envision the angle they turned through. Some attempted to measure the angle with their feet or arms. Participants were fairly evenly split as to which version of the task (Open field or Hallway) they felt was easier, but generally reported that the task was quite difficult. If anything, participants felt whichever version they completed second was easier, likely due to increased comfort with the task. Despite the counting distracter task in the distance reproduction task, participants reported attempting to count their steps. For the angle reproduction task, participants generally tried to "unturn" the angle they originally turned or gauged the angle relative to a reference angle of 90º, and some participants attempted to use their feet to measure

the angle. In all tasks, several participants said that they relied on intuition or visualization of the layout.

Simulations

Procedure

Triangle completion was simulated using Monte Carlo methods to sample from the distribution of errors from the distance and angle reproduction tasks. Simulations were designed to replicate the encoding-error model (Fujita et al., 1993). Triangle completion was simulated for each participant individually; each participant had an individual distribution for each distance and angle tested in the reproduction tasks, with their mean and standard/angular deviation taken from their individual data. Note that for angle encoding, we only used the Same condition, since this follows the encoding-error model's assumption that reproduction errors reflect only errors in encoding. However, as we will discuss below, the reproduction tasks also likely reflect a degree of execution error. This sampling method assumes a normal distribution. Each distance and angle were successfully fit to models of a normal or lognormal distributions when combining all of the trials for all participants, and so the assumption of a normal distribution was deemed justified.

For each iteration of the simulation, a value for the encoding distance of both legs and the turn angle was sampled randomly from a normal distribution, with the mean taken from the participant's mean reproduction data at that distance or angle, and the standard deviation taken from the participant's standard deviation at that distance or angle. For example, a triangle with an actual outbound path of a 4m leg, followed by a 60° interior turn, and a second leg of 6m might be sampled as 4.58m, 73.85°, and 5.92m, respectively. These values were then combined linearly using the law of cosines to compute the location where the simulated participant encoded the second vertex. Assuming no integration error and no execution error on the homeward

trajectory, as proposed by the encoding-error model, the model then used trigonometry to predict the location where the participant should walk when returning to the start location. The simulation was iterated 10,000 times for each of the six triangle types for each participant. The mean simulated final position was calculated, as well as the path length and turn angle for the homeward trajectory.

As Fujita et al. (1993) point out, the observed and model-predicted path length and turn angle will always be correlated to a certain extent, as a longer observed path will generally also have a longer model-predicted path. Path length error and turn angle error, however, will not necessarily be correlated and are thus more sensitive tests of the model's predictive power. Thus, the observed path length and turn angle errors were plotted against the models' path length and turn angle errors by combining the data across all triangle types for all participants (hallway condition only), yielding 78 points of comparison (6 triangle types x 13 participants). The slope and intercept of this relationship was computed and the effect size of the relationship was indicated by the r value; we report whether this relationship was significantly different from 0. Error ellipses were computed for a 95% confidence interval of the final simulated locations. In addition to simulating the encoding-error model, we simulated triangle completion using four other possible models (described below). We conducted a test on the r value of each alternative model's correlation, transformed to Fisher's z, compared with the baseline encoding-error model's r (transformed to z) value. Table 1 shows equations, r values, and significance levels for the encoding-error model and each of the four alternative models.

Results

The results for all five simulations are summarized in Table 1.

1. Encoding-error model. The simulations of triangle completion using the encoding-error

model yielded 10,000 estimated final positions for each participant for each triangle type. The mean estimated final position was then found for each set; the mean path length and turn angle errors were computed for that location. Figures 11 and 12 provide examples of two representative participants in the study. Part a) of these figures show the results for the encoding-error model simulations. Participant 11 (Figure 11) was not very accurate at triangle completion and was highly variable in both triangle completion and the reproduction tasks, with large error ellipses representing the final position. It should be noted, however, that even with the large ellipses for the simulated final positions for Participant 11 in panel a), the actual data collected from that participant generally still laid far outside of that ellipse. Participant 10 (Figure 12) was much more accurate and precise for all tasks, with a very small cluster of simulated final positions generally quite close to the participant's final position. However, several of the empirically collected data points for Participant 10 also laid outside of the simulation position (Figure 12a).

Figure 13a summarizes all the simulations, plotting model-predicted errors against actual errors for each participant. The relationship between actual and predicted path length errors is described by the equation $y = .45 x + 0.024$, ($r(76) = 0.4818$, $p < 0.001$), while for turn angle errors the relationship is described by $y = -.14x – 8.1$ ($r(76) = -0.2282$, $p = 0.044$). Note that the correlation coefficient for turn angle errors is negative, indicating that the model negatively predicted actual errors. In general, the model predicted path lengths that were longer than those taken by the participants, and compressed the range of turn angles compared to the range taken by participants. Although the model was close for participants who were more accurate or who had less variability in their responses, it proved less successful at predicting the responses of participants with high variability or low accuracy.

