

Sources of path integration error **in young and aging humans**

Matthias Stangl^{1*}, Ingmar Kanitscheider^{2*}, Martin Riemer^{1,3}, Ila Fiete^{2,4**}, Thomas Wolbers^{1,3**}

¹ German Center for Neurodegenerative Diseases (DZNE), Aging & Cognition Research Group, Magdeburg, Germany

² Center for Learning and Memory, Department of Neuroscience, The University of Texas, Austin, USA

³ Center for Behavioral Brain Sciences, Magdeburg, Germany

⁴ Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology (MIT), Cambridge, USA

* These authors have contributed equally to this work

** These authors have also contributed equally to this work

Correspondence to:

M. Stangl (stangl.mat@gmail.com) or I. Kanitscheider (ingmar@openai.com)

Abstract

Path integration is a vital function in navigation: it enables the continuous tracking of one's position in space by integrating self-motion cues. Path integration abilities vary across individuals but tend to deteriorate in old age. The specific causes of path integration errors, however, remain poorly characterized. Here, we combined tests of path integration performance with a novel analysis based on the Langevin diffusion equation, which allowed us to decompose errors into distinct causes that can corrupt path integration computations. Across age groups, the dominant errors were due to noise and a bias in speed estimation. Noise-driven errors accumulated with travel distance not elapsed time, suggesting that the dominant noise originates in the velocity input rather than within the integrator. Age-related declines were traced primarily to a growth in this unbiased noise. Together, these findings shed light on the contributors to path integration error and the mechanisms underlying age-related navigational deficits.

Keywords:

spatial navigation, human, computational modeling, path integration error, error sources, aging

Introduction

Spatial navigation is a complex behavior that combines many computations, including the storage and recall of information, the integration of information from multiple sensory and non-sensory brain areas, planning, prediction, and decision making. A vital component of navigation-related computations is path integration – the integration over time of a self-motion estimate, in the strict sense of vector calculus, to maintain an updated estimate of one's position and orientation while moving through space.

Self-motion estimates themselves derive from a sophisticated pooling over multiple sensory modalities, and rely on proprioceptive and vestibular information, visual optic flow signals (i.e., the pattern of apparent motion of objects, surfaces and edges), as well as motor efference copies that are produced during movement (Etienne & Jeffery, 2004). After being processed in their respective low-level sensory systems, these cues are integrated in brainstem nuclei as well as cortical structures (area MST) to allow an overall estimation of angular and linear movement velocity (Bassett & Taube, 2001; Biazoli et al., 2006; Angelaki & Cullen, 2008; Britten, 2008; Clark et al., 2012; Cullen, 2012; Butler & Taube, 2015). The integration of these cues is an error-prone process, and previous studies have demonstrated that path integration abilities therefore vary largely across individuals (Loomis et al., 1993; Klatzky et al., 1999; Chrastil et al., 2017). However, we have only a limited understanding of the specific sources of error that may corrupt path integration computations. In this work, we obtain quantitative measurements of path integration performance in subjects of different ages and develop and apply a method to decompose the observed path integration errors into components that can shed light on the mechanisms that underlie the observed errors (cf. Brunton et al., 2013).

A circuit that functions as a path integrator for two-dimensional space must do the following: take as input the given two-dimensional velocity signal; remember the previous integrated state; increment the previous integrated state by adding to it a quantity proportional to the instantaneous velocity input. There are thus several natural sources of error: First, the velocity estimate might be wrong, with systematic bias or unbiased noise. Second, the integrator might remember its past states in a leaky way, so that there is a decay of information over time. Third, the velocity input-based increments might be summed with a scaling or gain prefactor that differs from the value required to match the instantaneous displacement. Fourth, the integrator might itself be noisy.

These errors accrue over the course of a spatial movement trajectory, and the net localization error at path's end will depend on the details of the trajectory. Thus, properly modeling and decomposing these errors requires iteration of a temporal dynamics, a statistical model that incorporates these dynamics, and sufficiently rich and varied spatial trajectories in the input data. One final error arises when a downstream neural circuit or the human experimenter attempt to obtain a readout or report of the internal state of the integrator.

Our goal in the present work is not only to understand the contributors to path integration error, but also to reveal sources of age-related degradation in navigation performance. Specifically, aging has deleterious effects on path integration ability,

with declines in the triangle completion task (Loomis et al., 1993) – a standard assay of path integration performance (Allen et al., 2004; Mahmood et al., 2009; Adamo et al., 2012; Harris & Wolbers, 2012). Older adults are less accurate in reproducing travel distances or rotations (Mahmood et al., 2009; Adamo et al., 2012; Harris & Wolbers, 2012), and they exhibit worse path integration performance even if additional landmark information is available (Harris & Wolbers, 2012; Bates & Wolbers, 2014). Despite the sizeable body of research on losses in path integration performance with age, little is known about which specific aspects of the path integration computation or process are most affected in old age.

To address this important question, we combine an immersive virtual reality path integration experiment with a novel mathematical approach to reveal the sources of path integration error. We characterize the different contributors to error across subjects, and study group differences between young and older adults.

Results

Young and older adults experienced a virtual reality environment from a first-person perspective via a head-mounted display (HMD). When participants moved in the real world, their movements were translated into movements (i.e., changes in location and viewing orientation) in the virtual environment, allowing them to walk around within the virtual world and use both body-based and visual self-motion cues to estimate their changing location.

For the path integration task, subjects were asked to track their own position and orientation as they were guided through this environment along pre-defined but unmarked curved paths (**Figure 1**). Each path had four intermediate stopping points, at which participants were asked to stop and report their estimate of the direct distance and direction to the path's starting point.

Most participants showed a characteristic increase in path integration error over the course of their trajectories (**Figure 2A**). We first pooled path integration errors across individuals, separately for the group of young and older adults, and evaluated whether participants' performance in the path integration task was better than random guessing. Indeed, estimates of location were highly correlated with true location (**Figure 2B**; $r = 0.64$ to 0.94 , all $p < 0.00001$) while shuffled responses across trials (corresponding to different trajectories) per stopping point exhibited much larger squared errors (**Figure 2C**).

Dynamical model of errors

Next, we built and fit a temporally-resolved computational model of participants' responses to disentangle different sources of path integration error. Path integration was modeled as a continuous updating of an internal location estimate by an integrator receiving a time-varying velocity estimate. The process was assumed to

be corrupted over time by the following error sources: under- or overestimation of velocity (velocity gain bias), leaky integration of the velocity signal (memory decay or leak), an additive location bias (additive bias, AB), and ongoing zero-mean Gaussian additive noise (accumulating noise, AN), which accumulates and could be interpreted as originating in either the velocity input to the integrator or within the integrator. In our default “full model”, the accumulating noise is naturally interpreted as driven by the velocity input, as it accumulates during the trajectory and in proportion to travel distance, but does not accumulate during stopping points. In an alternative formulation that we tested, the noise accumulates over time instead of travel distance (i.e. it also accumulates during stopping points), and thus would be more naturally interpreted as internal to the integrator (as described in more detail below). In addition, we assumed that the participants’ reports of distance and angle to the starting point are imperfect and corrupted by reporting noise (RN), with angular and radial components.

Model parameters per participant were obtained by the best fit across all paths and trials (**Supplementary Figure S1**). The model captured not only the magnitude of errors averaged across paths (**Figure 3A**), but also predicted with high accuracy the full, time-resolved, signed errors at different portions of the several individual paths (**Figure 3B**).

We then quantified the support for the detailed structure of the full model by comparing it to other variants with fewer parameters or different noise models (i.e., ongoing noise that accumulated over time instead of travel distance, or a noise that remained constant instead of accumulating). We considered reporting noise that was proportional in magnitude to the reported variable, or constant in magnitude, or absent (**Figure 3C**). Model comparisons were carried out using both Bayesian Information Criterion (BIC) and leave-one-out cross-validation (LOOCV), which penalize overly rich models that do not improve prediction performance (see Methods section for more details about different model variants and BIC/LOOCV model comparisons).

The full model was highly favored (“very strong” evidence in support, or $\Delta BIC (\Delta LOOCV) \gg 10$) relative to alternatives, including models with no reporting noise (Full-AN+CN-AB-RN, Full-RN) or non-accumulating (constant) noise (Full-AN+CN-AB-RN, Full-AN+CN), consistently across both age groups (Young: Full vs. Full-AN+CN-AB-RN, $\Delta BIC = 36303$, $\Delta LOOCV = 34743$; Full vs. Full-RN, $\Delta BIC = 27103$, $\Delta LOOCV = 26089$; Full vs. Full-AN+CN: $\Delta BIC = 2035$, $\Delta LOOCV = 2021$; Old: Full vs. Full-AN+CN-AB-RN, $\Delta BIC = 30731$, $\Delta LOOCV = 32124$; Full vs. Full-RN, $\Delta BIC = 22579$, $\Delta LOOCV = 23577$; Full vs. Full-AN+CN: $\Delta BIC = 1963$, $\Delta LOOCV = 1957$). Specifically, the full model outcompeted the Full-AN+CN-AB-RN variant, which – with non-accumulating noise, no additive bias in integration, no reporting noise, but biased and leaky velocity integration – is the closest analogue to a leading existing model of human path integration performance (Lappe et al., 2007, 2011).

The full model was also much better supported than the alternatives when parameters were fit individually for each participant, even after accounting for the much larger number of parameters than fitting a common set of model parameters by age group (**Supplementary Figure S4**; Young: $\Delta BIC = 7412$, Old: $\Delta BIC = 5367$). However, the relative preference for an additive bias in the integrator was

inconclusive, and depended on both the comparison method (BIC and LOOCV) and subject group (**Supplementary Figure S3**).

We next sought to quantify whether accumulating noise in the integrator is better explained by ongoing noise as a function of travel distance or elapsed time. In principle, the former would be a movement-dependent noise that is likely to arise from external velocity inputs to a neural integrator, while the latter is likely to arise within the integrator, due for instance to noise within the grid cell circuit in entorhinal cortex (Hafting et al., 2005; Burak & Fiete, 2009). We therefore compared the full model, which assumes the accumulated noise scales with traveled distance, with the “time model” variant that assumes a scaling with elapsed time, and found much stronger support for the full model across both subject groups (**Figure 4A** and Methods; Young: $\Delta BIC = 194$, $\Delta LOOCV = 222$; Old: $\Delta BIC = 525$, $\Delta LOOCV = 533$).

More directly, we compared total error on trajectories in which subjects stopped versus did not stop at intermediate stopping points. Participants completed 48 paths in total, out of which 18 involved a stop only at the endpoint; in the remaining 30 paths, subjects also stopped at three intermediate stopping points to report the distance and angle to the starting point (see Methods section for more details). Since the different paths had very similar total length (17.7 ± 0.1 meters, mean \pm SD), the total travel distances were similar over stopping and non-stopping trajectories, but the travel times differed substantially (88.7 ± 12.4 seconds versus 35.2 ± 3.9 seconds, mean \pm SD). Nevertheless, path integration errors were very similar for stopping and non-stopping trajectories (**Figure 4B**), indicating that errors were mainly determined by the traveled distance instead of elapsed time, and therefore suggesting that the dominant source of accumulating noise is in the velocity inputs rather than within the integrator.

Moreover, given that paths had similar total lengths, the time-scaling model would predict a negative correlation between walking speed and path integration error: walking faster permits faster completion of the trajectory. However, we found little evidence for such negative correlation in the data (**Figure 4C**).

We next used the full model to assess the relative importance of the different sources of error during the task. To do so, we calculated the relative influence of each bias and noise parameter on the predicted square error (see Methods section for more details). We found the largest influence on total squared error to be from accumulating unbiased noise (50-55%) and the velocity gain bias (25-26%), followed by radial (14-15%) and angular (12-13%) reporting noise (**Figure 3D**). In contrast, the influence of both additive bias and memory leak were very small ($< 3\%$), suggesting that the integrator itself is well-tuned to eliminate leak and internal bias and that the errors are due to velocity misestimation, with contributions from both an unbiased ongoing noise and a biased multiplicative gain in estimating speed.

Note that the result that the largest contribution to the error in the full model is from accumulating noise (**Figure 3D**) does not contradict the result that the introduction of reporting noise causes the largest increase in model fit (**Figure 3C**). Intuitively, **Figure 3C** can be interpreted as a measure of 'error shape', namely how different sources of error grow with traveled distance and distance to the starting point, while **Figure 3D** measures 'error size' in the context of the full model. In models without

reporting noise, all errors have to be fit by a single noise source of incorrect shape, which causes the large discrepancy in **Figure 3C**.

Age-related differences in path integration

Older adults performed less well in the path integration task compared to young adults. Absolute path integration errors were significantly higher in older adults by the first stopping point, and continued to be higher at all subsequent stopping points along the path (**Figure 5A**; stopping point #1: $p = 0.016$; #2: $p = 0.004$; #3: $p = 0.005$; #4: $p = 0.005$). Moreover, incremental path integration errors or the gain in error between adjacent stopping points (pooled over all stopping points; see Methods section for more details) were significantly higher for older relative to young adults (**Figure 5B**; $p = 0.001$).

To determine the underlying reason for the differential performance of older and young adults, we fit our computational model parameters individually across subjects, and then compared the extracted parameters between age groups. Older adults had a significantly larger additive bias ($p = 0.001$), a significantly larger amount of accumulating noise ($p = 0.018$), and greater memory leak ($p = 0.035$) than young adults (**Figure 5C**). However, some of these parameters had only a small overall contribution to the total error; comparing each parameter's contribution to overall path integration error between age groups revealed that only the accumulating noise ($p = 0.012$) had a significantly higher contribution to error in older relative to young adults (**Figure 5D**), suggesting that velocity estimation degrades in relatively unbiased ways, to become noisier in older relative to young humans.

Discussion

We used a novel immersive virtual reality path integration task in which young and older adults tracked their own pose (position and orientation) using visual and body-based motion cues while travelling along sinuous paths. Simultaneously, we developed a powerful analysis approach based on stochastic differential equations (the Langevin equation) to decompose path integration errors into temporally resolved gain, leak, bias, and noise terms and to estimate, on a trial-to-trial basis at different times along the path, how these different sources of error contribute to the location estimation error. In addition to sources of accumulating error, the analysis also included the possibility of errors in the generation of an explicit report of an internal estimate of the displacement vector, as subjects are asked to provide (at each stopping point). We performed mathematical inference of model parameters using an approach based on the Extended Kalman Filter. Disentangling the different sources of error allowed us to compare their influence on path integration errors across participants and between age groups.

A central assumption of our analysis is that observers track and use self-motion cues to continuously update their internal estimates of pose. However, path integration

performance can also rely on a “configural strategy”, in which participants store the configuration of a path (i.e. segment lengths and turn angles) and only compute a homing response when required (Wiener et al., 2011). This strategy is often observed when the outbound path can be easily segmented into turns and distances – such as in the popular triangle completion task – and it can induce systematic biases such as a tendency to regularize turns and distances to canonical values (e.g., isosceles triangles or right-angle turns; Sadalla & Montello, 1989). To eliminate these confounds, we used irregularly shaped sinuous paths, in which translations and rotations were combined into curved trajectories. In addition, we asked subjects to report their internal estimates of the homing vector at intermediate stopping points. These strategies strongly encourage participants to continuously update their displacement estimates based on motion cues over the task.

