

# Sources of path integration error in young and aging humans

Matthias Stangl <sup>1,2,7</sup>✉, Ingmar Kanitscheider<sup>3,4,7</sup>✉, Martin Riemer<sup>2,5</sup>, Ila Fiete <sup>3,6,8</sup> & Thomas Wolbers<sup>2,5,8</sup>

Path integration plays a vital role in navigation: it enables the continuous tracking of one's position in space by integrating self-motion cues. Path integration abilities vary widely across individuals, and tend to deteriorate in old age. The specific causes of path integration errors, however, remain poorly characterized. Here, we combine tests of path integration performance in participants of different ages with an analysis based on the Langevin equation for diffusive dynamics, which allows us to decompose errors into distinct causes that can corrupt path integration computations. We show that, across age groups, the dominant error source is unbiased noise that accumulates with travel distance not elapsed time, suggesting that the noise originates in the velocity input rather than within the integrator. Age-related declines are primarily traced to a growth in this noise. These findings shed light on the contributors to path integration error and the mechanisms underlying age-related navigational deficits.

<sup>1</sup>Department of Psychiatry and Biobehavioral Sciences, Jane and Terry Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles (UCLA), Los Angeles, CA, USA. <sup>2</sup>German Center for Neurodegenerative Diseases (DZNE), Aging & Cognition Research Group, Magdeburg, Germany. <sup>3</sup>Center for Learning and Memory, Department of Neuroscience, The University of Texas, Austin, TX, USA. <sup>4</sup>OpenAI, San Francisco, CA, USA. <sup>5</sup>Center for Behavioral Brain Sciences, Magdeburg, Germany. <sup>6</sup>Department of Brain and Cognitive Sciences & McGovern Institute, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. <sup>7</sup>These authors contributed equally: Matthias Stangl, Ingmar Kanitscheider. <sup>8</sup>These authors jointly supervised this work: Ila Fiete, Thomas Wolbers. ✉email: [stangl@ucla.edu](mailto:stangl@ucla.edu); [ingmar@openai.com](mailto:ingmar@openai.com)

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. Critically, path integration also serves as a building block for cognitive maps<sup>1,2</sup>, because it allows us to associate environmental cues with positional estimates, and it plays a key role in spatial knowledge transfer from cognitive maps to route knowledge and wayfinding<sup>3</sup>. As a consequence, deficits in path integration could not only explain interindividual variability in navigational performance, but they could also contribute to the decline of navigational abilities commonly seen in old age<sup>4</sup>.

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<sup>5</sup>. After being processed in their respective low-level sensory systems, these cues are integrated in brainstem nuclei, as well as cortical structures, to allow for an overall estimation of angular and linear movement velocity<sup>6–12</sup>. The integration of these cues is an error-prone process, and previous studies have demonstrated that path integration abilities therefore vary largely across individuals<sup>13–15</sup>. 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 participants of different ages and we 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.<sup>16</sup>).

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, and 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.

Previous path integration models found that errors in path integration computations might stem from a leaky integration of self-motion cues, or from a bias in estimating self-motion speed<sup>17–19</sup>. These models, however, focused on only a small set of potential contributors to path integration errors, and they only tested path integration based on two-dimensional optic flow (i.e., automated movements shown on a computer screen). Here, we aim to build on this previous work, but take into account a broader range of potential sources of error (i.e., memory decay or

leak, velocity gain bias, additive location bias, accumulating or constant noise, and reporting errors) and quantify their influence and importance relative to each other, in order to obtain a more complete picture of the specific contributors to path integration performance in humans. Moreover, we use a path integration task in an immersive virtual reality setup that goes beyond passive visual path integration and also allows for providing richer body-based self-motion cues (including vestibular, proprioceptive, and motor cues), thus enabling an assessment of path integration performance under more realistic conditions.

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. Aging has deleterious effects on path integration ability<sup>20–23</sup>, with declines in the triangle completion task — a standard assay of path integration performance<sup>13</sup>. Moreover, older adults are less accurate in reproducing travel distances or rotations<sup>21–23</sup>, and they exhibit worse path integration performance even if additional landmark information is available<sup>23,24</sup>. Despite the sizeable body of research on losses in path integration performance with age, very little is known about which specific aspects of the path integration computation or process are most affected in old age. This knowledge, however, would significantly advance our understanding of navigational decline in old age and early dementia, which could ultimately promote the development of novel diagnostic tools and rehabilitative strategies to assess and improve navigational functioning.

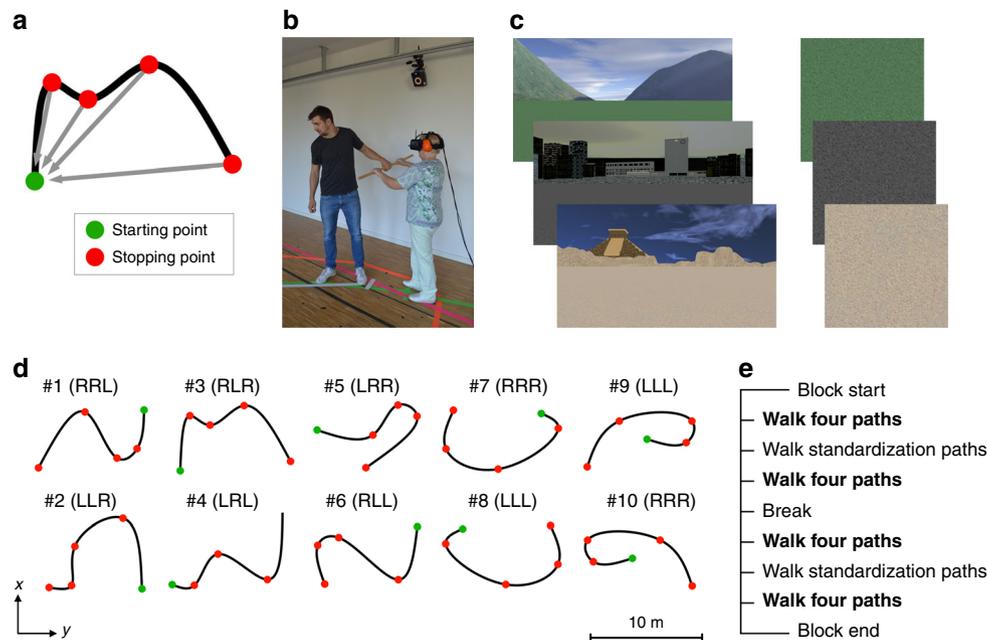
To address this important issue, we combine an immersive virtual reality path integration experiment with a mathematical approach to reveal the sources of path integration error. We characterize the different contributors to error across participants, and study group differences between young and older adults. The results of these analyses show that path integration errors in general, as well as age-related path integration deficits, are mainly driven by accumulating noise that originates in the velocity input to the path integrator.