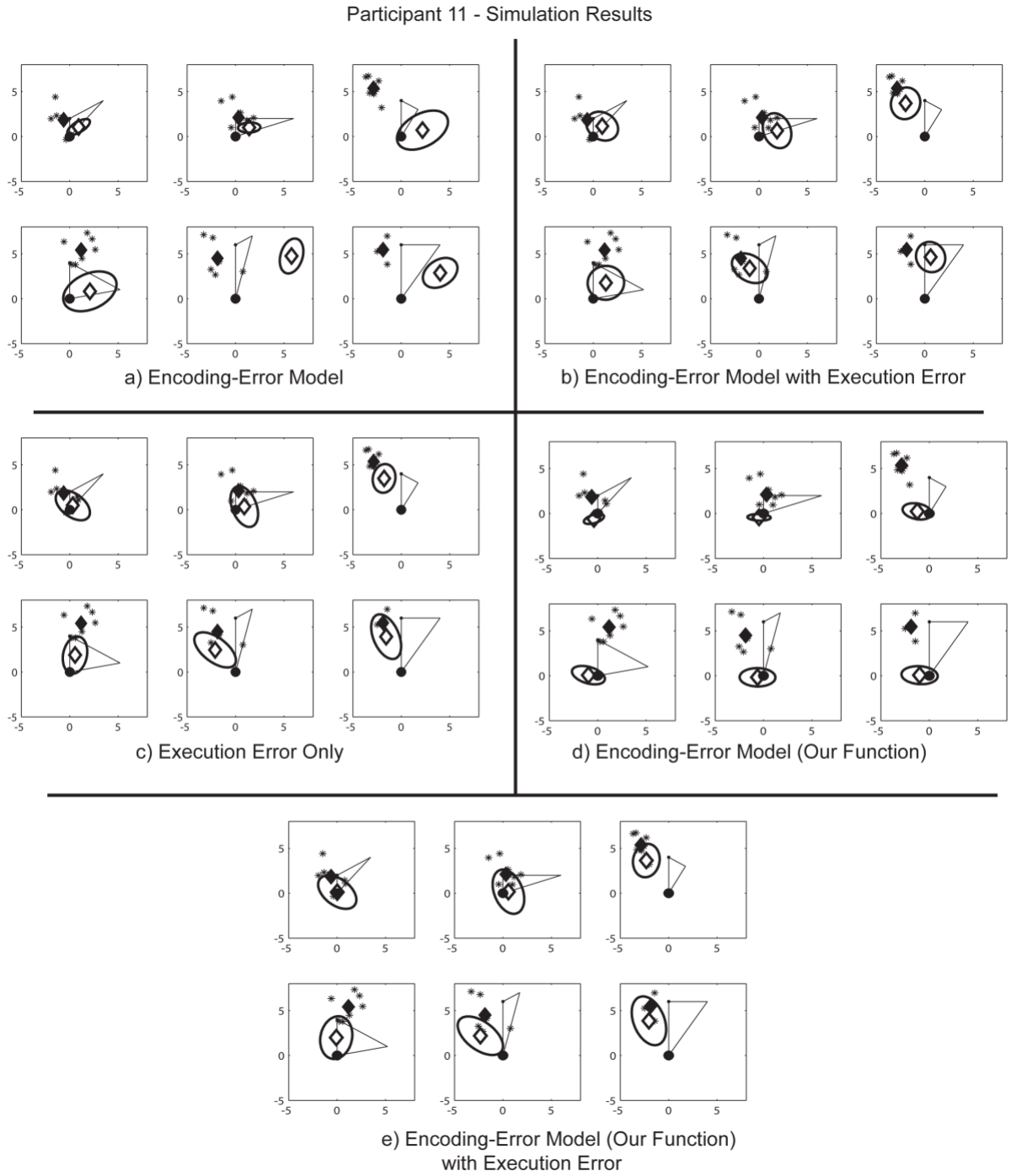

Figure 11. Results of simulations for Participant 11using three different path integration models. Filled circles are the start location, which is also the correct end location. Stars are the individual data points from the eight triangle completion hallway trials. Filled diamonds are the mean final location from the data. Open diamonds are the mean final location from the simulation. Ellipses indicate 95% confidence intervals for the simulation. The stars and filled diamonds are in the same locations across the five simulations because they represent empirical data, but the open diamonds and confidence ellipses differ based on the model. **(A)** Encoding-Error Model; **(B)** Encoding-Error Model with Execution Error; **(C)** Execution Error Only; **(D)** Encoding-Error Model using our new encoding function; **(E)** Encoding-Error Model using the new encoding function and including Execution Error.

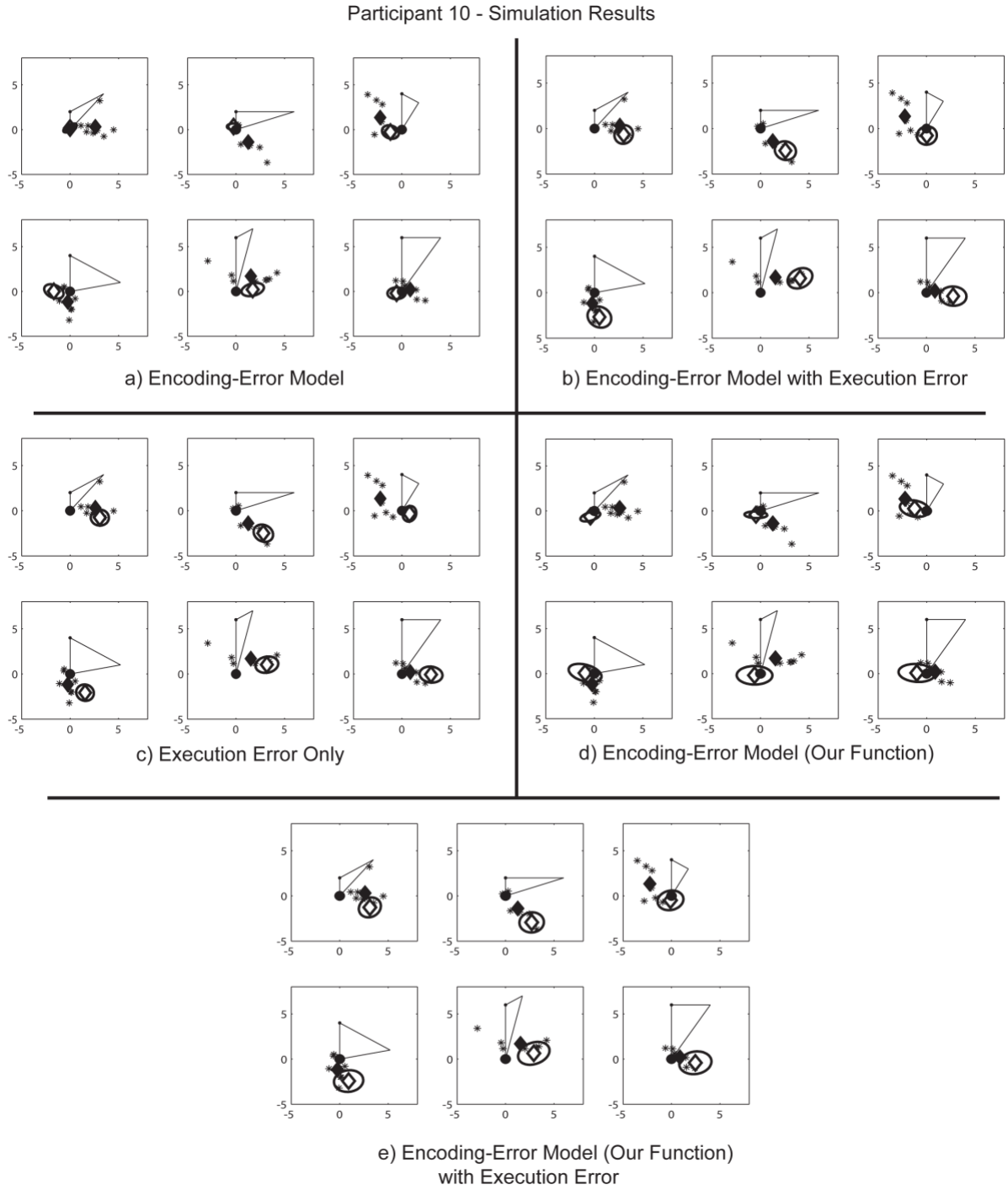

Figure 12. Results of simulations for Participant 10 using three different path integration models. Filled circles are the start location, which is also the correct end location. Stars are the individual data points from the eight triangle completion hallway trials. Filled diamonds are the mean final location from the data. Open diamonds are the mean final location from the simulation. Ellipses indicate 95% confidence intervals for the simulation. The stars and filled diamonds are in the same locations across the five simulations because they represent empirical data, but the open diamonds and confidence ellipses differ based on the model. **(A)** Encoding-Error Model; **(B)** Encoding-Error Model with Execution Error; **(C)** Execution Error Only; **(D)** Encoding-Error Model using our new encoding function; **(E)** Encoding-Error Model using the new encoding function and including Execution Error.

Model Name	Path Length Error Equation	Path Length Error r value	Turn Angle Error Equation	Turn Angle Error r value
1. Encoding-Error Model	y = .45 x + 0.024	0.4818***	y = -.14x – 8.1	-0.2282*
2. Encoding-Error Model with Execution Error	y = 0.73x + 0.078	0.6378***	y = 0.35x – 6.7	0.4837*** †††
3. Execution Error Only Model	y = 0.43x + 0.051	0.7351*** †	y = 0.33x - 3.1	0.4514*** †††
4. Encoding-Error Model (new fcns)	y = 0.049x + 0.78	0.3910***	y = 0.082x + 3.2	0.4355*** †††
5. Encoding-Error Model (new fncs) with Execution Error	y = 0.49x + 0.6	0.7509*** ††	y = 0.36x – 1.3	0.4626*** †††

Table 1. Summary of the simulation results. * indicates a significant correlation between the actual errors and the simulated errors, $p < 0.05$. *** indicates a significant correlation between the actual errors and the simulated errors, $p < 0.001$. † indicates significant difference in correlation between the alternative model and the encoding-error model (model 1). † $p < 0.05$, †† $p < 0.01$, ††† $p < 0.001$.