Previous work by others including Lappe and colleagues (2007, 2011) examined distance misestimation when subjects indicated the magnitude of their displacement along straight or veering outbound paths, reporting that subjects systematically underestimate longer displacements. By contrast, we find that the dominant error in estimating two-dimensional displacement vectors comes from unbiased noise; systematic biases in leak and velocity gain contribute only modestly to total error. There are at least two ways in which the setups differ: First, the studies use different models to decompose error. All observed errors in Lappe et al. were projected onto two terms, the degree of leak in integration and the gain in velocity estimation. Our model included these terms but also allowed for accumulating noise and reporting errors, permitting richer possible interpretations of the possible contributors to the total error. To address whether the richer model is justified by the data, we performed Bayesian model comparison and cross-validation, and showed that the richer model exhibits better performance on unseen data than the simpler one. Thus, it is implausible that subjects exhibit substantial biases in velocity gain that are not discovered by the analysis model. Second, the subjects in Lappe et al. (2007, 2011) have access only to optic flow for motion estimation, while subjects in our study additionally have access to richer, body-based motion cues including vestibular signals. In rodent studies, when motion cues are less rich (passive transport on trolleys; head-fixed animals in virtual environments), displacements are underestimated, similar to the finding in Lappe et al., suggesting that a decreased availability of sensory motion cues in Lappe et al. may account for the dominant contribution of a velocity estimation bias in their findings.

Previous work (e.g., Petzschner & Glasauer, 2011; Petzschner et al., 2015) has considered the possibility that subjects learn and subsequently exploit information about regularities in the tasks they must perform. In a Bayesian interpretation, this information can be incorporated into prior assumptions or biases on the values that variables and parameters can take. Subjects have likely not performed the tasks we designed enough times to form useful priors to improve task performance over naive path integration, and the tasks have little repeatable or regular structure to be exploited. Nevertheless, the ability of our analysis method to isolate different sources of error and their impact on individual path integration performance can make it possible for future studies to investigate the existence or learning of biases, including ones related to a priori assumptions about the structure of the world.

Our discovery that path integration errors in (both young and old) human subjects are mainly explained by an unbiased noise – resulting in a random diffusion of the estimated locations away from their true values – suggests that both velocity estimation and integration are well-tuned to be fairly unbiased processes, i.e. that velocity is estimated with a gain near unity, and that integration is largely non-leaky. The unbiased noise must arise at some stage along the path integration process, and thus could in principle arise within the integrator (Compte et al., 2000; Brody et al., 2003; Boucheny et al., 2005; Wu et al., 2008; Burak & Fiete, 2009, 2012) or in the velocity inputs delivered to it. Given the large body of research suggesting that grid cells may be the neuronal substrates of path integration (Hafting et al., 2005; Fuhs & Touretzky, 2006; McNaughton et al., 2006; Burgess et al., 2007; Guanella et al., 2007; Hasselmo, 2008; Burak & Fiete, 2009; Gil et al., 2018; Stangl et al., 2018) noise within the path integrator might correspond to stochastic processes in the neurons or synapses of the grid cell system. Noise in the velocity input is likely to have a more diffuse origin in the sensing and sensory processing systems that extract velocity estimates from diverse sensory cues across the visual, vestibular, and proprioceptive pathways (Angelaki & Hess, 2005; Angelaki & Cullen, 2008; Laurens & Angelaki, 2011).

To refine our understanding of the source of noise in the integration pathway, we compared the default version of our model, in which the unbiased path integration noise accumulates with the travel distance along a trajectory, with a model variant with time-scaling of this noise. Direct comparisons between these two models showed that internal path integration noise mainly scales with traveled distance rather than elapsed time. This finding suggests that the main part of the accumulating noise in the path integrated location estimate stems not from noise intrinsic to the path integrator, which would tend to accumulate over time regardless of input, but from the sensing or sensory processing systems that compute self-motion estimates, and whose estimates must be noisy. Together with similar findings in non-spatial (Kiani et al., 2013) and spatial evidence accumulation tasks (Pinto et al., 2018), these results suggest an emerging principle in the neurobiology of integrators: that the dominant source of noise in the output of neural integrators is in the inputs rather than within the integrator circuit.

The present work also shows that path integration performance is reduced in older as compared to young adults, in line with previous studies (Mahmood et al., 2009; Adamo et al., 2012; Harris & Wolbers, 2012; Bates & Wolbers, 2014). Further, we were able to determine the dominant sources of error in older subjects and thus determine which of the sources of error already found in young subjects is most magnified as subjects age. Comparing the components of error in young and older adults revealed a significantly higher magnitude of unbiased noise in path integration computations of older adults, while other sources of error were not significantly different between age groups. In other words, the biggest source of error in young adults – accumulating unbiased noise likely arising from imperfect velocity estimation but possibly with some additional contributions of noise internal to the integrator – is further magnified in aging adults, while the smaller sources of error are not significantly compromised with age. Notably, older adults do not appear, at least in our experimental setup, to acquire major additional biases in their speed estimates or become substantially more inaccurate in their reporting of their internal location

estimates. Rather, the most fragile part of the path integration process in younger subjects is most affected with aging.

What could potentially cause increased accumulating noise in older adults? Recent evidence from neuroimaging studies indicates that signatures of grid-cell-like activity in human entorhinal cortex are reduced in old age, and this reduction is associated with larger path integration errors (Stangl et al., 2018). Certain aspects of the vestibular system also deteriorate with age (for a review, see Allen et al., 2016), and vestibular loss affects path integration performance (Glasauer et al., 2002; Xie et al., 2017). These findings support the possibility that an increase in noise on a neuronal level, possibly from a loss of the signal-to-noise ratio from degrading circuitry in the grid cell or the vestibular system may be responsible for deficient computations of pose in old age. Alternatively, deficient processing of proprioceptive or optic flow information in cortical and sub-cortical structures could be responsible for noisy inputs to the path integration circuit in older adults. It therefore remains an important goal for future studies to determine which specific factors underlie increased internal noise in older adults' path integration computations.

Together, we have shown here that path integration error in both young and older adults is mainly caused by accumulating unbiased noise, whereas other error sources contribute only modestly to total error. Moreover, we found that this noise is further magnified in older adults, and therefore accounts for the majority of age-related path integration deficits. Given the importance of path integration computations for spatial navigation, these findings not only advance our understanding of the specific contributors to path integration error, but may also shed light on the mechanisms that underlie navigational decline in old age.

Methods

Participants

62 healthy humans took part in this study. They had no reported history of neurological or psychiatric disease and no reported motor deficits during normal walking or standing. All participants reported right-handedness and had normal or corrected-to-normal eyesight.

Informed consent was obtained from all participants in writing before the measurements, and the experiment received approval from the Ethics Committee of the University of Magdeburg.

Prior to the study, all participants underwent the Montreal Cognitive Assessment (MoCA) screening tool for mild cognitive impairment (Nasreddine et al., 2005). Six older adults who did not exceed a MoCA cut-off score of 23 (following Luis et al., 2009) were excluded from the study and did not participate in any further measurements. Consequently, the data of the remaining 56 participants was used

for data analyses: The group of young adults consisted of 30 participants (15 woman, 15 men) aged between 19 and 26 years ($M = 22.0$, $SD = 2.0$ years), whereas the group of older adults consisted of 26 participants (13 woman, 13 men) aged between 62 and 78 years ($M = 69.0$, $SD = 4.6$ years).

Path integration task

Each participant's path integration performance was measured using a behavioral path integration task, in which they had to track their own position during movement along pre-defined sinuous paths.

In commonly used path integration tasks for humans, such as the triangle completion task (Fujita et al., 1993; Harris & Wolbers, 2012), participants traverse a path and only estimate the distance and direction to the starting location at the end of the path. In the current study, we used a task in which participants were asked at four different points along the path to estimate the distance and direction to the path's starting point (**Figure 1A**). Multiple distance and direction judgments per path were used for three reasons: First, it results in a larger number of data points (i.e., participant responses) in a similar amount of time, enabling a more reliable estimation of path integration errors. Second, it allows us to characterize the accumulation of the path integration error along longer and more complex paths. Third, responses from multiple points along the path can allow for a more precise estimation of path integration errors. Specifically, when complex paths are used, a participant may become disorientated in some trials as they move along the path, and the chances of this occurring increase with the distance traversed. When only one response is collected at the end of the path, as per the traditional triangle completion task, the participant's estimate would be random and not provide a valid quantification of path integration performance. In contrast, our task samples from multiple points along the path meaning that, even if the participant has become disorientated at the path's end point, there are still other data points earlier in the path that provide more accurate estimates of path integration performance.

Prior to the task, participants received written information about the task, and completed several practice paths. Participants donned a head-mounted display (HMD; Oculus Rift Development Kit 2, Oculus VR LLC, www.oculus.com), so that they could not see anything outside the HMD. During the task, participants wore earmuffs in order to prevent them from hearing any background sounds. Furthermore, they were instructed to immediately inform the experimenter if they noticed any external cues that could help them to orient during the task (such as hearing, seeing, feeling or smelling something).

During the task, participants held a wooden stick and were guided by the experimenter along a path (**Figure 1B**). At each of four stopping points along the path, the distance to the starting point had to be estimated verbally in meters and centimeters, and participants turned their body on the spot to indicate the orientation to the starting point, which was measured by the built-in gyrometer of the HMD.

Via the HMD, participants saw a virtual environment, which consisted of a ground plane and distant landmark cues (**Figure 1C**). The ground plane was designed to provide optic flow information during movement, but did not contain any fixed reference points or landmark cues. The distal landmarks were rendered at infinity, so that participants could use them only to determine their heading direction but not their position or any distance information. The exact position of a participant was tracked throughout the task using the Vicon Motion Tracking System with 12 cameras of type T10 (Vicon, Oxford, UK). The participant's viewpoint within the virtual environment was constantly updated depending on their actual position and movement, so that participants could actively walk around in the virtual environment. Consequently, in order to keep track of their own position relative to the path's starting point, participants could use both body-based and visual self-motion cues to perform the path integration task. Specifically, body-based self-motion cues included proprioceptive and vestibular representations, as well as motor efference copies that are produced during movement, whereas visual self-motion cues included optic flow information from the virtual environment and directional information from the environment's distal landmarks (Etienne & Jeffery, 2004).

There were 10 different pre-defined paths (**Figure 1D**). Coordinates for each path were defined as follows: First, a 4-legged path was created that comprised four distances and three turning angles between them. Each distance was either 2, 3.5, 5, or 6.5 meters, and each angle was either 55, 80, or 105 degrees to the left or to the right. Various combinations of distances and angles were used, that fit into a rectangular area of approximately 10 x 8 meters (given by the tracking area and size of the room in which the experiment took place). On the basis of these 4-legged paths, we then created curved paths without corners by using the `cscvn`-function of MATLAB's curve fitting toolbox to calculate a natural interpolated cubic spline curve passing through the turning points of the 4-legged path.

Six paths comprised a mixture of left and right turns, respectively (see **Figure 1D**, path numbers 1 to 6). Two additional paths (path numbers 7 and 9) only comprised right turns or left turns, respectively, and these two paths were present also in their mirrored version (i.e., the path that had only left turns was present also in its mirrored version comprising only right turns, and vice-versa). Directions (left vs. right) of the three turning angles per path were counter-balanced between the different paths.

Critically, the experimenters ensured that participants did not see the real physical dimensions of the testing room and the paths before and during the experiment, by guiding the participants into the room only after they had donned the HMD.

Participants completed the path integration task in three blocks. Within each block, participants performed each of the 10 paths one time and, in addition, they performed the paths 1 to 6 (the ones which had both left and right turns) another time without stopping at the first three stopping points but only at the end of the path (i.e., only at stopping point 4). Consequently, each participant performed 16 paths per block. The order of paths was pseudo-randomized, but the same order was used for all participants. There were always at least three different paths between repeats of the same path. The virtual environment was different in each block (see **Figure 1C**) and the order of environments was randomized across participants.

After the 4th and after the 12th path of each block, participants completed three so-called “standardization paths”, which were needed for data analysis in order to correct each participant’s distance estimate for their ability in verbally reporting distances using meter/centimeter units (see Methods section on “Calculation of path integration errors”). The procedure during a standardization path was similar as during a normal path, but a standardization path had only one start point and one stopping point, which were connected by a straight line, and participants had to estimate the distance between starting and stopping point. Three different distances had to be estimated in the following order: 10 meters, 2 meters, 6 meters. Moreover, there were short breaks in the middle of each block and between blocks. **Figure 1E** gives an overview over the procedure for each block.

After completing the task, participants filled out a form in which they were asked whether they noticed any external cues that could have helped them to orient during the task (such as hearing, seeing, feeling or smelling something), but no participant reported such confounding sources of information. Further, all participants were asked whether they had recognized that some paths were repetitions of each other, but no participant did.

The path integration task was developed using the WorldViz Vizard 5.1 Virtual Reality Software (WorldViz LLC, www.worldviz.com).

Calculation of path integration errors

At every stopping point of a path, participants had to estimate the distance to the path's starting point verbally in meters and centimeters. Converting an internal estimate of location to a verbal estimate is known to be biased (Izard & Dehaene, 2008). Here we assume that the bias is multiplicative. To measure the bias, we ask subjects to walk on straight standardization paths of length 2m, 6m and 10m and to report verbally the distance to the starting point. The correction factor for the bias is then given by:

$$f_{corr} = \frac{d_{correct}}{d_{response}}$$

where $d_{correct}$ is the correct distance of the standardization path (2m, 6m, or 10m, respectively), $d_{response}$ is the responded distance, and f_{corr} is the resulting correction factor. For each participant, this led to three different correction factors, one each for shorter (derived from the 2m standardization path), middle (derived from the 6m standardization path) and longer distances (derived from the 10m standardization path). These factors were used to standardize the distance estimates this participant reported at normal paths: Whenever the participant’s response distance of a normal path was between 0m and 4m, the response was multiplied with the correction factor for shorter distances, whereas response distances between 4m and 8m were multiplied with the correction factor for middle distances, and response distances larger than 8m were multiplied with the correction factor for longer distances.

This standardization procedure was done for each block-half separately, in order to ensure that standardization was performed using an up-to-date correction factor that also accounts for potential temporal changes of a participant's perception of meter/centimeter units that might occur over the course of the experiment: Responses for the first half of each block (1st path to 8th path) were standardized using correction factors from the first set of standardization paths (i.e., carried out after the 4th path of a block), whereas responses for the second half of each block (9th path to 16th path) were standardized using correction factors from the second set of standardization paths (i.e., carried out after the 12th path of a block).