## Results

**Path integration performance across age groups.** 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 poses (locations and viewing orientations) were tracked using the Vicon Motion Tracking System (Vicon, Oxford, UK) and translated into movements (i.e., changes in pose) 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, participants were asked to track their own position and orientation as they were guided through this environment by holding onto a baton moved by an experimenter along 10 distinct pre-defined but unmarked curved paths (Fig. 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. Each participant performed three repetitions of each path. In addition, a subset of six paths was performed another three times without intermediate stopping at the first three stopping points but only at the path's end (see “Methods” section for more details).

Most participants showed a characteristic increase in path integration error over the course of their trajectories (Fig. 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



**Fig. 1 Path integration task.** **a** Example path from 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. **b** During the experiment, participants saw a virtual environment from first-person perspective via a head-mounted display (HMD). 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. 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** Three different virtual environments (left panel) used in the path integration task. Each environment comprised a ground plane and distant landmark cues. Landmark cues were rendered at infinity, in order to allow participants to determine their heading direction, but not 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. **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** Participants 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 addition, after the 4th and 12th path of each block, participants performed so-called "standardization-paths" (i.e., straight lines with a length of 2, 6, and 10 m), which were used to correct for each participant's bias in converting their internal location estimate to a verbal response. Text in bold in panel **e** indicates the phases used for data analyses. See "Methods" section for more details.

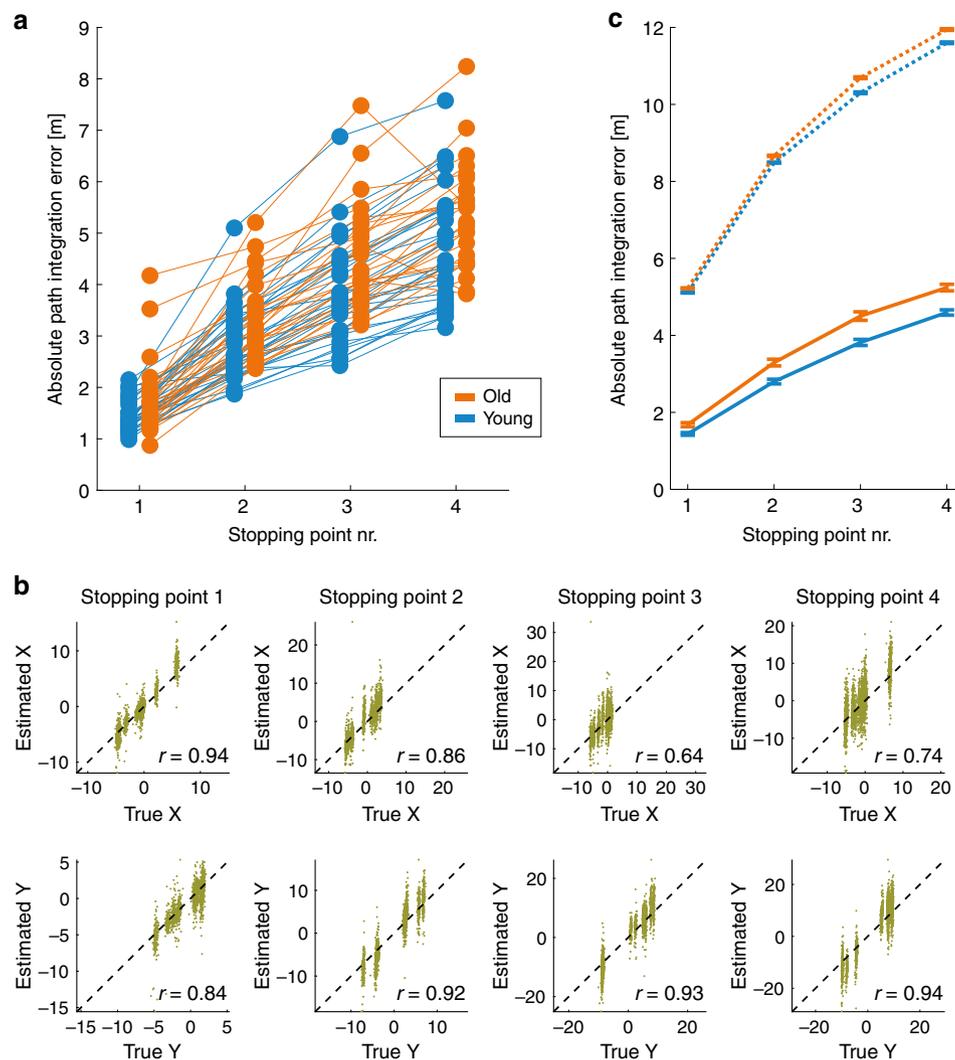
correlated with true location (Fig. 2b;  $r = 0.64-0.94$ , all  $p < 0.00001$ ) while shuffled responses across trials (corresponding to different trajectories) per stopping point exhibited much larger squared errors (Fig. 2c).

**Dynamical model of errors.** Next, we built and fit a temporally resolved computational model of the 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: underestimation or overestimation of velocity (velocity gain bias), leaky integration of the velocity signal (memory decay or leak), an additive bias (AB), and ongoing zero-mean Gaussian additive noise, 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" (Full), the accumulating noise (AN) 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 assume 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<sup>25-28</sup>. The mathematical details of the full model and different model variants are provided in the "Methods" section.

Model parameters per participant were obtained by the best fit across all paths and trials (Supplementary Fig. 1). The full model captured not only the magnitude of errors averaged across paths (Fig. 3a and b), but also predicted the full, time-resolved, signed errors at different portions of the different individual paths (Fig. 3c). Note that for the calculation of model predictions on the group level (Fig. 3b), participants in each age group were constrained to have the same model parameters, instead of fitting model parameters individually for each participant; thus, the variance around the fits includes inter-individual differences. However, when decomposing different sources of path integration error and age-related differences, we used models that were fitted individually for each participant (and we will show below that this approach provides better model-fit and was best supported by the data). For this reason, the performance of our model is best reflected by the fit between model predictions and empirical data for each participant individually (as shown Fig. 3a for two example participants, and in Supplementary Fig. 2 for each participant separately).