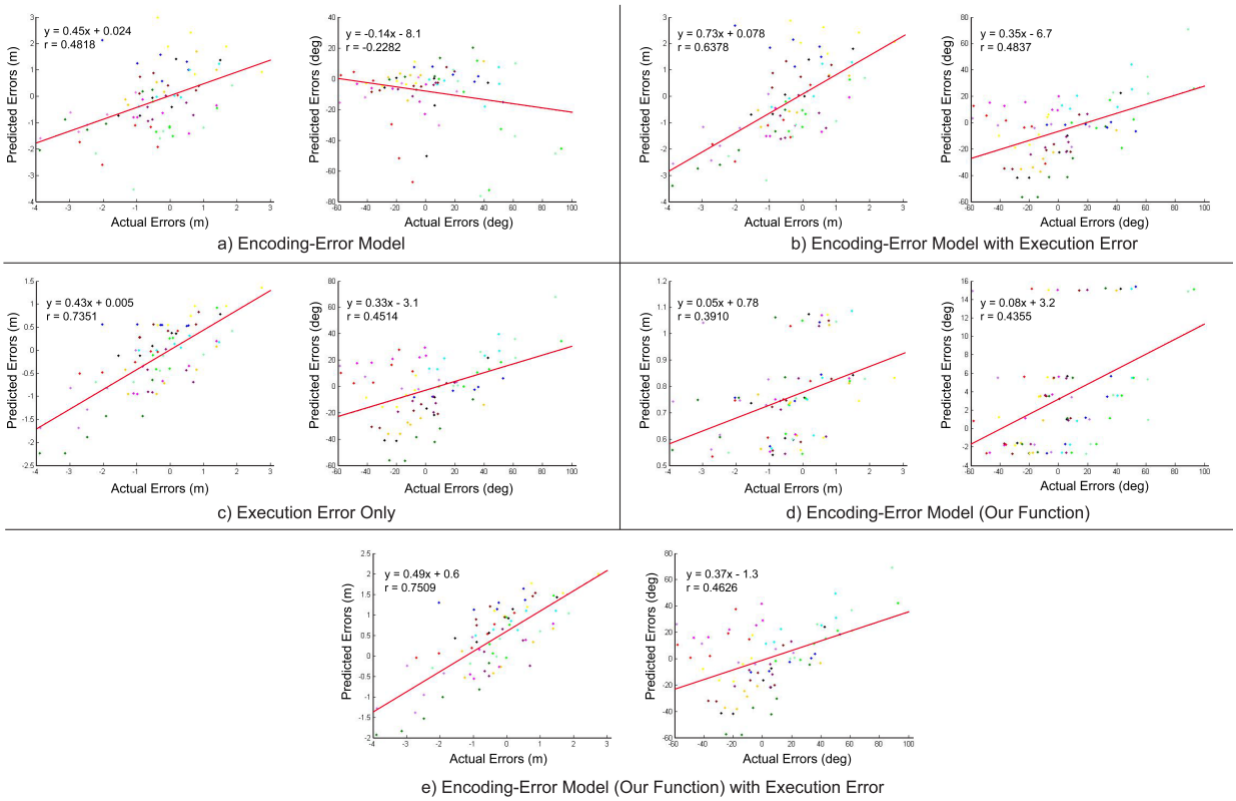

Figure 13. Path length errors (left column) and turn angle errors (right column) on the return path, comparing actual errors made by participants with the errors predicted by each model. Data for individual participants are plotted in different colors; there are six data points for each participant, representing the six triangle types. **(A)** Encoding-Error Model (model 1); **(B)** Encoding-Error Model with Execution Error (model 2); **(C)** Execution Error Only (model 3); **(D)** Encoding-Error Model using our new encoding function (model 4); **(E)** Encoding-Error Model using the new encoding function and including Execution Error (model 5).

The original encoding-error model used aggregate data to derive the encoding functions: average encoding functions predicted average participant performance (Fujita et al., 1993). In contrast, we used each participant's individual distance and angle reproduction data as their individual encoding function. To replicate the approach from the original encoding-error model, we used the average distance and angle errors derived from the reproduction tasks, combined with the average within-subject standard/angular deviations, to simulate the 6 triangles. Fujita et al. (1993) found very high correlations between their predicted and observed values: r^2 of .93 and .92 for distance and angle errors, respectively, with slopes of 1.17 for distance error and .98 for angle error. In contrast, we found a much reduced correlation between predicted and observed values. For distance error, we found $r(4) = 0.9702$, $p < 0.001$, with the equation $y = 0.35x + .0045$ describing the relationship. Although the r value is very high, the slope is much lower than those found by Fujita et al. (1993). For angle error, we found $r(4) = -0.4013$, $p > .4$, with the equation $y = -0.22x – 8.3$ describing the relationship. Thus, there was no relationship between the predicted angular errors and the actual angular errors when considering an "average" participant. These results are in discord with those of Fujita et al., suggesting that their findings either do not replicate or their assumptions are not sufficient to explain systematic errors in triangle completion.

2. Encoding-error model including execution error. Although the path length errors and turn angle errors in the encoding-error model alone were significantly correlated with actual errors, the correlation was negative for turn angle error, indicating the insufficiency of the encoding-error model. We attempted to improve upon this model by adding a component of execution error.

In the new model, the encoding of the outbound path was simulated the same way as in

the encoding-error model (model 1). Instead of assuming that the response path is perfectly integrated and executed, like the encoding-error model does, we added an error term in the execution of the homebound path. Each participant had an execution error term derived from their own reproduction data, as follows. A regression line[2] was fit to the distance reproduction tasks for each participant to describe the relationship between the desired distance and the actual distance walked. For distance, this regression line is essentially the same computation as for the distance reproduction task shown in Figure 9, but for each person individually. Likewise, a regression equation was derived from the combined Same and Opposite angle reproduction. The actual responses for these two conditions were plotted against the required response for each person, similar to Figure 10b, and a line was fit to those 10 data points. Thus, execution functions were fit for both distance and angle that depended on the *required response*, rather than the outbound distance or angle. Although this method is not perfect—it still includes some aspect of encoding—at present it provides a fairly close estimate for execution error. In addition, the findings from the angle reproduction segment of the experiment (Chrastil & Warren, 2017) suggest that errors in reproduction are primarily driven by the required execution, not the encoding angle. We therefore feel reasonably confident in these estimates for execution error.

The outbound path of this model (model 2) was computed in the same way as the encoding-error model (model 1). To add execution error in the homebound path, the desired path length and turn angle were entered into the execution functions described in the previous paragraph for the participant. The resulting path length and turn angle were used for the

[2] We used linear fits for both distance and angle. Given our relatively small distances and angles used, and to match the original encoding-error model, this assumption seemed reasonable. However, for larger distances and angles a logarithmic "leaky integrator" fit might be more appropriate.

simulation. For example, the encoding-error model might yield a desired homebound path of 7 m with a turn angle of 45°. These values were entered into the participant's individual execution functions, and now the homebound path might be something like 6.4 m with a turn angle of 61°. These values were used as the means for our sampling procedure, which added a measure of variability to these execution functions. First, we generated individual regression lines for standard and angular deviation from the empirical reproduction data, much like we did for the execution error functions. Next, we sampled using the means from the execution function and the standard deviations derived from the regression lines. Note that adding variance to the execution error does not change the mean values of the final position, but it does create a larger spread in the simulated results, as evidenced by the 95% confidence ellipses.