At each stopping point, the responded distance (multiplied with the respective correction factor f_{corr}) and orientation was used to calculate the "presumed starting point". The x and y coordinates of the presumed starting point according to the participant's response were calculated by:

$$x_{presumed} = x_{stop} + d_{standardized} * \cos(ori_{response})$$
$$y_{presumed} = y_{stop} + d_{standardized} * \sin(ori_{response})$$

where $d_{standardized}$ is the standardized response distance, and $ori_{response}$ is the responded orientation. x_{stop} and y_{stop} are coordinates of the stopping point, $x_{presumed}$ and $y_{presumed}$ are the resulting coordinates of the presumed starting point.

To calculate the so-called "absolute" path integration error Err_{abs} , we then calculated the Euclidean distance between the presumed starting point and the path's correct starting point by:

$$Err_{abs} = \sqrt{(x_{presumed} - x_{origin})^2 + (y_{presumed} - y_{origin})^2}$$

where x_{origin} and y_{origin} are the x and y coordinates of the path's correct starting point. According to this method, each absolute path integration error reflects the error that occurred between the path's starting point and the respective stopping point (i.e., at stopping point 1 it reflects the error between the starting point and stopping point 1; at stopping point 2 it reflects the error between the starting point and stopping point 2; and so on). Accumulation of this error measure (i.e., absolute path integration errors) across all available stopping points, however, would lead to an overrepresentation of errors that occurred on early path segments (because these errors would be included for both earlier and later stopping points).

In order to allow for accumulation of path integration errors across stopping points, we therefore also used an alternative method to calculate the so-called "incremental" path integration error Err_{inc} . For a given stopping point, the Euclidean distance between the presumed starting point (according to the participant's response at this respective stopping point) and the previously presumed starting point (according to the response at the previous stopping point) was calculated by:

$$Err_{inc} = \sqrt{(x_{presumed} - x_{prevPresumed})^2 + (y_{presumed} - y_{prevPresumed})^2}$$

where $x_{prevPresumed}$ and $y_{prevPresumed}$ are the x and y coordinates of the previously presumed starting point (according to the response at the previous stopping point). Note that the previously presumed starting point at stopping point 1 is the correct starting point of the path (i.e., $x_{prevPresumed} = x_{origin}$ and $y_{prevPresumed} = y_{origin}$). Consequently, this measure of the path integration error reflects only the incremental error that occurred on the latest path segment before the stopping point, but does not include the error that occurred on earlier segments of the same path. More specifically, at stopping point 1 it reflects the error that occurred between the starting point and stopping point 1, at stopping point 2 it reflects the error that occurred between stopping point 1 and stopping point 2 (not including the error between the starting point and stopping point 1), and so on. This method of calculating the path integration error allows, for each individual participant, to aggregate all error measures from all available stopping points, because each incremental path integration error measure includes only the incremental (i.e., unique) error contribution of one path segment.

Computational modeling

The computational model we use differs from previous models of path integration error (e.g., Lappe et al., 2007, 2011) in several ways: First, we use time-resolved models in which moment-by-moment errors during a trajectory can interact with the moment-by-moment unfolding of the trajectory, and detailed, signed errors can be predicted over time. The richer model allows us to distinguish a large number of sources of noise and bias, and take into account reporting errors in which subjects are only able to report an imperfect representation of their internal location estimates. Unlike previous models that fit path-integration biases using trial-averaged data by minimizing the mean square error (Lappe et al., 2007, 2011), we model both biases and variances using a well-defined log-likelihood. This approach has several advantages: We can fit a more heterogeneous dataset where each trajectory is only repeated a few times, location estimates are weighted inversely proportional to the model-predicted variance (mainly influenced by the traveled distance), making the fit less biased and more data-efficient, and the log-likelihood allows a systematic model-comparison using cross-validation and the Bayesian Information Criterion (BIC).

Throughout this section we use bold-faced letters to refer to two-dimensional vectors. We assume that any participant continuously updates an internal, two-dimensional estimate $\hat{\mathbf{x}}(t)$ of their location $\mathbf{x}(t)$ using an estimated walking velocity $\mathbf{v}(t)$. The update process is compromised by memory decay β , velocity gain α , additive bias \mathbf{b} , and Gaussian noise $\xi(t)$ with standard deviation σ_0 (where $\xi(t)$ is normally-distributed Gaussian noise) according to the following diffusion Langevin equation:

Eq. (1)

$$\frac{d\hat{\mathbf{x}}(t)}{dt} = -\beta\hat{\mathbf{x}}(t) + \alpha\mathbf{v}(t) + \mathbf{b} + \sigma_0\xi(t)$$

The parameters can be interpreted as follows:

- Memory decay or leak β : If $\beta = 0$, then $\hat{\mathbf{x}}(t)$ is the non-forgetful or perfect integral of the right-hand-side of the equation. If $\beta > 0$, then $\hat{\mathbf{x}}(t)$ will have forgotten about inputs $\mathbf{v}(t - \tau)$ with $\tau \gg 1/\beta$, thus the process is referred to as “leaky integration”.
- Multiplicative velocity gain or bias α : A value $\alpha > 1$ corresponds to a systematic overestimation of displacement given velocity $\mathbf{v}(t)$, while a value $\alpha < 1$ corresponds to an underestimate. Correct displacement estimation occurs when $\alpha = 1$.
- Additive bias (AB) \mathbf{b} : Specifies the bias direction along which the location estimate is pulled over time. Zero bias corresponds to $\mathbf{b} = 0$.
- Accumulating noise (AN) that is unbiased and additive with standard deviation σ_0 : This noise can be interpreted to originate from a noisy integrator, a noisy velocity estimate input, or a mixture of both, depending on whether it adds up over time regardless of travel speed or if it scales with speed. Non-noisy velocity estimation and integration occur when $\sigma_0 = 0$.

In our “full model”, we assume that the noise accumulates during displacements and thus grows in proportion to the travel distance. Therefore, the instantaneous value of σ_0 is taken to be proportional to the square root of the instantaneous velocity magnitude (speed) $|\mathbf{v}(t)|$. We consider variants in which this noise instead accumulates with elapsed time, independent of speed (see Methods section on “Model fitting, Extended Kalman Filter (EKF), and Bayesian model comparison” for more details). In a different variant, with constant noise (CN), noise does not accumulate at all but an overall unbiased Gaussian noise term whose total variance by the end of the trajectory does not scale with travel distance or time is added to the model estimate (see below). Within the accumulating noise models, the choice of an accumulating noise that scales with travel distance that we use in the full model, is better supported by our data (see Results section and **Figure 4A**).

Within the full model, we additionally assume that the subjects’ reports of estimated distance and angle to the starting point are corrupted by reporting noise (RN) (Schmidt et al., 1979; Jones et al., 2002; Faisal et al., 2008; Izard & Dehaene, 2008). Given an internally estimated distance d and angle φ , we assume that the reported distances \hat{d} and angles $\hat{\varphi}$ are given by:

Eq. (2)

$$\hat{d} = \exp(\log(d) + \sigma_d \eta_d)$$

$$\hat{\varphi} = \varphi + \sigma_\varphi \eta_\varphi$$

where σ_d and σ_φ are standard deviations of distance and angular noise, η_d is normally-distributed distance noise, and η_φ is normally-distributed angular noise. The parameterization of the distance reporting noise is chosen such that for fixed σ_d , the magnitude of the reporting error $|\hat{d} - d|$ increases approximately linearly with d “proportional or Weber-like reporting noise” (RN), in line with Weber’s law (Oberlin,

1936; Gaydos, 1958; Cornsweet & Teller, 1965; Fechner, 1966; Indow & Stevens, 1966; Izard & Dehaene, 2008). We find empirically that this Weber's law-type parameterization of the distance reporting error captures the data better than a linear parameterization, which we refer to as "constant reporting noise" (CRN) (see Results section and **Figure 3C**).

Model fitting, Extended Kalman Filter (EKF), and Bayesian model comparison

Participants report their location estimates only at stopping points after moving along path segments. Before we can fit our model parameters to those estimates we first need to integrate the stochastic differential equation (1) along segments, a calculation that can be performed analytically because eq. (1) describes an Ornstein-Uhlenbeck process (Uhlenbeck & Ornstein, 1930; Pavliotis, 2014). Assuming that participants walk along a trajectory segment for time t with constant velocity \mathbf{v} , the conditional distribution of the internal location estimate $\hat{\mathbf{x}}_{s+1}$ at the stopping point $s + 1$ given the estimate at the previous stopping point $\hat{\mathbf{x}}_s$ is given by the Gaussian distribution:

$$p(\hat{\mathbf{x}}_{s+1}|\hat{\mathbf{x}}_s) = \mathcal{N}(\hat{\mathbf{x}}_{s+1}|\boldsymbol{\mu}_{s+1}, \sigma_{s+1}^2 I_2)$$

where I_2 is the two-dimensional unity matrix and mean $\boldsymbol{\mu}_{s+1}$ and variance σ_{s+1}^2 are given by:

Eq. (3)

$$\boldsymbol{\mu}_{s+1} = \hat{\mathbf{x}}_s e^{-\beta t} + \frac{\alpha \mathbf{v} + \mathbf{b}}{\beta} (1 - e^{-\beta t})$$

$$\sigma_{s+1}^2 = \frac{\sigma_0^2}{2\beta} (1 - e^{-2\beta t})$$

This update equation for the distribution of internal estimates can also be expressed in terms of the true length $|\Delta \mathbf{x}|$ of the trajectory segment:

Eq. (4)

$$\boldsymbol{\mu}_{s+1} = \hat{\mathbf{x}}_s e^{-\tilde{\beta}|\Delta \mathbf{x}|} + \left(\alpha \frac{\Delta \mathbf{x}}{|\Delta \mathbf{x}|} + \tilde{\mathbf{b}} \right) \frac{1}{\tilde{\beta}} (1 - e^{-\tilde{\beta}|\Delta \mathbf{x}|})$$

$$\sigma_{s+1}^2 = \frac{\tilde{\sigma}_0^2}{2\tilde{\beta}} (1 - e^{-2\tilde{\beta}|\Delta \mathbf{x}|})$$

where we have rescaled three of the original parameters by the magnitude of the walking velocity $|\mathbf{v}|$:

Eq. (5)

$$\tilde{\beta} = \frac{\beta}{|\mathbf{v}|} \quad \tilde{\mathbf{b}} = \frac{\mathbf{b}}{|\mathbf{v}|} \quad \tilde{\sigma}_0^2 = \frac{\sigma_0^2}{|\mathbf{v}|}$$

Equations (3) and (4) are equivalent if the walking velocity $|\mathbf{v}|$ is truly constant across trajectory segments and trials. If the walking velocity does vary, holding the transformed parameters (5) fixed assumes that the path integration error of the internal location estimate mainly depends on the traveled *distance*, whereas the original model (3) assumes that the path integration error mainly depends on the elapsed walking *time*. In what follows, we will choose the *distance* model and hold the transformed parameters (5) fixed, in line with previous modeling of human path integration (Lappe et al., 2007, 2011). We also explicitly test that the *distance* model is better supported by the data than the *time* model (see Results section and **Figure 4A**).

Fitting model without reporting noise (Full-RN)

Here we explain how the parameters $\theta = (\tilde{\beta}, \alpha, \tilde{\mathbf{b}}, \tilde{\sigma}_0^2)$ related to integration and $\kappa = (\sigma_d^2, \sigma_\phi^2)$ related to reporting were fit to participants' performance by maximizing the likelihood. For simplicity, consider first a model without the reporting noise parameters κ . In this case the internal location estimate $\hat{\mathbf{x}}_s$ can be directly expressed in terms of participants' report of the distance \hat{d} and angle $\hat{\phi}$ to the starting point $\mathbf{x}_{\text{start}}$ of the current walking trajectory:

Eq. (6)

$$\hat{\mathbf{x}}_s = \begin{pmatrix} \hat{d} \cos(\hat{\phi}) \\ \hat{d} \sin(\hat{\phi}) \end{pmatrix} + \mathbf{x}_{\text{start}}$$

Without loss of generality we will set $\mathbf{x}_{\text{start}} = 0$. The log-likelihood of the data averaged over trials is given by:

Eq. (7)

$$LL_{\kappa=0}(\theta) = \left\langle \sum_{s=0}^3 \log p(\hat{\mathbf{x}}_{s+1} | \hat{\mathbf{x}}_s; \theta) \right\rangle_{\text{trials}} = \left\langle \sum_{s=0}^3 \log \mathcal{N}(\hat{\mathbf{x}}_{s+1} | \boldsymbol{\mu}_{s+1}(\hat{\mathbf{x}}_s, \theta), \sigma_{s+1}^2(\theta)) \right\rangle_{\text{trials}}$$

where $\boldsymbol{\mu}_{s+1}(\hat{\mathbf{x}}_s, \theta)$ and $\sigma_{s+1}^2(\theta)$ are given by the expressions in eq. (4). We then fit θ to the data by maximizing the log-likelihood numerically:

Eq. (8)

$$\theta_{\text{ML}} = \text{argmax}_{\theta} LL_{\kappa=0}(\theta)$$

Fitting model with reporting noise: the full model (Full)

With reporting noise, the expression for the log-likelihood as a function of $\Theta = (\theta, \kappa)$ is more involved, since the relationship between the reported estimates \hat{d} and $\hat{\varphi}$ and the internal location estimate $\hat{\mathbf{x}}_s$ is both stochastic and non-linear. We can nevertheless make progress by rephrasing the problem in terms of the well-studied Extended Kalman Filter (EKF), a framework that permits calculation of the log-likelihood by locally linearizing the non-linearities (Thrun et al., 2005). The EKF framework encompasses a stochastic state transition of a hidden variable $\hat{\mathbf{x}}_s$ whose distribution can be inferred using a noisy observation \mathbf{z}_s :

Eq. (9)

$$\hat{\mathbf{x}}_{s+1} = f(\hat{\mathbf{x}}_s) + \Sigma_x^{\frac{1}{2}} \boldsymbol{\xi}_x$$

$$\mathbf{z}_{s+1} = h(\hat{\mathbf{x}}_{s+1}) + \Sigma_z^{\frac{1}{2}} \boldsymbol{\xi}_z$$

where f and h are arbitrary non-linear functions and Σ_x and Σ_z are covariance matrices of Gaussian-distributed noise. In our case the state transition is linear in $\hat{\mathbf{x}}_s$ and is given as before by eq. (4):

Eq. (10)

$$f(\hat{\mathbf{x}}_s) = \boldsymbol{\mu}_{s+1}(\hat{\mathbf{x}}_s) = \hat{\mathbf{x}}_s e^{-\tilde{\beta}|\Delta\mathbf{x}|} + \left(\alpha \frac{\Delta\mathbf{x}}{|\Delta\mathbf{x}|} + \tilde{\mathbf{b}} \right) \frac{1}{\tilde{\beta}} (1 - e^{-\tilde{\beta}|\Delta\mathbf{x}|})$$

$$\Sigma_x = \sigma_{s+1}^2 \mathbf{I}_2 = \frac{\tilde{\sigma}_0^2}{2\tilde{\beta}} (1 - e^{-2\tilde{\beta}|\Delta\mathbf{x}|}) \mathbf{I}_2$$