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, such as models where the



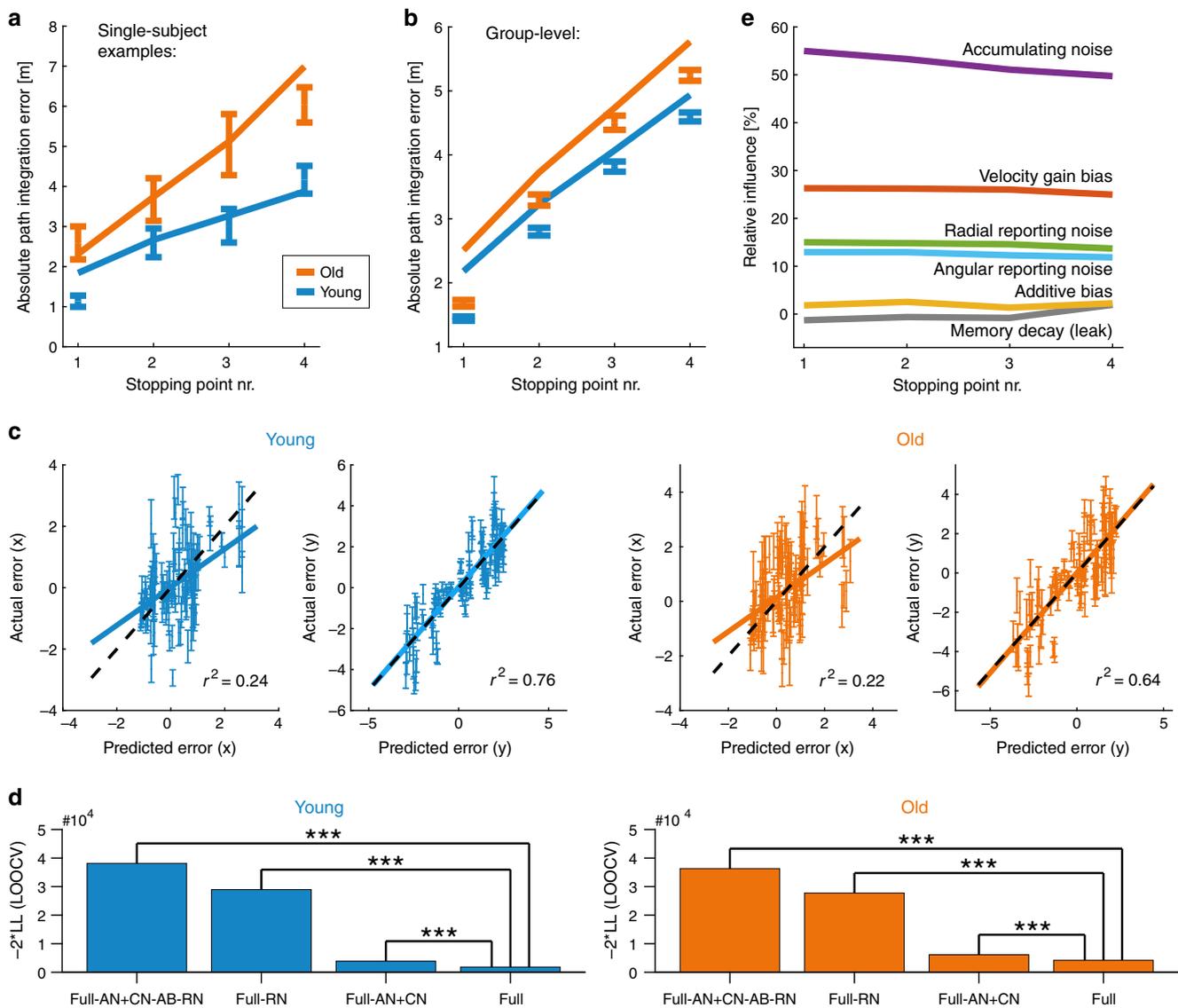
**Fig. 2** Path integration performance across both age groups. **a** Absolute path integration errors over four stopping points for young and older adults. Average errors per stopping point are shown for each participant separately by blue (young adults) and orange (older adults) dots, connected with 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  $p < 0.00001$ ). 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.

non-reporting portion of the noise remained constant (CN) instead of accumulating. We considered reporting noise that was proportional in magnitude to the reported variable, or constant in magnitude (CRN), or absent. 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 (Fig. 3d and Supplementary Fig. 3; 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, indicated by  $\Delta\text{BIC} \gg 10$  or  $\Delta\text{LOOCV} \gg 10$ ) 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,  $\Delta\text{BIC} = 36,303$ ;  $\Delta\text{LOOCV} = 34,743$ . Full vs. Full-RN,  $\Delta\text{BIC} = 27,103$ ;  $\Delta\text{LOOCV} = 26,089$ . Full vs. Full-AN+CN,  $\Delta\text{BIC} = 2035$ ;  $\Delta\text{LOOCV} = 2021$ . Old: Full vs. Full-AN+CN-AB-RN,

$\Delta\text{BIC} = 30,731$ ;  $\Delta\text{LOOCV} = 32,124$ . Full vs. Full-RN,  $\Delta\text{BIC} = 22,579$ ;  $\Delta\text{LOOCV} = 23,577$ . Full vs. Full-AN+CN,  $\Delta\text{BIC} = 1963$ ;  $\Delta\text{LOOCV} = 1957$ ). Specifically, the full model outperformed 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 analog to a leading existing model of human path integration performance<sup>17,18</sup>.