The results of the simulations of the encoding-error model including execution error show an improvement over the encoding-error model alone. For path length errors, the function $y = 0.73x + 0.078$ ($r(76) = 0.6378$, $p < 0.001$) related the model-predicted path length errors to the actual path length errors. For turn angle errors, the equation $y = 0.35x – 6.7$ ($r(76) = 0.4837$, $p< 0.001$) related the actual and model-predicted errors. The correlation coefficients were tested against those of the encoding-error model (model 1) using a Fisher's Z score (Table 1). The correlation coefficient for path length errors was not significantly different from model 1 ($Z = -1.4$, $p = .162$), but those for the turn angle errors were significantly different from model 1 ($Z = -4.65$, $p < 0.001$).

Figure 13b summarizes the data for all participants. Examination of Figures 11b and 12b reveals that the addition of execution error to the representative participants has improved the explanatory power of the simulations. Although the original encoding-error model did a reasonable job predicting errors for participant 10, it did not predict participant 11 very well. The

addition of execution error did not make the predictions worse for participant 10, but it dramatically improved those for participant 11. Note that the predicted values for this model are closer to the actual data than for the encoding-error model alone, but they do not overlap perfectly. The addition of execution error makes the model much more predictive of actual errors, but still does not explain all of the variance.

3. **Execution error only.** Our findings from angle reproduction suggest that execution error may be the primary source of error in path integration. The poor results from the encoding-error model (model 1), coupled with the large improvement from the addition of execution error (model 2) suggest that a model that only incorporates execution error may be sufficient to describe the errors in path integration.

For this model, we assumed that encoding error was negligible. This assumption implies that participants were completely accurate at encoding the outbound path, and that their encoded location at the end of the second leg was their actual location. Under this assumption, we used the execution functions from model 2 based on the reproduction tasks (distance and the combined Same and Opposite for angle) as an estimate of pure execution error. We then added execution error in the same manner as in model 2. Note that because there is no distribution of outbound locations in this simulation (since encoding was perfectly accurate), there was only one predicted outcome of the homebound path. Sampling from the distribution of the execution error using the means and standard deviations added some variance, but it was much reduced from that of model 2.

The results of this simulation for all participants are shown in Figure 13c, with details of two representative participants shown in Figures 11c and 12c. For path length errors the equation $y = 0.43x + 0.051$ (, $r(76) = 0.7351$, $p < 0.001$) described the relationship between actual and

model-predicted errors. For turn angle errors the equation y = 0.33x - 3.1 ($r(76) = 0.4514$, $p < 0.001$) described that relationship. The correlation coefficients were tested against those of model 1 using a Fisher's Z score (Table 1). The correlation coefficient for both path length errors and turn angle errors were significantly different from model 1 (path length: $Z = -2.54$, $p = .011$; turn angle: $Z = -4.4$, $p < 0.001$). Based on the significant improvement in the correlation value compared to the baseline encoding-error model, this model appears to be better than the combined model (model 2) at predicting the path length errors, and the turn angle errors are virtually unchanged from model 2. However, we also note that the slope for model 2 is greater than for model 3. This simpler model 3 describes the errors as well as the more complicated model that includes encoding error, therefore these results suggest that most of the error can be attributed to execution error.

4. Encoding-error model with independently estimated encoding functions. As we noted above, it may not be appropriate to use reproduction data to estimate encoding error, for reproduction error may include both encoding and execution errors. Elsewhere we have derived encoding functions for distance and angle based on independent data, which may provide better estimates (Chrastil & Warren, 2014a; unpublished data, see Supplement). Briefly, we began with reproduction data and subtracted out errors from tasks that more closely reflected execution error, such as blind walking to a target.

For distance, Chrastil & Warren (2014a) had participants walk an outbound distance, and then turn and reproduce that distance. This formed the estimate for reproduction. Participants also viewed a target and then turned and walked an equivalent distance. This formed the estimate for production; because participants did not encode the target distance by walking, this task provides a useful – although not perfect – estimate of execution error. We assume that encoding

error from vision is minimal, but not necessarily zero. Encoding distance from vision minimizes the primary problem of canceling errors in pure reproduction, but does not eliminate this issue altogether. Under the assumption of linear combination, we subtracted the value of execution error from reproduction error to estimate the encoding error at each distance. The encoding function was estimated by a linear regression of error (m) on distance (m), yielding $y = 0.9156x + 0.703$ ($r_4 = 0.916$).

For angle, we used a modified reproduction task (unpublished data, see Supplement for details of the task). We removed the second hallway in between the encoding and reproduction turns, reducing any memory decay that could occur between the two turns. We also changed the instructions to "reverse the total of the turns to face in the original direction" rather than "walk parallel to the original path", which could lead to other errors. Although this task is still essentially a reproduction task, it yields an estimate that is closer to production/execution error than the reproduction task ("Same" condition) reported above. Encoding error was then computed as the difference between the reproduction data (Same condition) in the present study and our earlier production data at each turn angle. A linear regression of encoding error on turn angle yielded $y = 0.8356x + 21.217$ ($r_4 = 0.999$).

Note that these derived encoding functions have slopes much closer to 1 than both the functions used by Fujita et al. (1993) and the functions derived from reproduction data above (see Results), implying that encoding error is quite low.

The model simulation procedure was similar to that for the encoding-error model (model 1). However, instead of using individual encoding functions for each participant, we applied the new encoding functions to group averages. Thus, all participants were modeled with the same encoding function, using the overall mean standard deviations from the reproduction data.

Because the new encoding functions predict fairly accurate encoding of the outbound path, the predicted performance should also be fairly accurate. In addition, since we used the same encoding function for each participant, the predicted errors were very similar for all participants (Figures 11d and 12d). Actual path length errors were correlated with predicted path length errors ($r(76) = 0.3910$, $p < 0.001$), with $y = 0.049x + 0.78$ describing the relationship (Figure 13d, Table 1). Actual turn angle errors were also correlated with predicted turn angle errors ($r(76) = 0.4355$, $p < 0.001$), with $y = 0.082x + 3.2$ describing the relationship. The correlation coefficient for path length errors was not significantly different from model 1 ($Z = 0.69$, $p = .490$), while those for the turn angle errors were significantly different from model 1 ($Z = -4.28$, $p < 0.001$). Although the correlation coefficient for angle is an improvement compared to model 1, the slopes for both turn angle and path length were nearly 0.

We also analyzed this encoding function using the overall average participant data by correlating the average turn angle and path length errors for the 6 triangles with the predicted errors from the model. We found the distance error was described by $y = 0.13x + 0.81$ ($r(4) = 0.6294$, $p = 0.181$). This result shows somewhat less predictive value than the encoding-error model. However, the results for the turn angle errors were much improved over the encoding-error model: $y = 0.39x + 2.1$ ($r(4) = 0.9565$, $p = 0.003$). Overall, this model appears to predict errors better than the original encoding error model, at least for an "average" participant.

5. Encoding-error model with independently estimated encoding functions, including execution error. The final model added execution error to model 4, using the execution error functions from model 2. All participants were again modeled with the same encoding function because it was not possible to derive individual encoding functions.

The results of this model showed much improvement over model 4 (Figures 11e, 12e,

13e). The addition of execution errors yielded $y = 0.49x + 0.6$ ($r(76) = 0.7509$, $p < 0.001$) describing the relationship between actual and predicted path length errors. For turn angle errors, the equation $y = 0.36x – 1.3$ ($r(76) = 0.4626$, $p < 0.001$) described the relationship between actual and predicted errors. The correlation coefficient for path length errors was significantly different from model 1 ($Z = -2.75$, $p = .006$), as was that for turn angle errors ($Z = -4.49$, $p < 0.001$) (refer to Table 1). This result is also a slight improvement over model 3 (execution error only). The path length errors were described better by model 5, but the turn angle errors are almost identical. Thus, execution error appears to account for most of the error, but the encoding functions we derived also describe the data better than the encoding-error model.