To derive the non-linear observation function we need to find a coordinate transformation such that in the transformed frame the noise is added linearly. According to eq. (2), the noise is added linearly in log-polar coordinates. The observation function $h(\hat{\mathbf{x}}_{s+1})$ therefore corresponds to the transformation from cartesian to log-polar coordinates:

Eq. (11)

$$h(\hat{\mathbf{x}}_{s+1}) = \begin{pmatrix} d(\hat{\mathbf{x}}_{s+1}) \\ \varphi(\hat{\mathbf{x}}_{s+1}) \end{pmatrix} = \begin{pmatrix} \log|\hat{\mathbf{x}}_{s+1}| \\ \text{atan2}((\hat{\mathbf{x}}_{s+1})_2, (\hat{\mathbf{x}}_{s+1})_1) \end{pmatrix}$$

$$\Sigma_z = \begin{pmatrix} \sigma_d^2 & 0 \\ 0 & \sigma_\varphi^2 \end{pmatrix}$$

and the observation \mathbf{z}_{s+1} is related to the reports \hat{d} and $\hat{\varphi}$ by:

Eq. (12)

$$\mathbf{z}_{s+1} = \begin{pmatrix} \log \hat{d} \\ \hat{\varphi} \end{pmatrix}$$

The EKF framework permits the calculation of two important distributions using Gaussian approximations: the posterior distribution of the hidden variable $\hat{\mathbf{x}}_{s+1}$ given the observations \mathbf{z}_1 to \mathbf{z}_s (predictive distribution), and the posterior distribution of $\hat{\mathbf{x}}_{s+1}$ given \mathbf{z}_1 to \mathbf{z}_{s+1} (updated distribution). We denote the mean and covariance of these posterior distributions as

Eq. (13)

$$p(\hat{\mathbf{x}}_{s+1} | \mathbf{z}_1, \dots, \mathbf{z}_s) = \mathcal{N}(\hat{\mathbf{x}}_{s+1} | \boldsymbol{\mu}_{s+1|s}, P_{s+1|s}) \quad (\text{predictive distribution})$$

$$p(\hat{\mathbf{x}}_{s+1} | \mathbf{z}_1, \dots, \mathbf{z}_{s+1}) = \mathcal{N}(\hat{\mathbf{x}}_{s+1} | \boldsymbol{\mu}_{s+1|s+1}, P_{s+1|s+1}) \quad (\text{updated distribution})$$

Mean and covariance of both distributions can be calculated recursively over stopping points using the standard EKF update equations (Thrun et al., 2005):

Eq. (14)

$$\boldsymbol{\mu}_{s+1|s} = f(\boldsymbol{\mu}_{s|s})$$

$$P_{s+1|s} = F_{s+1} P_{s|s} F_{s+1}^T + \Sigma_x$$

$$S_{s+1} = H_{s+1} P_{s+1|s} H_{s+1}^T + \Sigma_z$$

$$K_{s+1} = P_{s+1|s} H_{s+1}^T S_{s+1}^{-1}$$

$$\boldsymbol{\mu}_{s+1|s+1} = \boldsymbol{\mu}_{s+1|s} + K_{s+1}(\mathbf{z}_{s+1} - h(\boldsymbol{\mu}_{s+1|s}))$$

$$P_{s+1|s+1} = (I_2 - K_{s+1} H_{s+1}) P_{s+1|s}$$

where the matrices F_{s+1} and H_{s+1} are the Jacobian matrices of transition and observation function evaluated at the previous updated mean $\boldsymbol{\mu}_{s|s}$ and predictive mean $\boldsymbol{\mu}_{s+1|s}$ respectively:

Eq. (15)

$$F_{s+1} = \left. \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\boldsymbol{\mu}_{s|s}} = e^{-\tilde{\beta}|\Delta \mathbf{x}|} I_2$$

$$H_{s+1} = \left. \frac{\partial h(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\boldsymbol{\mu}_{s+1|s}} = \frac{1}{|\boldsymbol{\mu}_{s+1|s}|^2} \begin{pmatrix} \mu_{s+1|s,1} & \mu_{s+1|s,2} \\ -\mu_{s+1|s,2} & \mu_{s+1|s,1} \end{pmatrix}$$

At the starting point ($s = 0$), we initialize $\boldsymbol{\mu}_{s=0|s=0} = \mathbf{x}_{\text{start}} = 0$ and $P_{s=0|s=0} = 0$. Next, we calculate the predicted distribution of the next measurement \mathbf{z}_{s+1} given the previous measurements \mathbf{z}_1 to \mathbf{z}_s by integrating out the internal estimate $\hat{\mathbf{x}}_{s+1}$:

Eq. (16)

$$p(\mathbf{z}_{s+1} | \mathbf{z}_1, \dots, \mathbf{z}_s) = \int d\hat{\mathbf{x}}_{s+1} p(\mathbf{z}_{s+1} | \hat{\mathbf{x}}_{s+1}) p(\hat{\mathbf{x}}_{s+1} | \mathbf{z}_1, \dots, \mathbf{z}_s)$$

$$\begin{aligned}
 &= \int d\hat{\mathbf{x}}_{s+1} \mathcal{N}(\mathbf{z}_{s+1}|h(\hat{\mathbf{x}}_{s+1}), \Sigma_z) \mathcal{N}(\hat{\mathbf{x}}_{s+1}|\boldsymbol{\mu}_{s+1|s}, P_{s+1|s}) \\
 &\approx \int d\hat{\mathbf{x}}_{s+1} \mathcal{N}(\mathbf{z}_{s+1}|h(\boldsymbol{\mu}_{s+1|s}) + H_{s+1}(\hat{\mathbf{x}}_{s+1} - \boldsymbol{\mu}_{s+1|s}), \Sigma_z) \mathcal{N}(\hat{\mathbf{x}}_{s+1}|\boldsymbol{\mu}_{s+1|s}, P_{s+1|s}) \\
 &= \mathcal{N}(\mathbf{z}_{s+1}|h(\boldsymbol{\mu}_{s+1|s}), S_{s+1})
 \end{aligned}$$

where we have used the linearization approximation of the EKF at the 3rd line. This allows us to express the full log-likelihood as:

Eq. (17)

$$\text{LL}(\boldsymbol{\theta}) = \left\langle \sum_{s=0}^3 \log p(\mathbf{z}_{s+1}|\mathbf{z}_1, \dots, \mathbf{z}_s; \boldsymbol{\theta}) \right\rangle_{\text{trials}}$$

where the dependency on the parameters $\boldsymbol{\theta}$ is introduced through f , its Jacobian F_{s+1} , Σ_x and Σ_z . In analogy to (8), we find the maximum likelihood (ML) estimate for $\boldsymbol{\theta}$ by numerically maximizing the log-likelihood:

$$\boldsymbol{\theta}_{\text{ML}} = \text{argmax}_{\boldsymbol{\theta}} \text{LL}(\boldsymbol{\theta})$$

Numerical parameter optimization was performed using the `fminunc`-function of MATLAB's optimization toolbox.

Incorporating trials without participant responses at intermediate stopping points

For a fraction of the trials, a response is not collected at intermediate stopping points, but only at the end of the trajectory. For these trials the observations \mathbf{z}_{s+1} are missing for $s \in \{0,1,2\}$ and therefore the EKF update equations (14) need to be adapted. This can be achieved using the infinite observation noise limit $\Sigma_z \rightarrow \infty$, under which the predicted and updated posterior distributions become identical:

$$\begin{aligned}
 \boldsymbol{\mu}_{s+1|s} &= f(\boldsymbol{\mu}_{s|s}) \\
 P_{s+1|s} &= F_{s+1} P_{s|s} F_{s+1}^T + \Sigma_x \\
 \boldsymbol{\mu}_{s+1|s+1} &= \boldsymbol{\mu}_{s+1|s} \\
 P_{s+1|s+1} &= P_{s+1|s}
 \end{aligned}$$

For $s = 3$, the observation at the last stopping point \mathbf{z}_{s+1} is defined, and eq. (14) can be used as usual.

Model predictions

We simulated participants' responses by sampling 100 repetitions of model trajectories for each participant and trial from eq. (9) given the fitted parameters $\Theta = \Theta_{ML}$ and the trajectory parameters Δx for each segment. Each repetition generates stochastic observations \hat{d}_{model} and $\hat{\phi}_{model}$ via eq. (12) that can be analyzed analogously to the actual data. The model prediction for the square error is calculated by averaging the square error of the simulated data over trials and repetitions. The model prediction for the bias on individual trials is calculated by averaging the simulated data over repetitions.

Model variants

Full model without additive bias, no reporting noise (Full-AB-RN): The non-zero parameters in this model are memory decay $\tilde{\beta}$, multiplicative velocity gain α and noise $\tilde{\sigma}_0^2$. The additive bias (AB) $\tilde{\mathbf{b}}$ and reporting noise (RN) parameters $\kappa = (\sigma_d^2, \sigma_\phi^2)$ are set to zero. The log-likelihood is computed using eq. (7) instead of eq. (17).

Full model, no reporting noise (Full-RN): This model has non-zero parameters $\theta = (\tilde{\beta}, \alpha, \tilde{\mathbf{b}}, \tilde{\sigma}_0^2)$ but the reporting noise parameters $\kappa = (\sigma_d^2, \sigma_\phi^2)$ are set to zero. The log-likelihood is computed using eq. (7) instead of eq. (17).

Non-accumulating noise, no reporting error (Full-AN+CN-AB-RN, Full-AN+CN-RN): These models assume that the total amount of noise is independent of distance, time, and stopping points, and the reporting noise parameters $\kappa = (\sigma_d^2, \sigma_\phi^2)$ are set to zero. The fitting procedure for the non-noise (bias) parameters is equivalent to minimizing the square error in predicting the mean location estimates averaged over trials with similarly shaped trajectories. We replace Eq. (4) by:

$$\begin{aligned}\boldsymbol{\mu}_{s+1} &= \boldsymbol{\mu}_s e^{-\tilde{\beta}|\Delta x|} + \left(\alpha \frac{\Delta \mathbf{x}}{|\Delta \mathbf{x}|} + \tilde{\mathbf{b}} \right) \frac{1}{\tilde{\beta}} (1 - e^{-\tilde{\beta}|\Delta x|}) \\ \sigma_{s+1}^2 &= \tilde{\sigma}_0^2\end{aligned}$$

Note that $\boldsymbol{\mu}_{s+1}$ depends on the previous predicted mean $\boldsymbol{\mu}_s$ instead of the measured internal estimate $\hat{\mathbf{x}}_s$ as in eq. (4). Correspondingly the conditional distribution of each internal location estimate does not depend on the estimate at the previous stopping point, so that $p(\hat{\mathbf{x}}_{s+1} | \hat{\mathbf{x}}_{s+1}) = p(\hat{\mathbf{x}}_s)$. Maximizing the log-likelihood in eq. (7) corresponds to uniformly minimizing the square error across stopping points:

$$\begin{aligned}LL_{\kappa=0}(\theta) &= \left\langle \sum_{s=0}^3 \log p(\hat{\mathbf{x}}_{s+1}; \theta) \right\rangle_{trials} = \left\langle \sum_{s=0}^3 \log \mathcal{N}(\hat{\mathbf{x}}_{s+1} | \boldsymbol{\mu}_{s+1}(\theta), \sigma_{s+1}^2(\theta)) \right\rangle_{trials} \\ &= \left\langle - \sum_{s=0}^3 \frac{1}{2\tilde{\sigma}_0^2} (\hat{\mathbf{x}}_{s+1} - \boldsymbol{\mu}_{s+1}(\theta))^2 - \frac{1}{2} \log(2\pi\tilde{\sigma}_0^2) \right\rangle_{trials}\end{aligned}$$

We fit two versions of the constant or non-accumulating noise model, one without any additive bias ($\tilde{\mathbf{b}} = 0$; Full-AN+CN-AB-RN), and one with an additive bias ($\tilde{\mathbf{b}}$ can vary; Full-AN+CN-RN). The model without additive bias (Full-AN+CN-AB-RN) is the closest match to the model proposed in Lappe et al. (2007, 2011), as it contains leak and bias.

Non-accumulating noise with reporting noise (Full-AN+CN): As above, this variant assumes that the unbiased noise is independent rather than accumulating over time or distance, but does include reporting noise with non-zero reporting noise parameters $\kappa = (\sigma_d^2, \sigma_\phi^2)$, with Weber-like structure in which the reporting noise is proportional to the magnitude of the reported variable. This model can be fit using a few adjustments from the full model.

As there is no accumulating noise that induces correlations across stopping points, observations $\mathbf{z}_1, \dots, \mathbf{z}_s$ at previous stopping points are uninformative for the next location estimate $\hat{\mathbf{x}}_{s+1}$, and both predictive and updated distribution in eq. (13) are equal to the prior distribution:

$$p(\hat{\mathbf{x}}_{s+1} | \mathbf{z}_1, \dots, \mathbf{z}_s) = p(\hat{\mathbf{x}}_{s+1} | \mathbf{z}_1, \dots, \mathbf{z}_{s+1}) = p(\hat{\mathbf{x}}_{s+1}) = \mathcal{N}(\hat{\mathbf{x}}_{s+1} | \boldsymbol{\mu}_{s+1}, \Sigma_x)$$

Consequently, there is no need to distinguish between predictive and updated mean and variance parameters. Instead, eq. (14) is replaced by:

$$\boldsymbol{\mu}_{s+1} = f(\boldsymbol{\mu}_s) = \boldsymbol{\mu}_s e^{-\tilde{\beta}|\Delta\mathbf{x}|} + \left(\alpha \frac{\Delta\mathbf{x}}{|\Delta\mathbf{x}|} + \tilde{\mathbf{b}} \right) \frac{1}{\tilde{\beta}} (1 - e^{-\tilde{\beta}|\Delta\mathbf{x}|})$$

$$S_{s+1} = H_{s+1} \Sigma_x H_{s+1}^T + \Sigma_z$$

where

$$\Sigma_x = \tilde{\sigma}_0^2 I_2$$

$$\Sigma_z = \begin{pmatrix} \sigma_d^2 & 0 \\ 0 & \sigma_\phi^2 \end{pmatrix}$$

$$H_{s+1} = \left. \frac{\partial h(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\boldsymbol{\mu}_{s+1}} = \frac{1}{|\boldsymbol{\mu}_{s+1}|^2} \begin{pmatrix} \mu_{s+1,1} & \mu_{s+1,2} \\ -\mu_{s+1,2} & \mu_{s+1,1} \end{pmatrix}$$

The log-likelihood is approximated as

$$LL(\Theta) = \left\langle \sum_{s=0}^3 \log p(\mathbf{z}_{s+1}; \Theta) \right\rangle_{trials} = \left\langle \sum_{s=0}^3 \log \mathcal{N}(\mathbf{z}_{s+1} | h(\boldsymbol{\mu}_{s+1}), S_{s+1}) \right\rangle_{trials}$$

Model with constant reporting noise (Full-RN+CRN): This model is the same as the full model (eq. (10)), except that the reporting error is drawn from a distribution of constant size, instead of being Weber-like (proportional to the reported quantity). Thus, eq. (2) is replaced by:

Eq. (18)

$$\hat{d} = d + \sigma_d \eta_d$$

$$\hat{\varphi} = \varphi + \sigma_\varphi \eta_\varphi$$

The model can be fit in the same way as the full model, when reporting noise is proportional to the internal estimate, except that noise is added linearly in polar coordinates instead of log-polar coordinates. Specifically, the first component of the observation \mathbf{z}_s defined as the reported distance \hat{d} instead of its logarithm $\log(\hat{d})$, so that eq. (12) is replaced by:

Eq. (19)

$$\mathbf{z}_{s+1} = \begin{pmatrix} \hat{d} \\ \hat{\varphi} \end{pmatrix}$$

and we replace the observation function $h(\hat{\mathbf{x}}_{s+1})$ in eq. (11) by the transformation from cartesian to polar coordinates:

Eq. (20)

$$h(\hat{\mathbf{x}}_{s+1}) = \begin{pmatrix} d(\hat{\mathbf{x}}_{s+1}) \\ \varphi(\hat{\mathbf{x}}_{s+1}) \end{pmatrix} = \begin{pmatrix} |\hat{\mathbf{x}}_{s+1}| \\ \text{atan2}((\hat{\mathbf{x}}_{s+1})_2, (\hat{\mathbf{x}}_{s+1})_1) \end{pmatrix}$$

and the Jacobian H_{s+1} in eq. (15) by:

$$H_{s+1} = \left. \frac{\partial h(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\boldsymbol{\mu}_{s+1|s}} = \frac{1}{|\boldsymbol{\mu}_{s+1|s}|^2} \begin{pmatrix} \mu_{s+1|s,1} |\boldsymbol{\mu}_{s+1|s}| & \mu_{s+1|s,2} |\boldsymbol{\mu}_{s+1|s}| \\ -\mu_{s+1|s,2} & \mu_{s+1|s,1} \end{pmatrix}$$

The rest of the calculation of the log-likelihood function is exactly the same as for the full model.