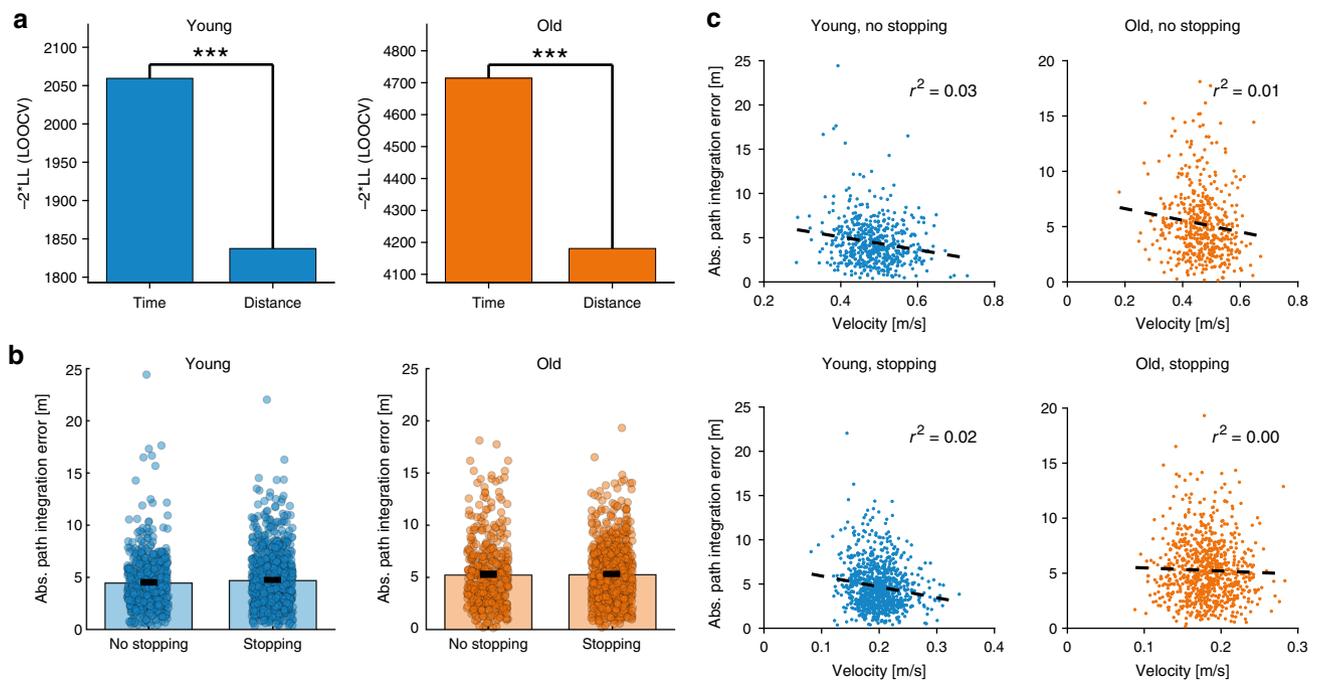
In addition, we carried out model comparisons using alternative cross-validation variants, in which (i) we trained the model on data from all but one trajectory and then tested it on all trials of the remaining, never-seen trajectory and (ii) we trained the model on data from only the first three stopping points of each path and then tested on the never-seen fourth stopping point (Supplementary Fig. 4). All cross-validation analyses led to highly similar results, in which the full model outperforms all other model variants. We therefore conclude that the full model is the most informative of all tested variants, across many cross-validation measures.



**Fig. 3 Computational modeling results.** **a** Path integration errors of two example participants (error bars) versus model fits (solid lines). Error bars represent mean  $\pm$  SEM over trials. See Supplementary Fig. 2 for all participants. **b** Average path integration errors per age group (error bars) versus model fits (solid lines). Error bars represent mean  $\pm$  SEM over participants. **c** Single-trial path integration error vectors versus error vectors predicted by the model. Predicted position is computed individually per participant per trajectory; datapoints show the per-trajectory predicted position, averaged across participants of the same age group on the same trajectory and trial (to reduce scatter). Error bars represent mean  $\pm$  SEM at a single trial across participants. Dashed black lines indicate perfect prediction; solid lines represent the best-fitting linear regression fit. Units are meters. **d** 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 ( $\Delta BIC$  or  $\Delta LOOCV \gg 10$ ; see also “Methods” section on model comparison, and Supplementary Fig. 3). Key to model names: The “full model” (Full) is our default, with ongoing “accumulating noise” (AN) that is proportional to the length of the traveled 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. +/- refers to the addition/removal of that contribution to the model, respectively. **e** Impact of 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).

To check whether our results might be a consequence of “overfitting” (i.e., a too close fit of an overly complex model to a limited set of training data, leading to poor generalization on unseen testing data), we carried out a quantitative comparison between training and test error. As shown in Supplementary Fig. 5, we find that in general, training and test error are very similar, and in particular, that for any pair of models the training error of the worse-fitting model is

larger than the test error of the better-fitting model. In addition, the agreement or consistency between the BIC and LOOCV results (see Supplementary Fig. 3) would likely not be as strong if each of their results were due to separate overfitting. Together, these results demonstrate that the estimated model fits (and the higher performance of the full model as compared to other variants) cannot be attributed to overfitting.



**Fig. 4** Time-scaling versus 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 ( $\Delta BIC$  or  $\Delta LOOCV \gg 10$ ; see “Methods” section on model comparison, and Supplementary Fig. 7). **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 mean  $\pm$  SEM (young adults:  $n = 540$  trials without stopping vs. 900 trials with stopping; older adults:  $n = 468$  trials without stopping vs. 780 trials with stopping). **c** Walking velocity versus path integration error for trials with and without stopping and for both age groups.

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 Fig. 6; 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 age group (see Supplementary Fig. 3).

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 neuronal noise<sup>29</sup>. 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 age groups (Fig. 4a and Supplementary Fig. 7; Young:  $\Delta BIC = 194$ ,  $\Delta LOOCV = 222$ ; Old:  $\Delta BIC = 525$ ,  $\Delta LOOCV = 533$ ).