Discussion

This experiment and accompanying simulations examined sources of error in path integration by segregating encoding and execution errors. Participants reproduced distances and angles in virtual hallways, and we used that reproduction data to model their performance in a triangle completion task. We found minimal differences in path integration performance between an open field and a hallway environment for triangle completion. Simulations revealed that the encoding-error model of path integration was insufficient to explain errors in path integration, but the inclusion of execution error significantly improved the model. Indeed, a model that only included execution error—and no encoding error—predicted the empirical data substantially better than the encoding-error model.

Triangle Completion

Errors from the triangle completion task indicate that participants tended to overturn small angles and underturn large angles. They also compressed the range of the length of the

homebound path. These results fit the general pattern seen in other triangle completion studies (Kearns et al., 2002; Klatzky et al., 1990; Loomis et al., 1993; Peruch et al., 1997). Maurer & Séguinot (Maurer & Séguinot, 1995) noted that most animals overturn on the return path, which puts them back on the outbound path and provides a safety mechanism. In the present study, two triangle types produced underturns, contradicting those observations.

Other accounts of the systematic turn biases observed in triangle completion have been proposed. Maurer & Séguinot (1995) identified the ratio between the first and second legs of the triangle as key to turn angle error. After examining several triangle completion tasks in the literature, they found that undershoots and small overshoots were predicted when the first leg was shorter than or equal to the second leg (ratio ≤1). Large overshoots were predicted when the second leg was longer than the first leg (ratio >1). The results from the present study are not incompatible with this proposal, but the linear relationship found by Maurer & Séguinot does not hold. Although our triangle types with underturns had ratios less than 1, the triangle type with the highest ratio—the 6,2,120 type with a ratio of 3—had an overshoot of less than 10°. In contrast, the 4,2,60 type—with a ratio of 2—had the greatest magnitude of overshoot. Therefore, the leg ratio seems to have some predictive power for turn angle errors, but it does not appear to be as good a predictor as the magnitude of the turn response, described in the next section.

In general, the open field and hallway conditions produced similar results, with some notable exceptions. Turn angle errors in the open field condition were generally closer to zero than those in the hallway condition because participants tended to undershoot and overshoot more in the hallway condition. Participants also generally walked somewhat further on the homebound leg in the hallway condition than in the open field. Variability for all measures was similar in both open field and hallway conditions, although the open field showed occasional

increased variability for angular measures. While the data do not indicate that these conditions are completely equivalent, they do suggest that the paradigms are similar enough to be able to generalize results from the hallway to those in an open field. The correspondence of errors between the open field and hallway conditions in triangle completion implies that the additional optic flow information in the hallway environment may not aid path integration greatly. These results agree with those of Kearns et al. (2002), who found similar accuracy in triangle completion between arenas with full optic flow information and those with reduced optic flow. The relatively small FOV (60^{o} horizontal x 46.8^{o} vertical) in the HMD may have also reduced the amount of optic flow overall.

Reproduction Tasks

Distance reproduction errors show that participants compressed the response space, generally overshooting short distances and undershooting long distances. These results agree with most previous work on distance reproduction (Israël et al., 2004a; May & Klatzky, 2000; M.-L. Mittelstaedt & Mittelstaedt, 2001; Schwartz, 1999; Sun et al., 2004). Moreover, the average regression equation for the distance reproduction task in this study is very similar to the encoding function used by Fujita et al. (1993). As seen in the tests of the encoding-error model by Klatzky et al. (1999), the compression of distance seen in previous work seems to be a regression to the mean, depending on the context (Petzschner & Glasauer, 2011; Schwartz, 1999).

Angle reproduction errors indicate that participants tended to overturn small *responses*—not necessarily small encoding angles—and underturned large responses (for detailed discussion of angle estimation, see Chrastil & Warren, 2017). As a result, it appears that the errors in angle reproduction stemmed less from encoding the turn angle than from executing the required

response. The compression for the response angles was not centered around the mean, as might be expected from previous studies. Instead, participants overturned small responses by a large margin and only underturned large responses by a small amount. Thus, instead of having the most accurate responses being at 90° as expected for the mean value, the most accurate responses came when participants had to turn 120° to 135°, leaving a 60°–45° angle internal angle.

These angle reproduction errors yield a striking result when examined in light of the turn angle errors from the triangle completion task. In triangle completion, the two most accurate triangle types (triangle types 4,6,60 and 6,2,120; compare results shown in Fig. 8b with the patterns seen in Fig. 10b), called for homebound turns of 139.11° and 133.90°, respectively, which fall within or near this accurate range. Likewise, the errors for the 6,4,90 triangle type should be fairly small; the data agree with this prediction. The 4,2,60 triangle type calls for a 90° turn. According to the angle reproduction task, 90° should be overshot by 15°–20°, which falls within the standard error of the actual overshoot of 27°. Finally, the two triangle types for which participants tended to turn not far enough (types 2,4,120 and 2,6,90) called for response turns of 160.89° and 161.56°, respectively. Based on the angle reproduction errors, these responses should be underturned by about 10°, which is in line with the actual errors seen in the triangle completion task. Thus, the turn required to make an accurate response in triangle completion may dictate the pattern of turn angle errors. This observation is supported by the results from the simulations indicating that execution error makes the largest contribution to the errors in path integration.

Encoding-Error Model

Our simulations demonstrate that the encoding-error model of path integration did not successfully describe empirical errors in triangle completion. The predicted path length and turn

angle errors in the homebound path had a weak correlation with actual distance errors, and a negative correlation with actual angle errors. Yet, both our average distance and angle reproduction functions are similar to those of the encoding-error model. Thus, the source of the differing results in these two studies must lie either in the application of these encoding functions, or in the performance of triangle completion itself. Fujita and colleagues applied their aggregate functions to both aggregate and individual triangle completion data. Their aggregate data matched well, but the individual data did not match as well. The current experiment used individual encoding functions and applied them to individual triangle completion data, with poor results for the encoding-error model. We also used our average encoding functions to predict an "average" participant's performance. The results of those simulations produced high correlations for distance errors, but still yielded negative correlations for angle errors. Thus, when simulating both individual participant performance and overall average performance, the encoding-error model did not capture the errors observed in triangle completion.

In addition to poor simulation results, the experimental data contradict the encoding-error model in other ways. Violations of either the first assumption (satisfaction of Euclidean axioms) or the fourth assumption (no integration or execution error) were seen in instances in which participants had negative turn angles (i.e. they turned > 180^{o}) for the homebound trajectory (e.g., Figure 11, upper right triangle). Even if encoded poorly, no Euclidean triangle would call for such a response. In addition, the results from the angle reproduction task violate the encoding-error model's assumption of no execution error: a 90^{o} turn should be executed fairly accurately because it lies on a reference axis, but participants overshot this turn by 15^{o}–20^{o}.

As noted in the introduction, other researchers have attempted to use the encoding-error model to explain their data. These previous efforts have been mixed in their assessment. The

original encoding-error model (Fujita et al., 1993) and subsequent positive tests (Klatzky et al., 1999; May & Klatzky, 2000) stemmed from blind walking tasks in Loomis et al. (Loomis et al., 1993) and used blindfolded participants. Thus, the encoding functions they found may only apply to non-visual navigation. Even so, Klatzky et al. (1999) found inconsistencies in one of the assumptions of the encoding-error model using a blind-walking paradigm. In purely visual tasks, Péruch et al. (Peruch et al., 1997) successfully applied the encoding functions to triangle completion in desktop VR, but Riecke et al. (Riecke et al., 2002) did not have the same success. However, they used feedback during training on triangle completion, possibly influencing performance. We used both visual and idiothetic information in the reproduction and triangle completion tasks. Previous research suggests that both visual and idiothetic information make similar contributions to path integration (Chrastil, Nicora, & Huang, 2019; Kearns et al., 2002; Tcheang et al., 2011). Although our methods differed from those who used blindfolded walking, the basic assumptions of the encoding-error model should still hold. However, our findings went largely against this model.