Fitting by age group: For this analysis, instead of fitting model parameters individually for each participant, participants in each age group are constrained to have the same model parameters.

Full model with time accumulation (ongoing noise is proportional to elapsed time rather than displacement; same reporting noise model as for the full model): This model assumes that the mean and variance of the internal location estimate is determined by the elapsed time of each trajectory segment, eq. (3), instead of the distance of each trajectory segment, eq. (4). In the case of zero leak, the time model predicts that the variance of the internal location estimate increases proportionally to elapsed time instead of traveled distance.

To fit the time model we replace eq. (10) by:

Eq. (21)

$$f(\hat{\mathbf{x}}_s) = \boldsymbol{\mu}_{s+1}(\hat{\mathbf{x}}_s) = \hat{\mathbf{x}}_s e^{-\beta \Delta t} + \left(\alpha \frac{\Delta \mathbf{x}}{\Delta t} + \mathbf{b} \right) \frac{1}{\beta} (1 - e^{-\beta \Delta t})$$

$$\Sigma_{\mathbf{x}} = \sigma_{s+1}^2 I_2 = \frac{\sigma_0^2}{2\beta} (1 - e^{-2\beta\Delta t}) I_2$$

where Δt is the elapsed time of each trajectory segment. In addition, the Jacobian of the transition function F_{s+1} in eq. (15) is replaced by:

$$F_{s+1} = \left. \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\boldsymbol{\mu}_{s|s}} = e^{-\beta\Delta t} I_2$$

For trials without intermediate stopping points, only the total elapsed time of the trajectory, but not the elapsed time Δt of individual segments was recorded. For these trials we estimated Δt by linear interpolation using the traveled distance $|\Delta \mathbf{x}|$ and assuming a constant walking speed.

The observation function $h(\hat{\mathbf{x}}_{s+1})$ and its Jacobian H_{s+1} are identical to the standard Weber reporting noise model as specified in eq. (11) and eq. (15) respectively.

Model comparison using Bayesian Information Criterion and leave-one-out cross-validation

The Bayesian Information Criterion (BIC) is a scheme to compare models with different numbers of parameters: Models with lower BIC are preferred over models with higher BIC, and large BIC differences between models ($\Delta \text{BIC} \gg 10$) can be interpreted as “very strong” evidence against the model with lower BIC (Raftery, 1995; Konishi & Kitagawa, 2008). The BIC corrects for the higher expressibility of models with larger number of parameters using an additive compensation term. The formula for the BIC is given by:

$$\text{BIC} = -2 \text{LL}(\Theta_{ML}) + \log(n)k$$

where n is the number of observations and k is the number of parameters. The number of parameters for different models is given by:

Model	Number of parameters k
Full model, individually fit, young adults	7 x n = 210
Full model, individually fit, older adults	7 x n = 182
Full model, group-level fit	7
Full-AB-RN, Full-AN+CN-AB-RN, individually fit, young adults	3 x n = 90
Full-AB-RN, Full-AN+CN-AB-RN, individually fit, older adults	3 x n = 78
Full-RN, Full-AN+CN-RN, individually fit, young adults	5 x n = 150
Full-RN, Full-AN+CN-RN, individually fit, older adults	5 x n = 130
Full-AN+CN, individually fit, older adults	7 x n = 210
Full-AN+CN, individually fit, young adults	7 x n = 182
Full-RN+CRN, individually fit, young adults	7 x n = 210
Full-RN+CRN, individually fit, older adults	7 x n = 182
Full model with time accumulation, individually fit, young adults	7 x n = 210
Full model with time accumulation, older adults	7 x n = 182

In addition to BIC, we compare models using leave-one-out cross-validation (LOOCV). Given T trajectories for each model and subject, we train the model parameters on a training dataset of $T - 1$ trajectories, evaluate it on the held-out test trajectory and average the result over the T distinct training-test splits. To allow numerical comparison with BIC we use as evaluation measure twice the negative log-likelihood:

$$\text{LOOCV} = -\frac{2}{T} \sum_{k=1}^T \text{LL}_k(\Theta_{\text{ML}}^{-k})$$

where LL_k is the log-likelihood corresponding to the k -th trajectory, and Θ_{ML}^{-k} are the ML parameters on the training set excluding the k -th trajectory.

Relative influence of model parameters on predicted square error

The detailed computational model allows us to measure the influence of each type of bias and noise parameter on the square error predicted by the model. For each parameter type we calculated a reduced square error that is generated by setting this parameter type to its ideal value corresponding to unbiased, noiseless integration, while keeping the remaining parameters at their ML estimates:

$$\text{error}_{\tilde{\beta}}^2 = \text{error}^2(\tilde{\beta} = 0, \alpha_{\text{ML}}, \tilde{\mathbf{b}}_{\text{ML}}, \tilde{\sigma}_0^2_{\text{ML}}, \sigma_d^2_{\text{ML}}, \sigma_\phi^2_{\text{ML}})$$

$$\text{error}_{\tilde{\alpha}}^2 = \text{error}^2(\tilde{\beta}_{\text{ML}}, \alpha = 1, \tilde{\mathbf{b}}_{\text{ML}}, \tilde{\sigma}_0^2_{\text{ML}}, \sigma_d^2_{\text{ML}}, \sigma_\phi^2_{\text{ML}})$$

$$\text{error}_{\tilde{\mathbf{b}}}^2 = \text{error}^2(\tilde{\beta}_{\text{ML}}, \alpha_{\text{ML}}, \tilde{\mathbf{b}} = 0, \tilde{\sigma}_0^2_{\text{ML}}, \sigma_d^2_{\text{ML}}, \sigma_\phi^2_{\text{ML}})$$

$$\text{error}_{\tilde{\sigma}_0^2}^2 = \text{error}^2(\tilde{\beta}_{\text{ML}}, \alpha_{\text{ML}}, \tilde{\mathbf{b}}_{\text{ML}}, \tilde{\sigma}_0^2 = 0, \sigma_d^2_{\text{ML}}, \sigma_\phi^2_{\text{ML}})$$

$$\text{error}_{\sigma_d^2}^2 = \text{error}^2(\tilde{\beta}_{\text{ML}}, \alpha_{\text{ML}}, \tilde{\mathbf{b}}_{\text{ML}}, \tilde{\sigma}_0^2_{\text{ML}}, \sigma_d^2 = 0, \sigma_\phi^2_{\text{ML}})$$

$$\text{error}_{\sigma_\phi^2}^2 = \text{error}^2(\tilde{\beta}_{\text{ML}}, \alpha_{\text{ML}}, \tilde{\mathbf{b}}_{\text{ML}}, \tilde{\sigma}_0^2_{\text{ML}}, \sigma_d^2_{\text{ML}}, \sigma_\phi^2 = 0)$$

The relative influence of each reduced error in percent is then calculated as:

$$\text{infl}_i = 100 \frac{\text{error}^2(\Theta_{\text{ML}}) - \text{error}_i^2}{\text{error}^2(\Theta_{\text{ML}})}$$

Note that the relative influence can be negative if the reduced square error is larger than the square error of the full model. This can be true in particular for the memory leak parameter $\tilde{\beta}$: For example, a memory leak value $\tilde{\beta}_{\text{ML}} < 1$ that draws location estimates towards the starting point can partly compensate for a velocity bias $\alpha_{\text{ML}} > 1$ that draws location estimates away from the starting point. Setting $\tilde{\beta} = 1$ when

$\alpha_{ML} > 1$ can therefore lead to a larger “reduced” square error and a negative relative influence.

Also note that due to the nonlinearity of the model, the relative influences do not have to sum to 100%.

Author contributions

MS, IK, IF and TW conceptualized the work. MS and MR programmed and implemented the virtual reality path integration task. MS acquired the data. IK and IF developed the computational model for dissecting path integration errors. MS and IK analyzed the data, visualized the results, and drafted the manuscript. All authors edited the manuscript. IF and TW supervised the work.

Acknowledgements

MS, MR, and TW express their gratitude to Falko Eckardt, Mareen Hanelt, Patrick Hauff, Anita Hökelmann, Marko Kirbach, Claudia Marx, Swantje Petersen, Mona Reißberg, and Uwe Sobieray for their help with data acquisition, project administration, and technical assistance. IK and IF thank members of the Fiete group for helpful discussions and comments.

This work was supported by a Collaborative Research in Computational Neuroscience Grant (01GQ1303) from the National Science Foundation (NSF) and the German Ministry of Education and Research (BMBF) to TW and IF; by the Simons Foundation through a SCGB grant to IF; and by the European Research Council Starting Investigator Grant AGESPACE (335090) to TW. IF is an HHMI Faculty Scholar and a CIFAR Senior Fellow.

References

- Adamo, D. E., Briceño, E. M., Sindone, J. A., Alexander, N. B., & Moffat, S. D. (2012). Age differences in virtual environment and real world path integration. *Frontiers in Aging Neuroscience*, 4, 26.
- Allen, D., Ribeiro, L., Arshad, Q., & Seemungal, B. M. (2016). Age-Related Vestibular Loss: Current Understanding and Future Research Directions. *Frontiers in Neurology*, 7, 231.
- Allen, G. L., Kirasic, K. C., Rashotte, M. A., & Haun, D. B. M. (2004). Aging and path integration skill: Kinesthetic and vestibular contributions to wayfinding. *Perception & Psychophysics*, 66, 170–179.
- Angelaki, D. E., & Cullen, K. E. (2008). Vestibular System: The Many Facets of a Multimodal Sense. *Annual Review of Neuroscience*, 31, 125–150.
- Angelaki, D. E., & Hess, B. J. M. (2005). Self-motion-induced eye movements: effects on visual acuity and navigation. *Nature Reviews Neuroscience*, 6, 966–976.

- Bassett, J. P., & Taube, J. S. (2001). Neural correlates for angular head velocity in the rat dorsal tegmental nucleus. *The Journal of Neuroscience*, *21*, 5740–5751.
- Bates, S. L., & Wolbers, T. (2014). How cognitive aging affects multisensory integration of navigational cues. *Neurobiology of Aging*, *35*, 2761–2769.
- Biazoli, C. E., Goto, M., Campos, A. M. P., & Canteras, N. S. (2006). The supragenual nucleus: A putative relay station for ascending vestibular signs to head direction cells. *Brain Research*, *1094*, 138–148.
- Boucheny, C., Brunel, N., & Arleo, A. (2005). A continuous attractor network model without recurrent excitation: Maintenance and integration in the head direction cell system. *Journal of Computational Neuroscience*, *18*, 205–227.
- Britten, K. H. (2008). Mechanisms of Self-Motion Perception. *Annual Review of Neuroscience*, *31*, 389–410.
- Brody, C. D., Romo, R., & Kepecs, A. (2003). Basic mechanisms for graded persistent activity: Discrete attractors, continuous attractors, and dynamic representations. *Current Opinion in Neurobiology*, *13*, 204–211.
- Brunton, B. W., Botvinick, M. M., & Brody, C. D. (2013). Rats and Humans Can Optimally Accumulate Evidence for Decision-Making. *Science*, *340*, 95–98.
- Burak, Y., & Fiete, I. R. (2009). Accurate Path Integration in Continuous Attractor Network Models of Grid Cells. *PLoS Computational Biology*, *5*, e1000291.
- Burak, Y., & Fiete, I. R. (2012). Fundamental limits on persistent activity in networks of noisy neurons. *Proceedings of the National Academy of Sciences*, *109*, 17645–17650.
- Burgess, N., Barry, C., & O'Keefe, J. (2007). An oscillatory interference model of grid cell firing. *Hippocampus*, *17*, 801–812.
- Butler, W. N., & Taube, J. S. (2015). The nucleus prepositus hypoglossi contributes to head direction cell stability in rats. *The Journal of Neuroscience*, *35*, 2547–2558.
- Chrastil, E. R., Sherrill, K. R., Aselcioglu, I., Hasselmo, M. E., & Stern, C. E. (2017). Individual Differences in Human Path Integration Abilities Correlate with Gray Matter Volume in Retrosplenial Cortex, Hippocampus, and Medial Prefrontal Cortex. *eNeuro*, *4*, 0346-16.
- Clark, B. J., Brown, J. E., & Taube, J. S. (2012). Head direction cell activity in the anterodorsal thalamus requires intact supragenual nuclei. *Journal of Neurophysiology*, *108*, 2767–2784.
- Compte, A., Brunel, N., Goldman-Rakic, P. S., & Wang, X. J. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cerebral Cortex*, *10*, 910–923.
- Cornsweet, T. N., & Teller, D. Y. (1965). Relation of increment thresholds to brightness and luminance. *Journal of the Optical Society of America*, *55*, 1303–1308.
- Cullen, K. E. (2012). The vestibular system: Multimodal integration and encoding of self-motion for motor control. *Trends in Neurosciences*, *35*, 185–196.
- Etienne, A. S., & Jeffery, K. J. (2004). Path integration in mammals. *Hippocampus*, *14*, 180–192.
- Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, *9*, 292–303.
- Fechner, G. (1966). *Elements of Psychophysics*. (D. H. Howes, Ed.). New York City: Holt, Rinehart and Winston.
- Fuhs, M. C., & Touretzky, D. S. (2006). A Spin Glass Model of Path Integration in Rat Medial Entorhinal Cortex. *The Journal of Neuroscience*, *26*, 4266–4276.
- Fujita, N., Klatzky, R. L., Loomis, J. M., & Golledge, R. G. (1993). The Encoding-Error Model of Pathway Completion without Vision. *Geographical Analysis*, *25*, 295–314.
- Gaydos, H. F. (1958). Sensitivity in the Judgment of Size by Finger-Span. *The American Journal of Psychology*,

71, 557.