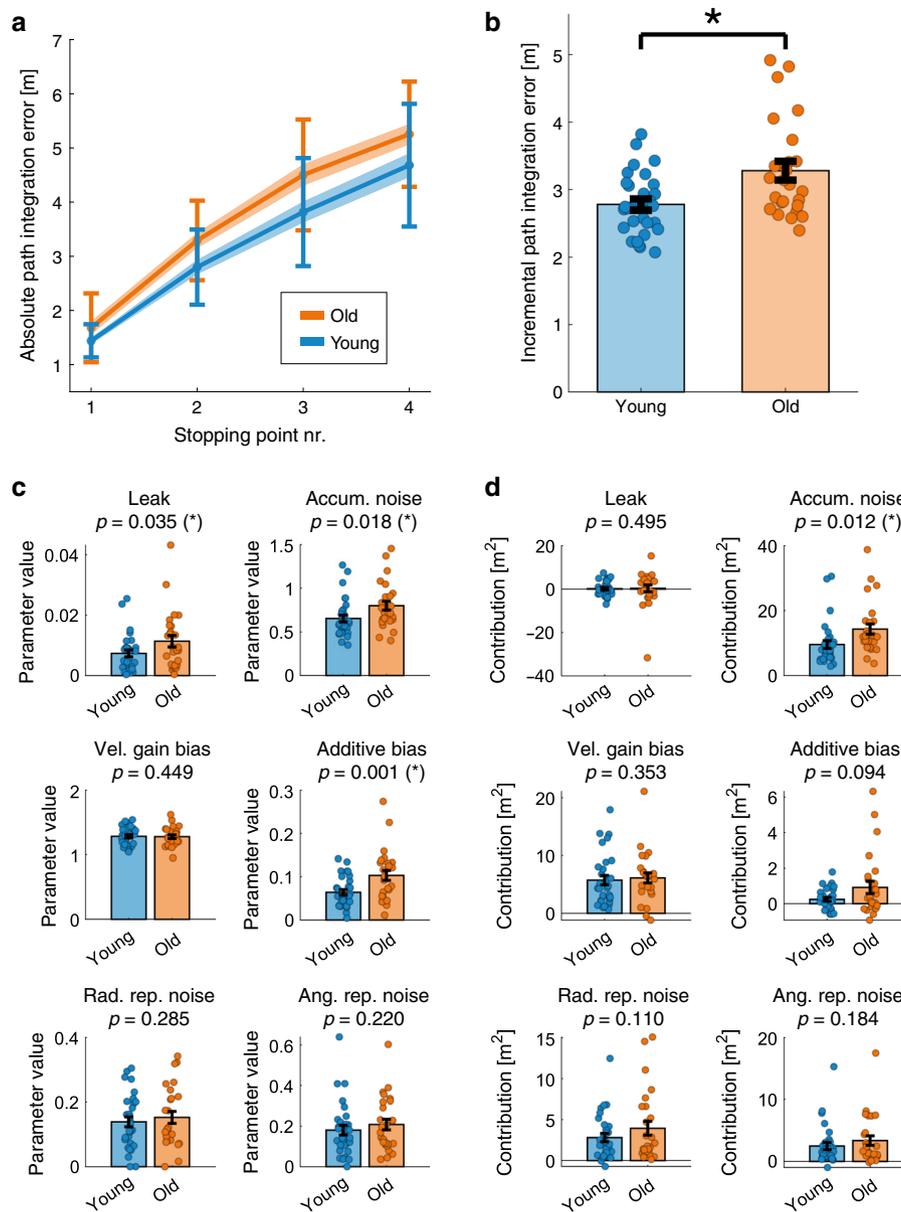
More directly, we compared total error on trajectories in which participants stopped versus did not stop at intermediate stopping points. Participants completed 48 paths in total, out of which 18 involved a stop only at the endpoint; in the remaining 30 paths, participants also stopped at three intermediate stopping points to report the distance and angle to the starting point (see “Methods” section for more details). Since the different paths had a very similar total length ( $17.7 \pm 0.1$  m, mean  $\pm$  SD), the total travel distances were similar over stopping and non-stopping trajectories, but the travel times differed substantially ( $88.7 \pm 12.4$  versus  $35.2 \pm 3.9$  s, mean  $\pm$  SD). Nevertheless, path integration

errors were very similar for stopping and non-stopping trajectories (Fig. 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.

Walking speed during the path integration task was quantified by traveled distance (as measured with the Vicon motion tracking system) divided by elapsed time between a path’s starting point and each stopping point. 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 (Fig. 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 (Fig. 3e). 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 (Fig. 3e) does not contradict the result that the introduction of reporting noise



**Fig. 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 shaded lines indicate group mean  $\pm$  SEM, error bars indicate SD. **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 ( $p = 0.001$ ). **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 participants are shown in Supplementary Fig. 1. **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; note that due to the non-linearity of the model a parameter’s contribution can be negative (see “Methods” section for more details). Dots indicate data for individual participants. Error bars in panels **b–d** indicate mean  $\pm$  SEM ( $n = 30$  young vs. 26 older participants). \* Denotes a significant difference between age groups ( $p < 0.05$ ) in a one-sided permutation test with 10,000 permutations.

causes the largest increase in model fit (Fig. 3d). Intuitively, Fig. 3d 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 Fig. 3e 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 Fig. 3d.

**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 (Fig. 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 (Fig. 5b;  $p = 0.001$ ).

In order to test whether the age effect in path integration performance might be driven only by a few particularly high or

low performing individuals, we compared incremental path integration performance between young and older adults in a reduced dataset, in which we excluded the data for the best-performing and worst-performing individuals in each group. Path integration performance remained significantly different between young and older adults after excluding participants with the 10% highest and 10% lowest path integration performance ( $p < 0.001$ ), and this age effect still persisted after excluding the best and worst 20% of each age group ( $p < 0.001$ ). Moreover, young adults showed a significantly higher path integration performance relative to a subset of older adults, in which we excluded those older adults who ranged within the 10% or 20% lowest path integration performance (10%:  $p = 0.012$ ; 20%:  $p = 0.034$ ), suggesting that the performance difference between young and older adults is not driven by a number of particularly impaired older adults.

We also checked whether differences in head movements during walking, which affect vestibular motion estimation, might have impacted our results (Supplementary Fig. 8). We found that the degree of head movement was not significantly different between age groups ( $p = 0.540$ ), and that head movement was not correlated with individual path integration performance across both young and older adults ( $r = 0.107$ ,  $p = 0.354$ ), or within individual age groups (old:  $r = 0.168$ ,  $p = 0.401$ ; young:  $r = -0.054$ ,  $p = 0.776$ ). We therefore conclude that head movement during walking did not have a significant impact on path integration performance within and across age groups.

Next, we compared participants' average angular velocity between young and older adults and found that angular velocity was not different between age-groups (Supplementary Fig. 9A;  $p = 0.267$ ). In addition, we tested whether within-subject variations in angular velocity across path segments might have affected our results. For each participant separately, we performed a median-split of all path segments based on the angular velocity along each segment. We tested whether participants showed higher or lower path integration errors on path segments with high versus low angular velocity (Supplementary Fig. 9B and C), but found no such performance differences, for all participants ( $p = 0.302$ ), or separately within the group of young ( $p = 0.485$ ) or older adults ( $p = 0.250$ ). These results suggest that our results cannot be driven by variations in angular velocity between or within participants.

Finally, to determine the underlying reason for the differential performance of older and young adults, we fit our computational model parameters individually across participants, 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 (Fig. 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 (Fig. 5d), suggesting that velocity estimation degrades in relatively unbiased ways, to become noisier in older relative to young humans.