The results of our simulations, although contrary to some of the assumptions of the encoding-error model, do not completely discredit the model altogether. Evidence suggests that encoding functions may be context-dependent and can be modified by the experience of the navigator (Abdolvahab et al., 2015; Arthur et al., 2012; Chrastil & Warren, 2014a; Klatzky et al., 1999; Petzschner & Glasauer, 2011; Schwartz, 1999; Turvey et al., 2009), in opposition to the assumptions of the encoding-error model. However, in the current study, each participant had their own encoding functions, using a range of values for leg length and turn angle experienced in both the outbound and return paths of triangle completion. Furthermore, the size and structure of the environment were similar for both the triangle completion and reproduction tasks. Given

these restrictions, the assumption of a single encoding function for each person—at least for the scale and structure of this environment—seems reasonable. Although it is not necessarily the case that encoding plays no role in path integration, the simulation results indicate that encoding alone cannot account for all of the systematic errors seen in path integration.

Other Models of Path Integration

Because the encoding-error model of path integration proved inadequate to describe the errors observed during a triangle completion task, we extended and modified this model in several ways. We then simulated path integration data for these alternative models using the same Monte Carlo methods used to test the encoding-error model. The alternative models showed significant improvement over the encoding-error model.

The first alternative (model 2) incorporated execution errors into the basic encoding-error model. The addition of execution errors proved to be a significant improvement over the encoding-error model for predicting the angular errors. Although the correlation coefficient was not significantly better than the encoding-error model for distance errors, model 2 had the best slope of any alternative model. This improvement was so striking that we next attempted to explain the variance through execution errors alone. This model (model 3) also improved predictions for angular error in the simulations compared to the encoding error model. From these results, we determined that encoding may make a minor contribution to path integration errors, but the majority of error was explained by execution error alone.

Although the encoding-error model showed only limited success, the approach was not abandoned altogether. Instead, we considered the possibility that the reproduction task used to estimate the encoding functions for this model were not appropriate. As noted earlier, reproduction tasks confound encoding and execution errors. If the primary source of error in

reproduction tasks stems from execution rather than encoding, as suggested by the results of the angle tasks (Chrastil & Warren, 2017), then those encoding functions were heavily weighted toward execution error. This could be a reason that the encoding-error model (model 1) did not perform well. Model 4 returned to a model comprised of only encoding errors, but used encoding functions derived from independent data (Chrastil & Warren, 2014; unpublished data), yielding low, but not negligible, encoding error. The resulting simulations had higher correlations with angular errors, but the slopes for both distance and angle errors were nearly 0. Thus, although the fits appeared better, this model predicted similar errors for all triangle types. Finally, model 5 combined the encoding error model using our new encoding functions with the addition of execution error. This model was significantly better at predicting both distance and angle errors than the encoding-error model. These results suggest that a model which incorporates small encoding errors and large execution errors best describes the errors in path integration.

The slopes and correlation coefficients for angular error between models 2, 3, and 5 were very similar, suggesting that the addition of execution error was primarily responsible for the angular component of error (Table 1, last two columns). Unlike for angular error, distance error was affected by the encoding function used; the slopes for distance errors between models 2 and 5 differed quite a bit (.73 and .49, respectively, see Table 1). On one hand, model 2 has a slope closer to 1, but on the other hand, the correlation effect size is stronger in model 5. In either case, the addition of execution error is important for improving the predictive value of the model. These results suggest that encoding error may make a much smaller contribution than execution error, but that it is not completely negligible.

Although the encoding-error model showed only limited success, it is possible that the encoding functions used in the encoding-error model were not appropriate for this task. As noted

earlier, it is difficult to measure pure encoding or execution errors, and so both the encoding functions derived from reproduction data (models 1 and 3) and the new encoding functions derived from subtracting out production errors (models 4 and 5) could include a degree of execution error (and vice versa for the execution functions). Under a purely encoding-error model, reproduction tasks reflect encoding, but it is possible that encoding and execution compensate for each other in these tasks. In that case, reproduction tasks might underestimate both encoding and execution errors. Having two measures of execution in the angle reproduction tasks gives us greater confidence in those functions, as does the blind walking measure of production our distance encoding function. Our attempts at deriving new encoding functions, which minimize the contribution of execution, are a step in the direction toward separating these factors. However, we acknowledge that these functions are not perfect and still include both encoding and execution errors. Methods for separating encoding and execution are still not resolved and thus we cannot conclude with full confidence that execution error is the primary source of error. However, our results suggest it is a major source of error. Furthermore, the models demonstrate that encoding error alone is not sufficient to explain path integration errors.

Finally, it is important to consider other limitations of these models. First, we did not include an explicit term for integration error—the error involved in computing the trajectory for the homebound path—which could be substantial. At present, we have not found an appropriate method for isolating and quantifying integration error; thus, integration error cannot be directly added to the alternative models of path integration. Accurate encoding coupled with substantial errors in integration could pose a problem for the idea of a metric cognitive map. Although a metric map assumes that distances and angles can be measured accurately, it also requires accurate integration for a global metric embedding. However, integration errors would be

consistent with a labeled graph, which only requires roughly accurate encoding (Warren, 2019).

Second, all of the models assume that errors add linearly based on configural path integration. It is possible that a model involving a homing vector is more appropriate to describe path integration, although its step-wise error accumulation follows a similar pattern to that seen in the encoding-error model. As described in the introduction, the evidence is mixed as to whether human path integration follows a configurational or homing vector model (Fujita et al., 1990; Mou & Zhang, 2014; Muller & Wehner, 1988), although recent behavioral and neural evidence points to both systems being present (Chrastil et al., 2015; Chrastil, Sherrill, Hasselmo, & Stern, 2016; He & McNamara, 2018; Wiener et al., 2011). However, there are no current methods of predicting the individual differences observed in human path integration using a homing vector model, whereas configurational models allow for predicting errors on an individual basis. The dissociation between position and heading estimations with an allocentric homing vector (Mou & Zhang, 2014; Zhang & Mou, 2017) could prove informative as a hybrid model in the future.

Summary and Conclusions

A triangle completion task was used to test the encoding-error model of path integration. Errors from the Open Field and Hallway version of triangle completion were similar, indicating that findings in a hallway generalize to other environments. Participants generally underturned large required turns and overturned small required turns, and underestimated long distances and overestimated short distances. Angle reproduction showed that errors were not proportional to the outbound turn angle, but instead to the required response turn. When the errors from the distance and angle reproduction tasks were applied to simulations of the encoding-error model, it

did not adequately describe the systematic errors seen in a triangle completion task. Analysis of the alternative models revealed that both encoding and execution error contribute to errors in path integration, but with execution errors playing the dominant role. These results challenge the assumption that errors in both reproduction tasks and more complex path integration experiments stem solely from errors in encoding. Errors in triangle completion might not arise from failing to know where you are, but from an inability to get back home.