- Gil, M., Ancau, M., Schlesiger, M. I., Neitz, A., Allen, K., De Marco, R. J., & Monyer, H. (2018). Impaired path integration in mice with disrupted grid cell firing. *Nature Neuroscience*, *21*, 81–91.
- Glasauer, S., Amorim, M.-A., Viaud-Delmon, I., & Berthoz, A. (2002). Differential effects of labyrinthine dysfunction on distance and direction during blindfolded walking of a triangular path. *Experimental Brain Research*, *145*, 489–497.
- Guanella, A., Kiper, D., & Verschure, P. (2007). A model of grid cells based on a twisted torus topology. *International Journal of Neural Systems*, *17*, 231–240.
- Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, *436*, 801–806.
- Harris, M. A., & Wolbers, T. (2012). Ageing effects on path integration and landmark navigation. *Hippocampus*, *22*, 1770–1780.
- Hasselmo, M. E. (2008). Grid cell mechanisms and function: Contributions of entorhinal persistent spiking and phase resetting. *Hippocampus*, *18*, 1213–1229.
- Indow, T., & Stevens, S. S. (1966). Scaling of saturation and hue. *Perception & Psychophysics*, *1*, 253–271.
- Izard, V., & Dehaene, S. (2008). Calibrating the mental number line. *Cognition*, *106*, 1221–1247.
- Jones, K. E., Hamilton, A. F., & Wolpert, D. M. (2002). Sources of signal-dependent noise during isometric force production. *Journal of Neurophysiology*, *88*, 1533–1544.
- Kiani, R., Churchland, A. K., & Shadlen, M. N. (2013). Integration of direction cues is invariant to the temporal gap between them. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *33*, 16483–16489.
- 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.
- Konishi, S., & Kitagawa, G. (2008). *Information Criteria and Statistical Modeling*. New York: Springer.
- Lappe, M., Jenkin, M., & Harris, L. R. (2007). Travel distance estimation from visual motion by leaky path integration. *Experimental Brain Research*, *180*, 35–48.
- Lappe, M., Stiels, M., Frenz, H., & Loomis, J. M. (2011). Keeping track of the distance from home by leaky integration along veering paths. *Experimental Brain Research*, *212*, 81–89.
- Laurens, J., & Angelaki, D. E. (2011). The functional significance of velocity storage and its dependence on gravity. *Experimental Brain Research*, *210*, 407–422.
- 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*, 73–91.
- Luis, C. A., Keegan, A. P., & Mullan, M. (2009). Cross validation of the Montreal Cognitive Assessment in community dwelling older adults residing in the Southeastern US. *International Journal of Geriatric Psychiatry*, *24*, 197–201.
- Mahmood, O., Adamo, D., Briceno, E., & Moffat, S. D. (2009). Age differences in visual path integration. *Behavioural Brain Research*, *205*, 88–95.
- 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*, 663–678.
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American Geriatrics Society*, *53*, 695–699.

- Oberlin, K. W. (1936). Variation in intensive sensitivity to lifted weights. *Journal of Experimental Psychology*, *19*, 438–455.
- Pavliotis, G. A. (2014). *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations*. New York: Springer.
- 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*, *31*, 17220–17229.
- Petzschner, F. H., Glasauer, S., & Stephan, K. E. (2015). A Bayesian perspective on magnitude estimation. *Trends in Cognitive Sciences*, *19*, 285–293.
- Pinto, L., Koay, S. A., Engelhard, B., Yoon, A. M., Deverett, B., Thiberge, S. Y., Witten, I. B., Tank, D. W., & Brody, C. D. (2018). An Accumulation-of-Evidence Task Using Visual Pulses for Mice Navigating in Virtual Reality. *Frontiers in Behavioral Neuroscience*, *12*, 36.
- Raftery, A. E. (1995). Bayesian Model Selection in Social Research. *Sociological Methodology*, *25*, 111–163.
- Sadalla, E. K., & Montello, D. R. (1989). Remembering Changes in Direction. *Environment and Behavior*, *21*, 346–363.
- Schmidt, R. A., Zelaznik, H., Hawkins, B., Frank, J. S., & Quinn, J. T. (1979). Motor-output variability: A theory for the accuracy of rapid motor acts. *Psychological Review*, *47*, 415–451.
- Stangl, M., Achtzehn, J., Huber, K., Dietrich, C., Tempelmann, C., & Wolbers, T. (2018). Compromised Grid-Cell-like Representations in Old Age as a Key Mechanism to Explain Age-Related Navigational Deficits. *Current Biology*, *28*, 1108–1115.
- Thrun, S., Burgard, W., & Fox, D. (2005). *Probabilistic Robotics*. Cambridge, MA: MIT Press.
- Uhlenbeck, G. E., & Ornstein, L. S. (1930). On the Theory of the Brownian Motion. *Physical Review*, *36*, 823–841.
- Wiener, J. M., Berthoz, A., & Wolbers, T. (2011). Dissociable cognitive mechanisms underlying human path integration. *Experimental Brain Research*, *208*, 61–71.
- Wu, S., Hamaguchi, K., & Amari, S. (2008). Dynamics and Computation of Continuous Attractors. *Neural Computation*, *20*, 994–1025.
- Xie, Y., Bigelow, R. T., Frankenthaler, S. F., Studenski, S. A., Moffat, S. D., & Agrawal, Y. (2017). Vestibular Loss in Older Adults Is Associated with Impaired Spatial Navigation: Data from the Triangle Completion Task. *Frontiers in Neurology*, *8*, 173.

Figures

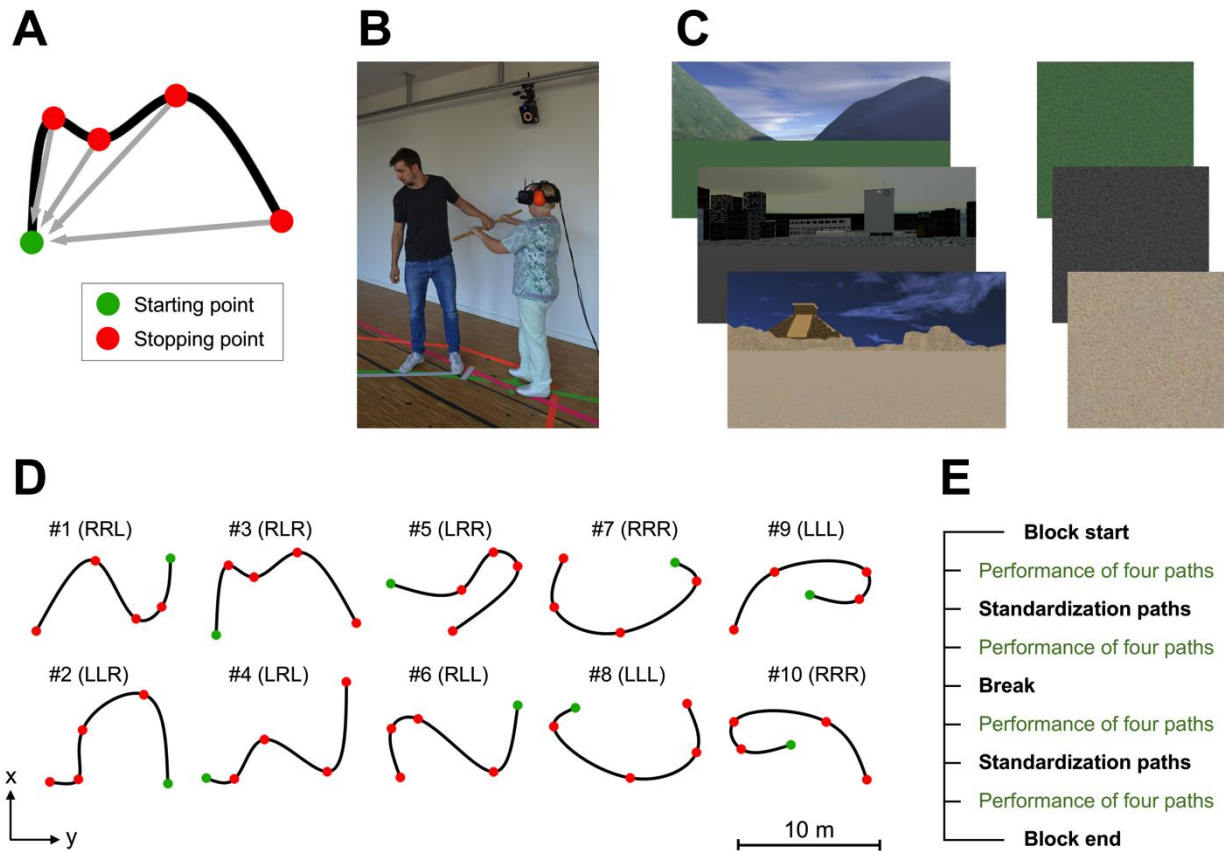


Figure 1: Path integration task.

A: Example path, viewed from a top-down perspective. Participants began at the starting point (green dot) and then walked along the path (curved black line). There were four stopping points (red dots) along each path; at these points, participants were asked to report their estimate of the direct distance and angle to the path's starting point (gray arrows).

B: During the experiment, participants saw a virtual environment from first-person perspective via a head-mounted display (HMD). All movements in the real world were tracked with a motion tracking system and translated to movements (i.e., changes in location and viewing orientation) in the virtual environment, so that participants could walk around and perform the task in the virtual environment. Participants held a wooden stick and were guided by the experimenter along a path. At each stopping point, the direct distance to the starting point had to be estimated verbally in meters and centimeters, and participants turned their body on the spot to indicate the orientation to the starting point.

C: The three different virtual environments (left panel) used in the path integration task. Each environment comprised a ground plane and distant landmark cues. As these landmark cues were rendered at infinity, they can facilitate the determination of heading direction, but do not allow participants to determine their position or distance information. One tile of each environment's ground plane is shown in the right panel. These tiles were textured to provide optic flow during movement, but were seamless (no visible border between adjacent tiles) and provided no fixed cues with positional information. Across environments, the ground planes differed only in color.

D: Overview of the 10 different paths used in the experiment. Each path contained three turns, and turn directions (i.e., left "L" and right "R" turns) were counter-balanced between paths.

E: Each participant performed three blocks of the path integration task. Each block consisted of 16 paths (paths #1-10, and paths #1-6 repeated without intermediate stopping at stopping points 1-3) in pseudo-randomized order. In addition, after the 4th and 12th path of each block, participants performed so-called "standardization-paths" (i.e., straight lines with a length of 2m, 6m, and 10m), which were used during data analysis to correct for each participant's abilities in verbally reporting distances using meter/centimeter units. Finally, there were short breaks between blocks and in the middle of each block. See Methods for more details.

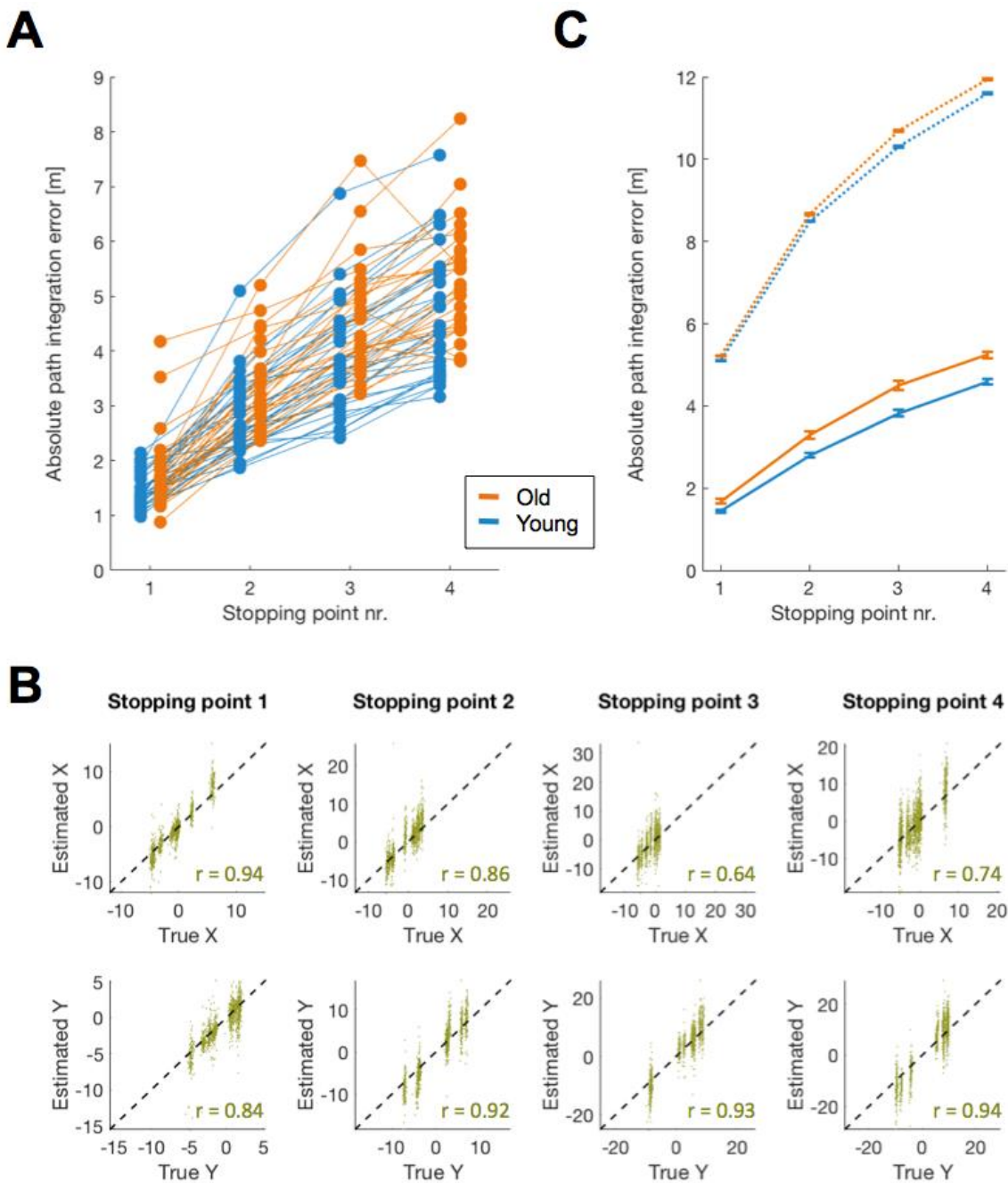


Figure 2: Path integration performance across both age groups.