## Discussion

We used an 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 traveling 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, as well as 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 generating an explicit report of an internal estimate of the displacement vector, as participants are asked to provide (at each stopping point). We performed mathematical inference of model parameters using an approach based on the extended Kalman filter (EKF). Disentangling the different sources of error allowed us to compare their influence on path integration errors across participants and between age groups. With this approach, we show that path integration computations are mainly corrupted by accumulating noise that mainly originates in the velocity input to the path integrator, and that an increase in this noise with age accounts for the majority of age-related path integration deficits. Existing path integration models (e.g., Lappe et al.<sup>17,18</sup> and Lakshminarasimhan et al.<sup>19</sup>) sought to explain path integration errors when participants indicated the magnitude of their displacement along straight or curved outbound paths. Their results suggested that path integration may be mainly corrupted by leaky integration or biased velocity estimation. 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. This difference in our results compared to previous findings might be explained at least by two ways in which the setups differ: First, the studies use different models to decompose error. All observed errors in previously applied path integration models were decomposed into a small number of terms (such as leak and velocity gain bias). A common problem that occurs when using models with less finely parsed error terms, is that a single error term might catch and contain other unmodeled sources of error. It is possible that the memory leak term in the simpler models of previous reports might have acted as a “catch-all” error term, which has, in turn, led to a belief that factors such as memory leak are major contributors to path integration errors. Our model built on this previous work but allowed for a range of additional parameters, most notably accumulating noise and reporting errors in addition to different bias factors, permitting richer possible interpretations of the 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 a range of simpler variants. Thus, it is implausible that participants exhibit substantial biases in velocity gain that were not discovered by the analysis model. Our results therefore suggest that sources of error that have been previously thought to play an important role (i.e., leaky integration and biased velocity estimation) may be less important than assumed, particularly in comparison to unbiased accumulating noise that was not included in previous models. Second, the participants in Lappe et al.<sup>17,18</sup> and Lakshminarasimhan et al.<sup>19</sup> formed motion estimates based only on optic flow during automated movements shown on computer or projection screens. In contrast, participants in our study actively walked in an immersive virtual environment with body-based self-motion cues including vestibular, proprioceptive and motor signals. In rodent studies, when motion cues are less rich (e.g., passive transport on trolleys; head-fixed animals in virtual environments), displacements are underestimated and firing of spatially selective cells is attenuated<sup>30–34</sup>, suggesting that a decreased availability of sensory motion cues in Lappe et al.<sup>17,18</sup> and Lakshminarasimhan et al.<sup>19</sup> may account for the dominant contribution of a velocity estimation bias in their findings.

Previous work<sup>35,36</sup> has considered the possibility that people 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. Participants 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 enable 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 older) human participants 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, for example due to neuronal noise within the brain's path integration circuit<sup>29,37–41</sup>. Alternatively, accumulating noise might arise from external velocity inputs to a neural integrator, which 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<sup>12,42,43</sup>. To refine our understanding of the origin 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. If the noise originated within the integrator, we would expect its accumulating variance to scale with elapsed time, whereas it is expected to scale with traveled distance if it originates outside the integrator (e.g., due to noise in the sensory systems or processing of sensory information). 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 might stem 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<sup>44</sup> and spatial integrators<sup>45,46</sup>, these results suggest an emerging principle in the neurobiology of integrators: that the dominant source of noise in the output of neural integrators originates in their inputs rather than from within the integrator circuit. This finding does not exclude, however, that there might still be a smaller contribution of internal neuronal noise, as suggested by previous studies<sup>47,48</sup>.

The present work also shows that path integration performance is consistently reduced across healthy older adults as compared to young adults, which is widely in line with previous studies<sup>21–24,47</sup>. Further, we were able to determine the dominant sources of error in older adults and thus determine which of the sources of error in young adults is most magnified as people 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 already present 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, what is already the most fragile part of the path integration process in younger adults is also most affected with aging.

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 cognitive mapping and spatial navigation more generally, 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.** Sixty-two 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<sup>49</sup>. Six older adults who did not exceed a MoCA cut-off score of 23 (following Luis et al.<sup>50</sup>) 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 (mean = 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 (mean = 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<sup>13,20–23,51</sup>, 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 (Fig. 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 disoriented 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 disoriented 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.

A central assumption of our analysis is that participants 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<sup>52</sup>. This strategy is often observed when the outbound path can be easily segmented into turns and distances — such as in the 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<sup>53</sup>). To eliminate these confounds, we used irregularly shaped sinuous paths, in which translations and rotations were combined into curved trajectories. In addition, we asked participants 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.

Prior to the task, participants received written information about the task, and completed several practice paths. Participants donned a HMD (Oculus Rift

Development Kit 2, Oculus VR LLC, [www.oculus.com](http://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 (Fig. 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. The orientation information for data analysis was derived from the HMD, which enables rotational tracking by a gyroscope, an accelerometer, and a magnetometer, and provides information about orientation per timepoint in three dimensions (yaw, pitch, roll), with a rotational update rate of 1000 Hz. For our analysis of participants' path integration performance, we used their orientation in the horizontal plane (yaw dimension).

Via the HMD, participants saw a virtual environment, which consisted of a ground plane and distant landmark cues (Fig. 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. Note that rendering landmarks at infinity was essential for our study, because if landmarks would provide positional and distance information, then participants would not need to path-integrate in order to determine their displacement, but they could simply use these landmarks (rather than self-motion cues) for localization. A setup in which landmarks were not rendered at infinity, would therefore not necessarily involve path integration computations, but would target a different cognitive/computational process.

The exact pose (locations and viewing orientation) of a participant was tracked in steps of 100 ms (10 Hz) 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<sup>5</sup>.

There were 10 different pre-defined paths (Fig. 1d). Coordinates for each path were defined as follows: First, a four-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 m, and each angle was either 55°, 80°, or 105° to the left or to the right. Various combinations of distances and angles were used, that fit into a rectangular area of approximately 10×8 m (given by the tracking area and size of the room in which the experiment took place). On the basis of these four-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 four-legged path.

Six paths comprised a mixture of left and right turns, respectively (see Fig. 1d, path numbers 1–6). 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 1–6 (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 (i.e., 48 paths in total: 30 with and 18 without intermediate stopping). 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 repeated instances of the same path. The virtual environment was different in each block (see Fig. 1c) and the order of environments was randomized across participants. Across the three different environments, the ground planes had identical textures (i.e., visual patterns) and differed only in color.