References

Abdolvahab, M., Carello, C., Pinto, C., Turvey, M. T., & Frank, T. D. (2015). Symmetry and order parameter dynamics of the human odometer. *Biological Cybernetics*, *109*(1), 63–73. https://doi.org/10.1007/s00422-014-0627-1

Arthur, J. C., Philbeck, J. W., Kleene, N. J., & Chichka, D. (2012). The role of spatial memory and frames of reference in the precision of angular path integration. *Acta Psychologica*, *141*(1), 112–121. https://doi.org/10.1016/j.actpsy.2012.07.003

Avraamides, M. N., Klatzky, R. L., Loomis, J. M., & Golledge, R. G. (2004). Use of cognitive versus perceptual heading during imagined locomotion depends on the response mode. *Psychological Science*, *15*(6), 403–408. https://doi.org/10.1111/j.0956-7976.2004.00692.x

Bakker, N. H., Werkhoven, P. J., & Passenier, P. O. (1999). The Effects of Proprioceptive and Visual Feedback on Geographical Orientation in Virtual Environments. *Presence: Teleoperators & Virtual Environments*, *8*(1), 36–53.

Bakker, N. H., Werkhoven, P. J., & Passenier, P. O. (2001). Calibrating Visual Path Integration in VEs. *Presence*.

Becker, W., Jürgens, R., & Boss, T. (2000). Vestibular perception of self-rotation in different postures: a comparison between sitting and standing subjects. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *131*(4), 468–476.

Benhamou, S., Sauve, J.-P., & Bovet, P. (1990). Spatial memory in large scale movements: efficiency and limitation of the egocentric coding process. *Journal of Theoretical Biology*, *145*, 1–12.

Benhamou, S., & Séguinot, V. (1995). How to Find one's Way in the Labyrinth of Path Integration Models. *Journal of Theoretical Biology*, *174*, 463–466.

Chen, X., He, Q., Kelly, J. W., Fiete, I. R., & McNamara, T. P. (2015). Bias in Human Path Integration Is Predicted by Properties of Grid Cells. *Current Biology : CB*. https://doi.org/10.1016/j.cub.2015.05.031

Cheng, K., Shettleworth, S. J., Huttenlocher, J., & Rieser, J. J. (2007). Bayesian Integration of Spatial Information. *Psychological Bulletin, 133*(4), 625–637.

Chrastil, E. R., Nicora, G. L., & Huang, A. (2019). Vision and proprioception make equal contributions to path integration in a novel homing task. *Cognition*, *192*, 103998. https://doi.org/10.1016/J.COGNITION.2019.06.010

Chrastil, E. R., Sherrill, K. R., Hasselmo, M. E., & Stern, C. E. (2015). There and back again: Hippocampus and retrosplenial cortex track homing distance during human path integration. *J. Neurosci.*, *35*(46), 15442–15452. https://doi.org/10.1523/JNEUROSCI.1209-15.2015

Chrastil, E. R., Sherrill, K. R., Hasselmo, M. E., & Stern, C. E. (2016). Which way and how far? Tracking of translation and rotation information for human path integration. *Human Brain Mapping*, *37*(10), 3636–3655. https://doi.org/10.1002/hbm.23265

Chrastil, E. R., & Warren, W. H. (2014a). Does the human odometer use an extrinsic or intrinsic metric? *Attention, Perception, & Psychophysics*, *76*(1), 230–246. https://doi.org/10.3758/s13414-013-0549-3

Chrastil, E. R., & Warren, W. H. (2014b). From cognitive maps to cognitive graphs. *PLoS ONE*, *9*(11), e112544. https://doi.org/10.1371/journal.pone.0112544

Chrastil, E. R., & Warren, W. H. (2017). Rotational error in path integration: encoding and execution errors in angle reproduction. *Experimental Brain Research*, *235*(6), 1885–1897. https://doi.org/10.1007/s00221-017-4910-y

Etienne, A. S., Maurer, R., & Saucy, F. (1988). Limitations in the Assessment of Path Dependent

Information. *Behaviour*, *106*, 81–111.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. In *Behavior Research Methods* (Vol. 39, pp. 175–191). https://doi.org/10.3758/BF03193146

Foo, P., Duchon, A. P., Warren, W. H., & Tarr, M. J. (2007). Humans do not switch between path knowledge and landmarks when learning a new environment. *Psychological Research*, *71*, 240–251.

Foo, P., Warren, W. H., Duchon, A., & Tarr, M. J. (2005). Do humans integrate routes into a cognitive map? Map- versus landmark-based navigation of novel shortcuts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(2), 195–215.

Fujita, N., Klatzky, R. L., Loomis, J. M., & Golledge, R. G. (1993). The encoding-error model of pathway completion without vision. *Geographical Analysis*, *25*(4), 295–314.

Fujita, N., Loomis, J. M., Klatzky, R. L., & Golledge, R. G. (1990). A minimal representation for dead-reckoning navigation: Updating the homing vector. *Geographical Analysis*, *22*(4), 326–335.

Gallistel, C. R. (1990). *The Organization of Learning*. Cambridge, MA: MIT Press.

Harris, L. R., Jenkin, M., & Zikovitz, D. C. (2000). Visual and non-visual cues in the perception of linear self motion. *Experimental Brain Research*, *135*(1), 12–21.

He, Q., & McNamara, T. P. (2018). Spatial updating strategy affects the reference frame in path integration. *Psychonomic Bulletin and Review*, *25*(3), 1073–1079. https://doi.org/10.3758/s13423-017-1307-7

Israël, I., Bronstein, a M., Kanayama, R., Faldon, M., & Gresty, M. a. (1996). Visual and vestibular factors influencing vestibular "navigation". *Experimental Brain Research.*

Experimentelle Hirnforschung. Expérimentation Cérébrale, *112*(3), 411–419.

Israël, I., Capelli, a, Sablé, D., Laurent, C., Lecoq, C., & Bredin, J. (2004a). Multifactorial interactions involved in linear self-transport distance estimate: a place for time. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *53*(1), 21–28. https://doi.org/10.1016/j.ijpsycho.2004.01.002

Israël, I., Capelli, A., Sablé, D., Laurent, C., Lecoq, C., & Bredin, J. (2004b). Multifactorial interactions involved in linear self-transport distance estimate: a place for time. *International Journal of Psychophysiology*, *53*(1), 21–28.

Israël, I., Sievering, D., & Koenig, E. (1995). Self-rotation estimate about the vertical axis. *Acta Oto-Laryngologica*, *115*(1), 3–8.

Ivanenko, Y., Grasso, R., Israël, I., & Berthoz, a. (1997). Spatial orientation in humans: perception of angular whole-body displacements in two-dimensional trajectories. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *117*(3), 419–427.

Jetzschke, S., Ernst, M. O., Moscatelli, A., & Boeddeker, N. (2016). Going round the bend: Persistent personal biases in walked angles. *Neuroscience Letters*, *617*, 72–75. https://doi.org/10.1016/j.neulet.2016.01.026

Jurgens, R., Nasios, G., & Becker, W. (2003). Vestibular, Optokinetic, and Cognitive Contribution to the Guidance of Passive Self-Rotation Toward Instructed Targets. *Experimental Brain Research*, *151*, 90–107.

Harootonian, S. K., Wilson, R. C., Hejtmanek, L., Ziskin, E. M., & Ekstrom, A. D. (in press). Path integration in large-scale space and with novel geometries: Comparing vector addition

and encoding-error models. *PLoS Computational Biology.*

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*(3), 349–374.

Klatzky, R. L., Beall, A. C., Loomis, J. M., Golledge, R. G., & Philbeck, J. W. (1999). Human navigation ability: Tests of the encoding-error model of path integration. *Spatial Cognition and Computation*, *1*, 31–65.

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.

Klatzky, R. L., Loomis, J. M., & Golledge, R. G. (1997). Encoding Spatial Representations Through Nonvisually Guided Locomotion: Tests of Human Path Integration. In D. L. Medin (Ed.), *Psychology of Learning and Motivation* (Vol. Volume 37, pp. 41–84). Academic Press. https://doi.org/DOI: 10.1016/S0079-7421(08)60499-5

Klatzky, R. L., Loomis, J. M., Golledge, R. G., Cicinelli, J. G., Doherty, S., & Pellegrino, J. W. (1990). Acquisition of route and survey knowledge in the absence of vision. *Journal of Motor Behavior*, *22*(1), 19–43.