A: Absolute path integration errors over four stopping points for young and older adults. Average errors per stopping point are shown for each participant separately by blue (young adults) and orange (older adults) dots, connected with lines between stopping points.

B: Each participant's location estimate (y-axis) versus their true location (x-axis) at each of the four stopping points (columns), separately for x-coordinates (top row) and y-coordinates (bottom row). Plots show data from all participants and all paths. The diagonal (dashed line) indicates perfect response (estimated location = true location). All correlation coefficients are statistically significant (all $p < 0.00001$). All units are meters.

C: Absolute path integration errors of young and older adults versus errors with shuffled responses. It is evident that the mean absolute path integration error of both groups (solid lines) is much lower than the errors obtained from shuffling each participant's responses across trials. Error bars indicate group mean \pm SEM.

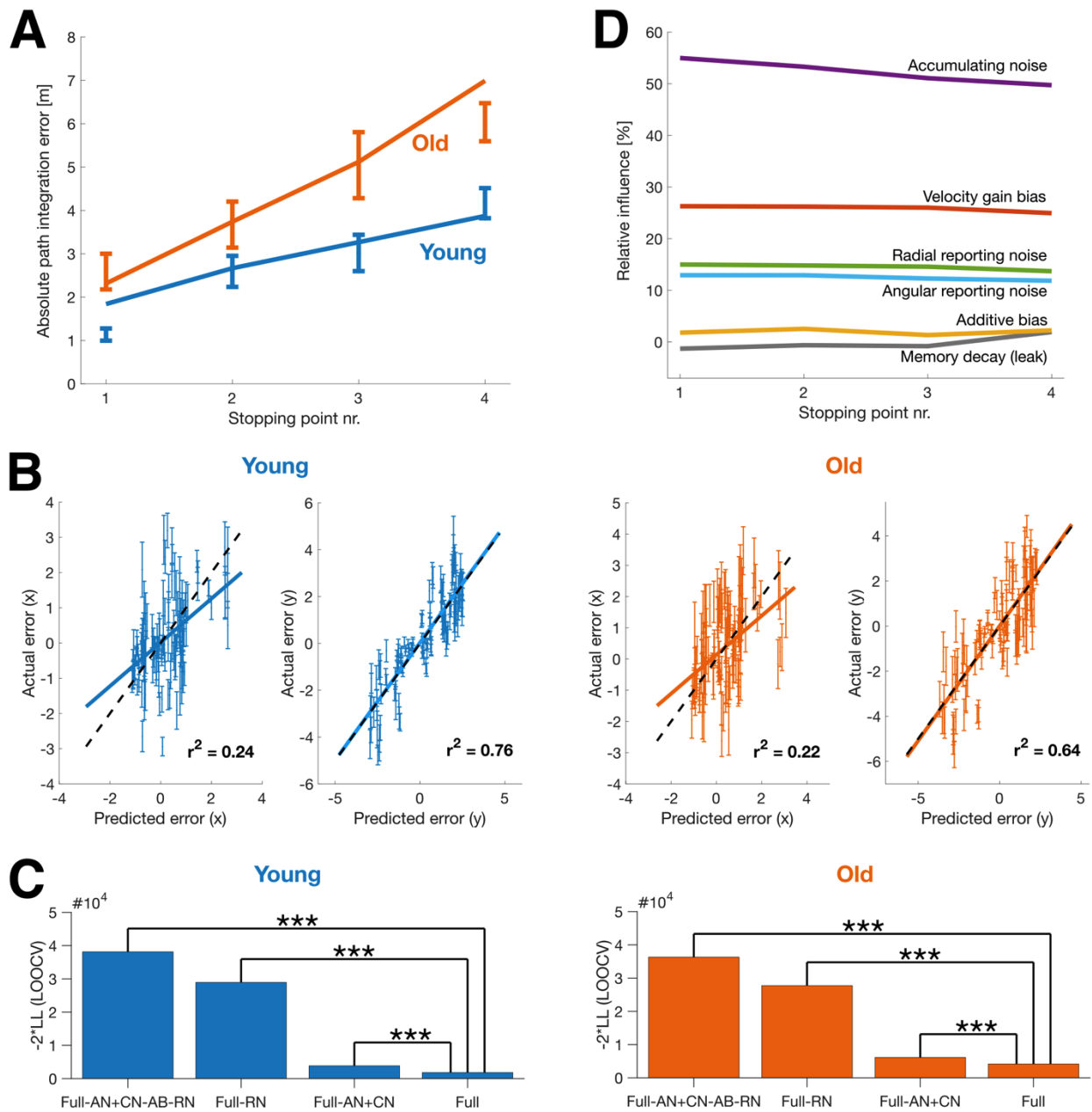


Figure 3: Computational modeling results.
(Figure Legend on next page)

Figure Legend (Figure 3)

A: Path integration errors, averaged across all trials, of one example participant from each of the two age groups (error bars), and model fits (solid lines). Error bars represent variability (SEM) over trials. Data and fits for all individual participants and average fits by age group are in **Supplementary Figure S2**.

B: Measured single-trial path integration error vectors versus error vectors predicted by the model. Predicted position is computed individually per subject per trajectory; datapoints show the per-trajectory predicted position, averaged across subjects of the same age group on the same trajectory and trial (to reduce scatter). Error bars represent the variability (SEM) at a single trial across participants (trial order was identical across participants). The dashed black (diagonal) line indicates perfect prediction; solid lines represent the best-fitting linear regression fit of the datapoints. All units are meters.

C: Model comparison: negative log-likelihood scores using LOOCV between models, with higher bars indicating a poorer model fit. *** denotes “very strong” evidence against the model relative to the full model ($\Delta BIC (\Delta LOOCV) \gg 10$; see also Methods section and **Supplementary Figure S3**). Key to model names: The full model (“Full”) is our default, with ongoing “accumulating noise” (AN) that is proportional to the length of the travelled path, nonzero additive bias (AB) and velocity gain bias parameters, and reporting noise (RN) that is proportional to the magnitude of the reported variable. CN refers to when the non-reporting portion of the noise is constant rather than accumulating. CRN refers to when the reporting noise is constant (rather than proportional to the magnitude of the reported variable). +/- refers to the addition/removal of that contribution to the model, respectively.

D: Impact of different model parameters on the predicted path integration error. Relative influence measures the predicted reduction in square error by setting a parameter to its ideal value corresponding to noiseless and unbiased integration. Note that due to the nonlinearity of the model, the relative influences do not have to sum to 100%, and that a parameter’s relative influence can be negative if the reduced square error is larger than the square error of the full model (see Methods section for more details). Ongoing noise, whose effect is cumulative and proportional to travelled distance, has the largest relative influence.

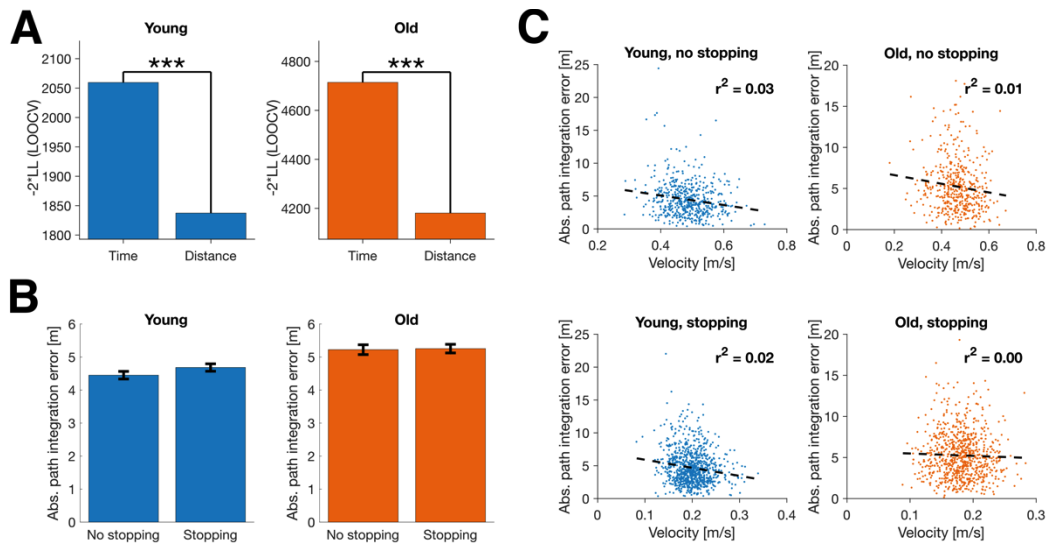


Figure 4: Time- vs distance-scaling of accumulating noise.

A: Model comparison using LOOCV between the full model with accumulated error proportional to total travel distance versus total time in trajectory. For both age groups, the full model is better supported by the data. Higher bars indicate poorer model-fit. *** denotes “very strong” evidence against the model with poorer fit (ΔBIC ($\Delta LOOCV$) $\gg 10$; see Methods and **Supplementary Figure S5**).

B: Average path integration error at the last stopping point, in trials with and without intermediate stopping points. The path integration error is very similar even though trials with stopping take much more time, indicating that the path integration error mainly scales with distance rather than time. Error bars indicate SEM.

C: Walking velocity versus path integration error for trials with and without stopping and for both age groups.

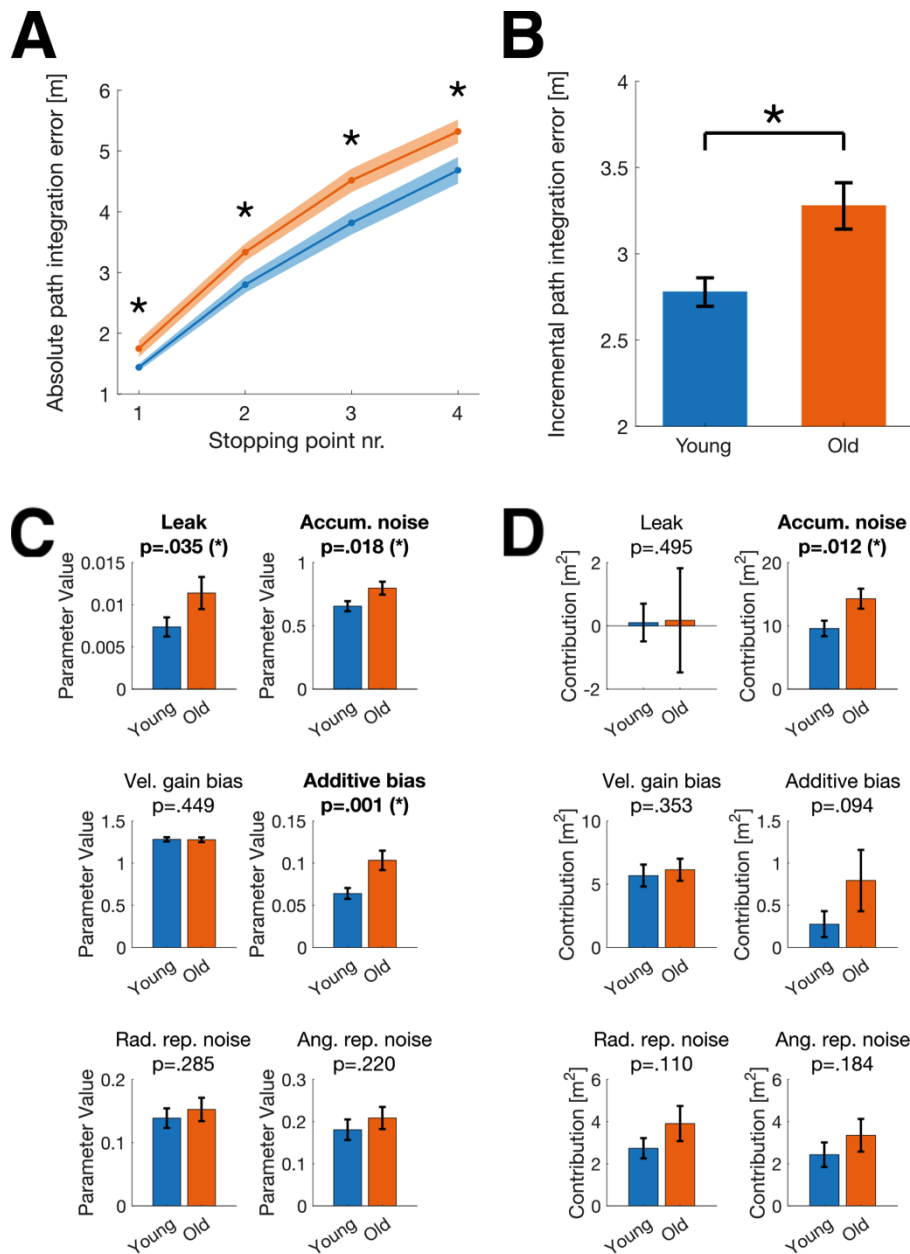


Figure 5: Path integration in older versus young adults.

A: On average, older adults showed a higher absolute path integration error than young adults at all stopping points. Blue and orange lines indicate group mean \pm SEM.

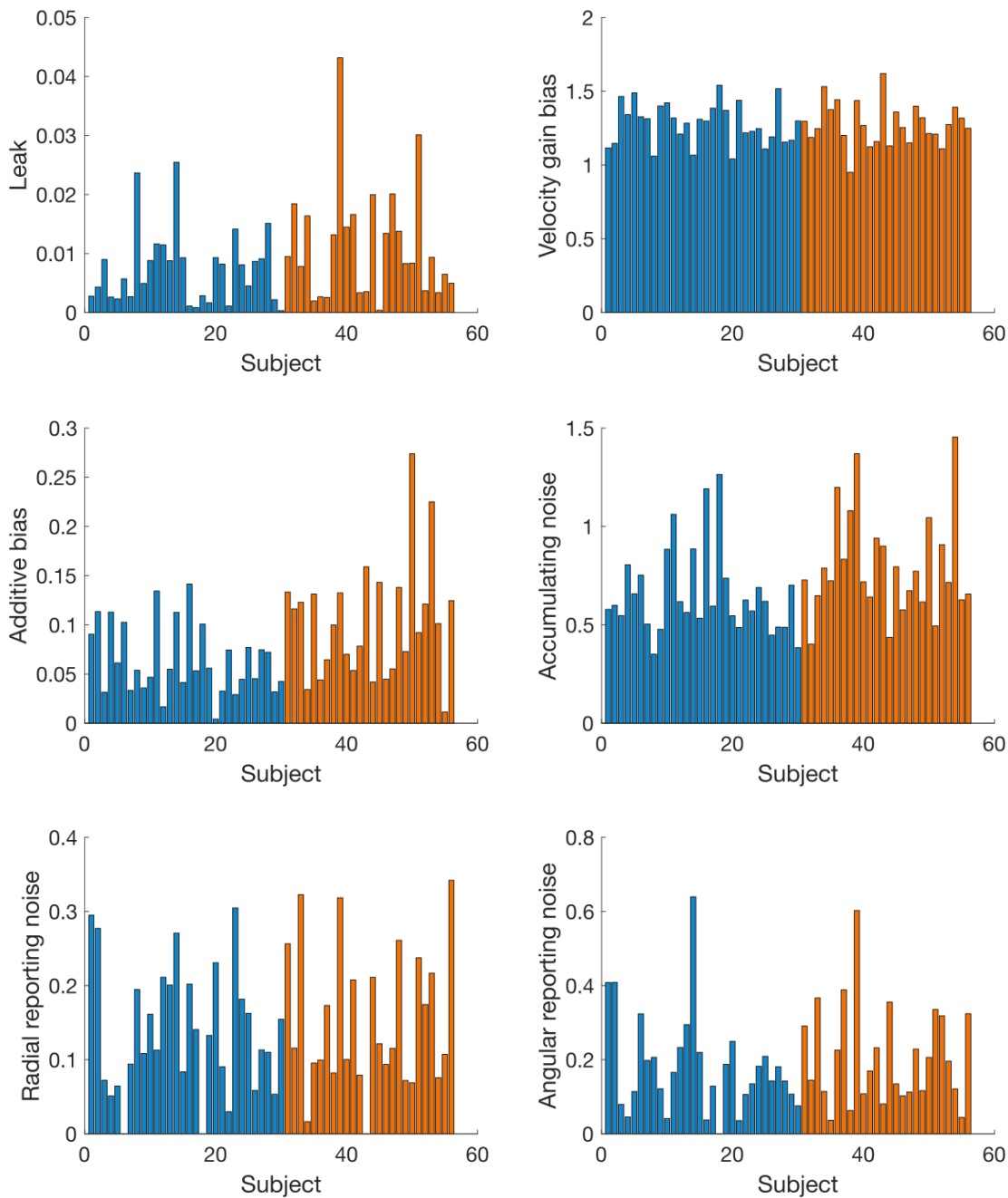
B: The incremental path integration error (i.e., the additional contribution to the path integration error for each segment between adjacent stopping points), averaged across stopping points, was higher for older than young adults.