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, 2, 6 m. 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, Inc.; [www.worldviz.com](http://www.worldviz.com)). The virtual environments were designed using Autodesk 3ds Max 2014 (Autodesk, Inc.; [www.autodesk.com](http://www.autodesk.com)).

For a discussion of the ecological validity of our path integration task, and its relevance for everyday navigation in the real world, please see Supplementary Note 1.

The authors affirm that human research participants provided informed consent for publication of the image in Fig. 1b.

**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<sup>25</sup>. Here we assume that the bias is multiplicative. To measure the bias, we ask participants to walk on straight standardization paths of length 2, 6, and 10 m and to report verbally the distance to the starting point. The correction factor for the bias is then given by

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

where  $d_{\text{correct}}$  is the correct distance of the standardization path (2, 6, or 10 m, respectively),  $d_{\text{response}}$  is the responded distance, and  $f_{\text{corr}}$  is the resulting correction factor. For each participant, this led to three different correction factors, one each for shorter (derived from the 2 m standardization path), middle (derived from the 6 m standardization path), and longer distances (derived from the 10 m 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 0 and 4 m, the response was multiplied with the correction factor for shorter distances, whereas response distances between 4 and 8 m were multiplied with the correction factor for middle distances, and response distances larger than 8 m were multiplied with the correction factor for longer distances.

While this approach to standardize verbal distance estimates has been established earlier<sup>47</sup>, here we extended this approach, and also tested whether verbal estimates after walking along a straight standardization path directly correspond to verbal estimates after walking along curved paths during the task, both within and across different age groups. As shown in Supplementary Fig. 10, participants' distance estimation performance was not significantly different between task and standardization paths, suggesting that distance estimates (i.e., the error and noise in verbally estimating distances) during standardization paths correspond in magnitude to verbal distance estimates in the path integration task, both for young as well as for older adults. Moreover, it is evident that older adults on average tend to show more undershooting in their distance estimations; hence it was critical to standardize participants' response, in order to allow an unbiased comparison between young and older adults.

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  $f_{\text{corr}}$ ) 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

$$\begin{aligned} x_{\text{presumed}} &= x_{\text{stop}} + d_{\text{standardized}} * \cos(\text{ori}_{\text{response}}) \\ y_{\text{presumed}} &= y_{\text{stop}} + d_{\text{standardized}} * \sin(\text{ori}_{\text{response}}) \end{aligned}$$

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

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

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

where  $x_{\text{origin}}$  and  $y_{\text{origin}}$  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  $\text{Err}_{\text{inc}}$ . 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

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

where  $x_{\text{prevPresumed}}$  and  $y_{\text{prevPresumed}}$  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.,  $x_{\text{prevPresumed}} = x_{\text{origin}}$  and  $y_{\text{prevPresumed}} = y_{\text{origin}}$ ). 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.<sup>17,18</sup>) 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 participants 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<sup>17,18</sup>, 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 BIC.

We assume that each participant continuously updates an internal, two-dimensional estimate  $\hat{\mathbf{x}}(t)$  of his or her location  $\mathbf{x}(t)$  using an estimate of the true walking velocity  $\mathbf{v}(t)$ . The update process is compromised by memory decay  $\beta$ , velocity gain  $\alpha$ , additive bias  $\mathbf{b}$ , and Gaussian noise  $\xi(t)$  with standard deviation  $\sigma_0$  (where  $\xi(t)$  is normally distributed Gaussian noise) according to the following diffusion Langevin equation (note that bold-faced letters refer to two-dimensional vectors):

$$\frac{d\hat{\mathbf{x}}(t)}{dt} = -\beta\hat{\mathbf{x}}(t) + \alpha\mathbf{v}(t) + \mathbf{b} + \sigma_0\xi(t) \quad (1)$$

The parameters can be interpreted as follows:

- Memory decay or leak  $\beta$ : 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  $\alpha$ : 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  $\mathbf{b}$ : Specifies the bias direction along which the location estimate is pulled over time. Zero bias corresponds to  $\mathbf{b} = 0$ .
- Accumulating noise that is unbiased and additive with standard deviation  $\sigma_0$ : 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  $\sigma_0$  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 below). 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 (described in more detail 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 Fig. 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<sup>25-28</sup>. Given an internally estimated distance  $d$  and angle  $\varphi$ , we assume that the reported distances  $\hat{d}$  and angles  $\hat{\varphi}$  are given by

$$\begin{aligned} \hat{d} &= \exp(\log(d) + \sigma_d\eta_d) \\ \hat{\varphi} &= \varphi + \sigma_\varphi\eta_\varphi \end{aligned} \quad (2)$$

where  $\sigma_d$  and  $\sigma_\varphi$  are standard deviations of distance and angular noise,  $\eta_d$  is normally distributed distance noise, and  $\eta_\varphi$  is normally distributed angular noise. The parameterization of the distance reporting noise is chosen such that for fixed  $\sigma_d$ , the magnitude of the reporting error  $|\hat{d} - d|$  increases approximately linearly with  $d$  "proportional or Weber-like reporting noise", in line with Weber's law<sup>25,54-58</sup>. We find empirically (see "Results" section and Fig. 3d) 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).

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<sup>59,60</sup>. 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:

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

where  $I_2$  is the two-dimensional unity matrix and mean  $\mu_{s+1}$  and variance  $\sigma_{s+1}^2$  are given by

$$\begin{aligned} \mu_{s+1} &= \hat{\mathbf{x}}_s e^{-\beta t} + \frac{\alpha\mathbf{v} + \mathbf{b}}{\beta} (1 - e^{-\beta t}) \\ \sigma_{s+1}^2 &= \frac{\sigma_0^2}{2\beta} (1 - e^{-2\beta t}) \end{aligned} \quad (3)$$

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:

$$\begin{aligned} \mu_{s+1} &= \hat{\mathbf{x}}_s e^{-\tilde{\beta}|\Delta\mathbf{x}|} + (\alpha \frac{\Delta\mathbf{x}}{|\Delta\mathbf{x}|} + \tilde{\mathbf{b}}) \frac{1}{\tilde{\beta}} (1 - e^{-\tilde{\beta}|\Delta\mathbf{x}|}) \\ \sigma_{s+1}^2 &= \frac{\tilde{\sigma}_0^2}{2\tilde{\beta}} (1 - e^{-2\tilde{\beta}|\Delta\mathbf{x}|}) \end{aligned} \quad (4)$$

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

$$\tilde{\beta} = \frac{\beta}{|\mathbf{v}|} \quad \tilde{\mathbf{b}} = \frac{\mathbf{b}}{|\mathbf{v}|} \quad \tilde{\sigma}_0^2 = \frac{\sigma_0^2}{|\mathbf{v}|} \quad (5)$$

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<sup>17,18</sup>. We also explicitly test that the distance model is better supported by the data than the time model (see "Results" section and Fig. 4a).