Lappe, M., & Frenz, H. (2009). Visual estimation of travel distance during walking. *Experimental Brain Research*, *199*(3–4), 369–375. https://doi.org/10.1007/s00221-009-1890-6

Lappe, M., Jenkin, M., & Harris, L. R. (2007). Travel distance estimation from visual motion by leaky path integration. *Experimental Brain Research*, *180*, 35–48.

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.

Marlinsky, V. V. (1999). Vestibular and vestibulo-proprioceptive perception of motion in the horizontal plane in blindfolded man--I. Estimations of linear displacement. *Neuroscience*, *90*(2), 389–394.

Maurer, R., & Séguinot, V. (1995). What is modelling for? a critical review of the models of path integration. *Journal of Theoretical Biology*, *175*(4), 457–475.

May, M., & Klatzky, R. L. (2000). Path integration while ignoring irrelevant movement. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I., & Moser, M.-B. (2006). Path integration and the neural basis of the "cognitive map". *Nature Reviews Neuroscience*, *7*(8), 663–678. https://doi.org/10.1038/nrn1932

Merkle, T., Rost, M., & Alt, W. (2006a). Egocentric path integration models and their application to desert arthropods. *Journal of Theoretical Biology*, *240*(3), 385–399. https://doi.org/10.1016/j.jtbi.2005.10.003

Merkle, T., Rost, M., & Alt, W. (2006b). Egocentric path integration models and their application to desert arthropods. *Journal of Theoretical Biology*, *240*(3), 385–399.

Mittelstaedt, H., & Mittelstaedt, M.-L. (1982). Homing by Path Integration. In F. Papi & H. G. Wallraff (Eds.), *Avian Navigation* (pp. 290–297). Berlin: Springer-Verlag.

Mittelstaedt, M.-L., & Mittelstaedt, H. (1980). Homing by path integration in a mammal. *Naturwissenschaften*, *67*, 566–567.

Mittelstaedt, M.-L., & Mittelstaedt, H. (2001). Idiothetic navigation in humans: estimation of path length. *Experimental Brain Research*, *139*(3), 318–332.

Mou, W., & Zhang, L. (2014). Dissociating position and heading estimations: Rotated visual

orientation cues perceived after walking reset headings but not positions. *Cognition*, *133*(3), 553–571. https://doi.org/10.1016/j.cognition.2014.08.010

Muller, M., & Wehner, R. (1988). Path integration in desert ants, Cataglyphis fortis. *Proceedings of the National Academies of Science*, *85*, 5287–5290.

Peruch, P., May, M., & Wartenberg, F. (1997). Homing in Virtual Environments: Effects of Field of View and Path Layout. *Perception*, *26*, 301–311.

Petzschner, F. H., & Glasauer, S. (2011). Iterative Bayesian estimation as an explanation for range and regression effects: a study on human path integration. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(47), 17220–17229. https://doi.org/10.1523/JNEUROSCI.2028-11.2011

Philbeck, J. W., Klatzky, R. L., Behrmann, M., Loomis, J. M., & Goodridge, J. (2001). Active control of locomotion facilitates nonvisual navigation. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 141–153. https://doi.org/10.1037/0096-1523.27.1.141

Redlick, F. P., Jenkin, M., & Harris, L. R. (2001). Humans can use optic flow to estimate distance of travel. *Vision Research*, *41*(2), 213–219.

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. https://doi.org/10.1162/105474602320935810

Schwartz, M. (1999). Haptic perception of the distance walked when blindfolded. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 852–865.

Seguinot, V., Maurer, R., & Etienne, A. S. (1993). Dead reckoning in a small mammal: the evaluation of distance. *Journal of Comparative Physiology A*, *173*, 103–113.

Siegler, I. (2000). Idiosyncratic orientation strategies influence self-controlled whole-body rotations in the dark. *Brain Research. Cognitive Brain Research*, *9*(2), 205–207.

Siegler, I., Viaud-Delmon, I., Israël, I., & Berthoz, a. (2000). Self-motion perception during a sequence of whole-body rotations in darkness. *Experimental Brain Research*, *134*(1), 66–73. https://doi.org/10.1007/s002210000415

Sommer, S., & Wehner, R. (2004). The ant's estimation of distance travelled: experiments with desert ants, Cataglyphis fortis. *Journal of Comparative Physiology A*, *190*(1), 1–6.

Sun, H. J., Campos, J. L., Young, M., Chan, G. S. W., & Ellard, C. G. (2004). The contributions of static visual cues, nonvisual cues, and optic flow in distance estimation. *Perception*, *33*(1), 49–65.

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. https://doi.org/10.1073/pnas.1011843108

Turvey, M. T., Romaniak-Gross, C., Isenhower, R. W., Arzamarski, R., Carello, C., & Harrison, S. (2009). Human odometer is gait-symmetry specific. *Proceedings of The Royal Society B. Biological Sciences*, *276*(1677), 4309–4314. https://doi.org/10.1098/rspb.2009.1134

Vidal, M., & Bülthoff, H. H. (2010). Storing upright turns: how visual and vestibular cues interact during the encoding and recalling process. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *200*(1), 37–49. https://doi.org/10.1007/s00221-009-1980-5

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.

https://doi.org/10.1080/13875868.2012.678521

Wang, R. F. (2015). Building a cognitive map by assembling multiple path integration systems. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-015-0952-y

Warren, W. H. (2019). Non-Euclidean navigation. *The Journal of Experimental Biology*, *222*(jeb187971), 1–10. https://doi.org/10.1242/jeb.187971

Warren, W. H., Rothman, D. B., Schnapp, B. H., & Ericson, J. D. (2017). Wormholes in virtual space: From cognitive maps to cognitive graphs. *Cognition*, *166*, 152–163. https://doi.org/10.1016/j.cognition.2017.05.020

Wehner, R., & Wehner, S. (1986). Path integration in desert ants: approaching a long-standing puzzle in insect navigation. *Monitore Zoologico Italiano*, *20*, 309–331.

Wiener, J. M., Berthoz, A., & Wolbers, T. (2011). Dissociable cognitive mechanisms underlying human path integration. *Experimental Brain Research*, *208*(1), 61–71. https://doi.org/10.1007/s00221-010-2460-7

Wiener, J. M., & Mallot, H. A. (2006). Path complexity does not impair visual path integration. *Spatial Cognition and Computation*, *6*(4), 333–346.

Wittlinger, M., Wehner, R., & Wolf, H. (2006). The ant odometer: stepping on stilts and stumps. *Science (New York, N.Y.)*, *312*(5782), 1965–1967. https://doi.org/10.1126/science.1126912

Yamamoto, N., Meléndez, J. A., & Menzies, D. T. (2014). Homing by path integration when a locomotion trajectory crosses itself. *Perception*, *43*(10), 1049–1060.

Zhang, L., & Mou, W. (2017). Piloting systems reset path integration systems during position estimation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(3), 472–491. https://doi.org/10.1037/xlm0000324

Zhao, M., & Warren, W. H. (2015a). Environmental stability modulates the role of path

integration in human navigation. *Cognition*, *142*, 96–109. https://doi.org/10.1016/j.cognition.2015.05.008

Zhao, M., & Warren, W. H. (2015b). How you get there from here: interaction of visual landmarks and path integration in human navigation. *Psychological Science*, *26*(6), 915–924. https://doi.org/10.1177/0956797615574952