C: Model parameter values, averaged over participants of the same age group. Parameter values for leak, accumulating unbiased noise, and additive bias were significantly higher in older relative to young adults. Individual parameter values for single subjects are shown in **Supplementary Figure S1**.

D: Each model parameter's contribution to the absolute square error, averaged over participants of the same age group. Only the accumulating unbiased noise resulted in a significant difference in error contribution between age groups. A parameter's contribution is calculated by measuring the reduction in square error when setting the parameter to its ideal value corresponding to unbiased, noiseless integration (see Methods).

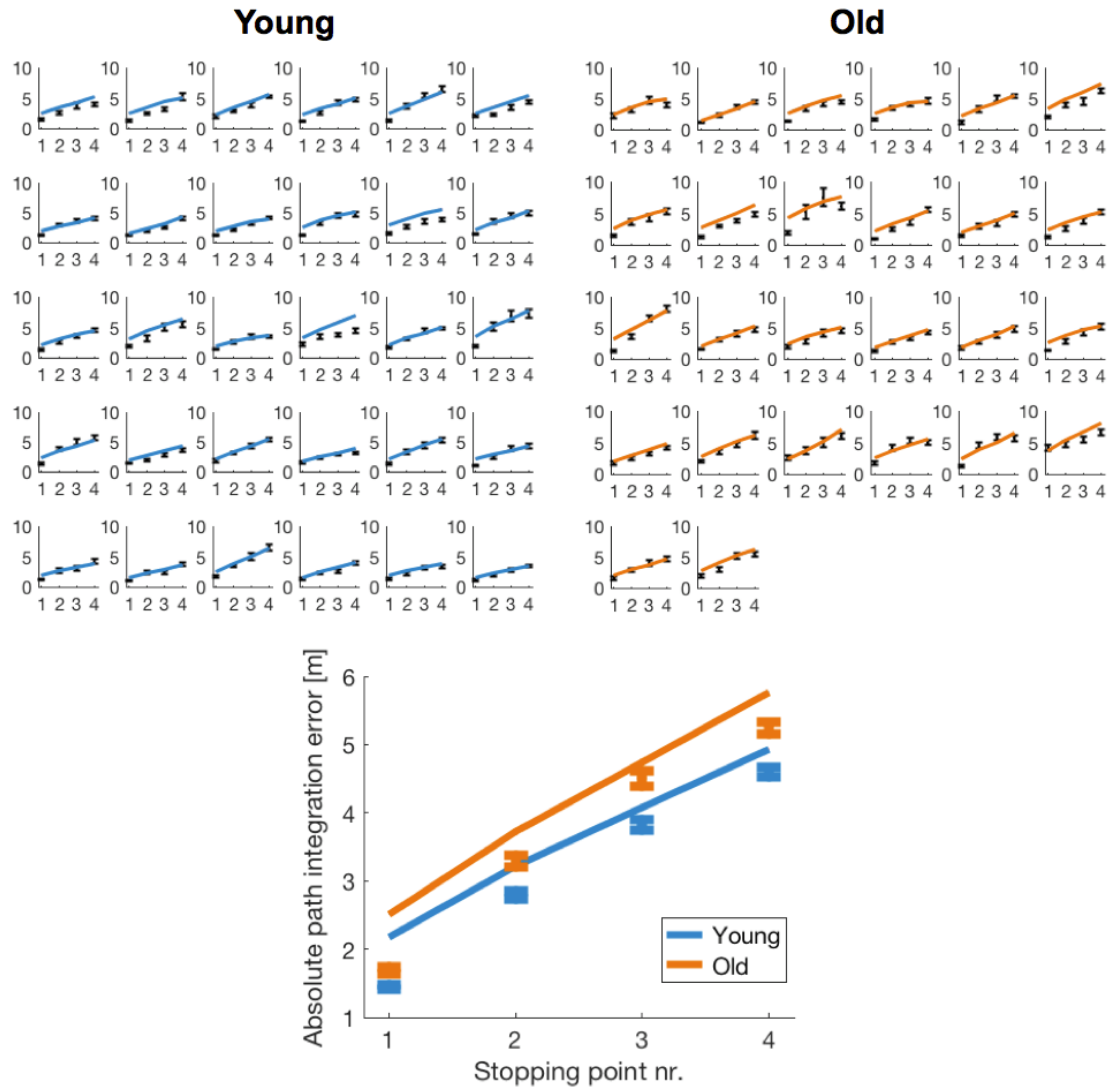
Error bars indicate SEM. * denotes a significant difference between age groups ($p < 0.05$) in a one-sided permutation test with 10000 permutations.

Supplementary information



Supplementary Figure S1: Individual parameter values for single subjects

Model parameter fit of the full model, shown for each subject individually. Blue bars indicate young subjects, orange bars indicate older subjects. Subject ordering is identical across plots.

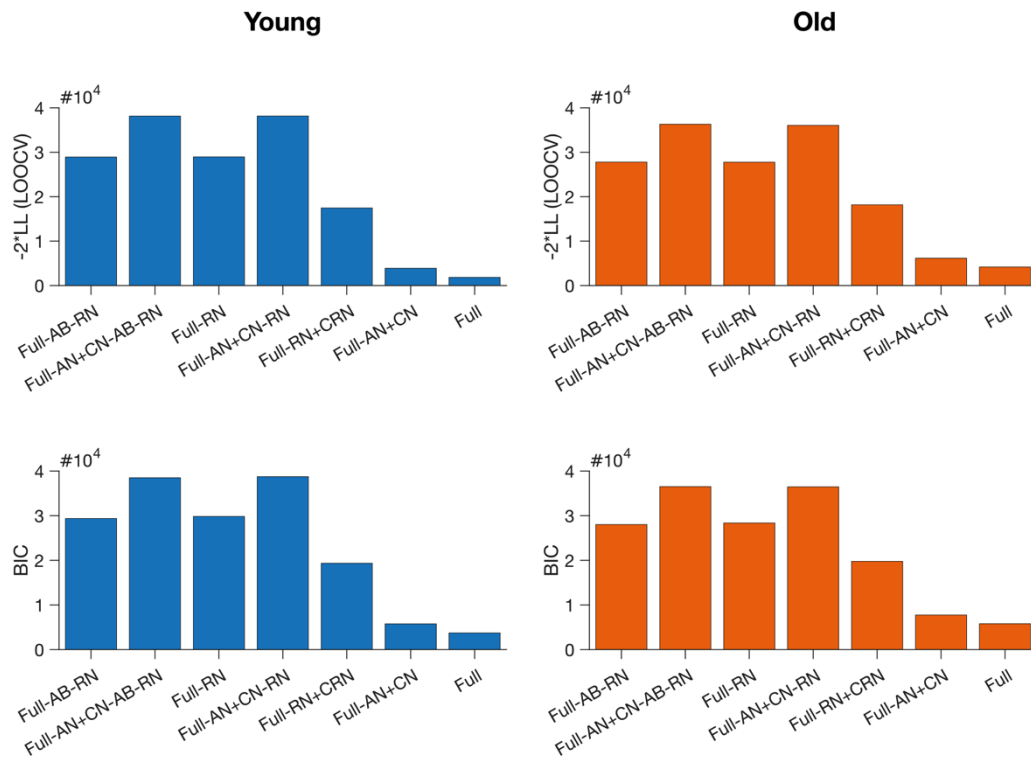


Supplementary Figure S2: Path integration errors of young and older adults versus errors predicted by the model

Top panel: Path integration errors of each individual participant (black error bars) versus errors predicted by the model (solid lines) using each participant's individual model parameters.

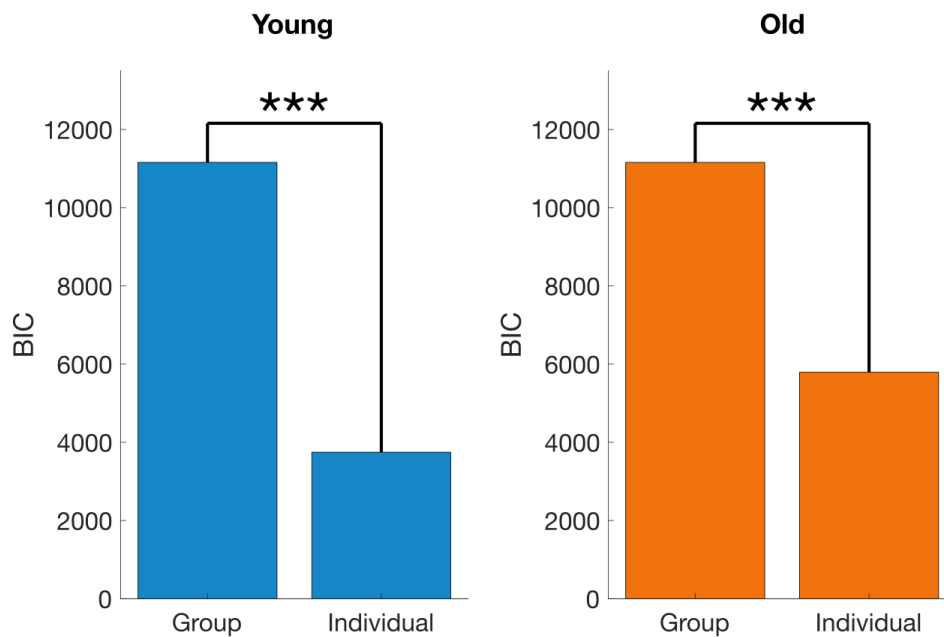
Bottom panel: Mean path integration errors per age group (error bars) versus errors predicted by the computational model (solid lines). In order to calculate group-level model predictions, participants in each age group are constrained to have the same model parameters, instead of fitting model parameters individually for each participant.

Error bars indicate mean \pm SEM.



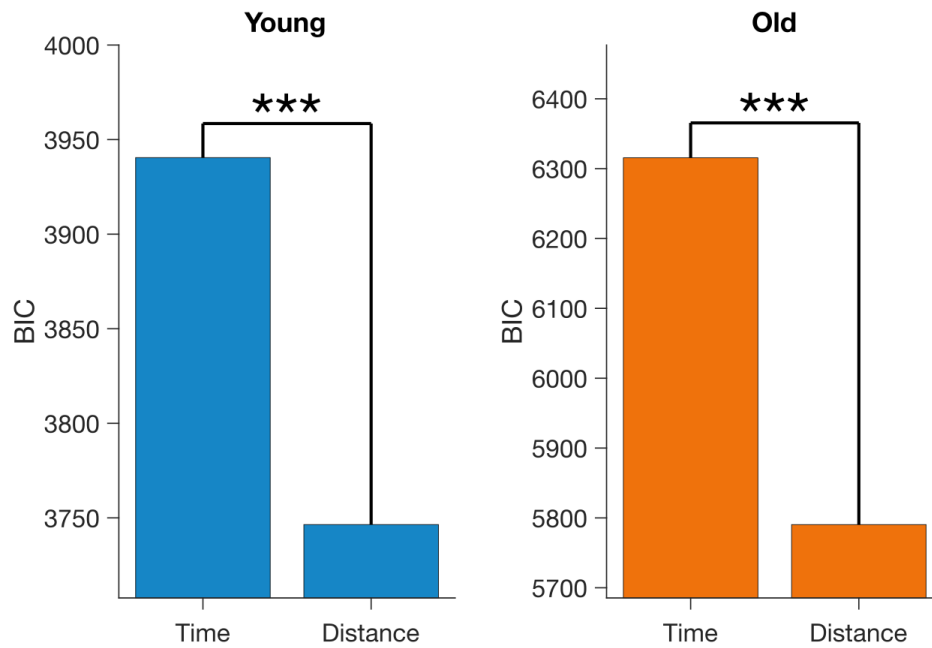
Supplementary Figure S3: Comparison between model variants using BIC and LOOCV.

Comparison between full model without additive bias, no reporting noise (Full-AB-RN); full model with constant instead of accumulating noise, no additive bias, no reporting noise (Full-AN+CN-AB-RN); full model without reporting noise (Full-RN); full model with constant instead of accumulating noise, no reporting noise (Full-AN+CN-RN); full model with constant reporting noise (Full-RN+CRN); full model with constant instead of accumulating noise (Full-AN+CN); and the default full model (Full). For both age groups, the full model was best supported by the data. Higher bars indicate poorer model-fit. More details about different model variants and BIC/LOOCV model comparisons are provided in the Methods section.



Supplementary Figure S4: Comparison between group-level and individual models.

Model comparison using BIC between models that were fitted at the group level and models that were fitted individually for each participant. For both age groups, the model with individual parameters per participant (i.e., the full model) was best supported by the data. Higher bars indicate lower model-fit. *** denotes "very strong" evidence against the model with lower model-fit ($\Delta BIC \gg 10$). More details about different model variants and BIC model comparisons are provided in the Methods section.



Supplementary Figure S5: Comparison between “time model” and “distance model” variant using BIC.

Model comparison using BIC between a model with time-scaling of the internal path integration error and the full model with distance-scaling. For both age groups, the full model is better supported by the data. Higher bars indicate lower model-fit. *** denotes “very strong” evidence against the model with lower model-fit ($\Delta BIC \gg 10$). More details about different model variants and BIC model comparisons are provided in the Methods section.