**Model fitting without reporting noise.** For our full model without reporting noise (Full-RN), we explain here how the parameters  $\theta = (\tilde{\beta}, \alpha, \tilde{\mathbf{b}}, \tilde{\sigma}_0^2)$  related to integration and  $\kappa = (\sigma_d^2, \sigma_\varphi^2)$  related to reporting were fit to participants' performance by maximizing the likelihood. For simplicity, consider first a model without the reporting noise parameters  $\kappa$ . 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  $\hat{\varphi}$  to the starting point  $\mathbf{x}_{\text{start}}$  of the current walking trajectory:

$$\hat{\mathbf{x}}_s = \begin{pmatrix} \hat{d} \cos(\hat{\varphi}) \\ \hat{d} \sin(\hat{\varphi}) \end{pmatrix} + \mathbf{x}_{\text{start}} \quad (6)$$

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

$$LL_{\kappa=0}(\theta) = \left\langle \sum_{s=0}^3 \log p(\hat{\mathbf{x}}_{s+1} | \hat{\mathbf{x}}_s; \theta) \right\rangle_{\text{trials}} = \left\langle \sum_{s=0}^3 \log \mathcal{N}(\hat{\mathbf{x}}_{s+1} | \mu_{s+1}(\hat{\mathbf{x}}_s, \theta), \sigma_{s+1}^2(\theta)) \right\rangle_{\text{trials}} \quad (7)$$

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

$$\theta_{\text{ML}} = \text{argmax}_{\theta} LL_{\kappa=0}(\theta) \quad (8)$$

**Model fitting with reporting noise.** For our full model with reporting noise (Full), 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 EKF, a framework that permits calculation of the log-likelihood by locally linearizing the non-linearities<sup>61</sup>. 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  $\mathbf{z}_s$ :

$$\begin{aligned} \hat{\mathbf{x}}_{s+1} &= f(\hat{\mathbf{x}}_s) + \Sigma_x^{\frac{1}{2}} \xi_x \\ \mathbf{z}_{s+1} &= h(\hat{\mathbf{x}}_{s+1}) + \Sigma_z^{\frac{1}{2}} \xi_z \end{aligned} \quad (9)$$

where  $f$  and  $h$  are arbitrary non-linear functions and  $\Sigma_x$  and  $\Sigma_z$  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):

$$\begin{aligned} f(\hat{\mathbf{x}}_s) &= \mu_{s+1}(\hat{\mathbf{x}}_s) = \hat{\mathbf{x}}_s e^{-\beta|\Delta\mathbf{x}|} + \left( \alpha \frac{\Delta\mathbf{x}}{|\Delta\mathbf{x}|} + \bar{\mathbf{b}} \right) \frac{1}{\beta} (1 - e^{-\beta|\Delta\mathbf{x}|}) \\ \Sigma_x &= \sigma_{s+1}^2 I_2 = \frac{\bar{\sigma}_0^2}{2\beta} (1 - e^{-2\beta|\Delta\mathbf{x}|}) I_2 \end{aligned} \quad (10)$$

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:

$$\begin{aligned} h(\hat{\mathbf{x}}_{s+1}) &= \begin{pmatrix} d(\hat{\mathbf{x}}_{s+1}) \\ \varphi(\hat{\mathbf{x}}_{s+1}) \end{pmatrix} = \begin{pmatrix} \log|\hat{\mathbf{x}}_{s+1}| \\ \text{atan2}((\hat{\mathbf{x}}_{s+1})_2, (\hat{\mathbf{x}}_{s+1})_1) \end{pmatrix} \\ \Sigma_z &= \begin{pmatrix} \sigma_d^2 & 0 \\ 0 & \sigma_{\varphi}^2 \end{pmatrix} \end{aligned} \quad (11)$$

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

$$\mathbf{z}_{s+1} = \begin{pmatrix} \log \hat{d} \\ \hat{\varphi} \end{pmatrix} \quad (12)$$

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

$$\begin{aligned} p(\hat{\mathbf{x}}_{s+1} | \mathbf{z}_1, \dots, \mathbf{z}_s) &= \mathcal{N}(\hat{\mathbf{x}}_{s+1} | \mu_{s+1|s}, P_{s+1|s}) \quad (\text{predictive distribution}) \\ p(\hat{\mathbf{x}}_{s+1} | \mathbf{z}_1, \dots, \mathbf{z}_{s+1}) &= \mathcal{N}(\hat{\mathbf{x}}_{s+1} | \mu_{s+1|s+1}, P_{s+1|s+1}) \quad (\text{updated distribution}) \end{aligned} \quad (13)$$

Mean and covariance of both distributions can be calculated recursively over stopping points using the standard EKF update equations<sup>61</sup>:

$$\begin{aligned} \mu_{s+1|s} &= f(\mu_{s|s}) \\ P_{s+1|s} &= F_{s+1} P_{s|s} F_{s+1}^T + \Sigma_x \\ S_{s+1} &= H_{s+1} P_{s+1|s} H_{s+1}^T + \Sigma_z \\ K_{s+1} &= P_{s+1|s} H_{s+1}^T S_{s+1}^{-1} \\ \mu_{s+1|s+1} &= \mu_{s+1|s} + K_{s+1} (\mathbf{z}_{s+1} - h(\mu_{s+1|s})) \\ P_{s+1|s+1} &= (I_2 - K_{s+1} H_{s+1}) P_{s+1|s} \end{aligned} \quad (14)$$

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  $\mu_{s|s}$  and predictive mean  $\mu_{s+1|s}$ , respectively:

$$\begin{aligned} F_{s+1} &= \left. \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mu_{s|s}} = e^{-\beta|\Delta\mathbf{x}|} I_2 \\ H_{s+1} &= \left. \frac{\partial h(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mu_{s+1|s}} = \frac{1}{|\mu_{s+1|s}|^2} \begin{pmatrix} \mu_{s+1|s,1} & \mu_{s+1|s,2} \\ -\mu_{s+1|s,2} & \mu_{s+1|s,1} \end{pmatrix} \end{aligned} \quad (15)$$

At the starting point ( $s = 0$ , we initialize  $\mu_{s=0|s=0} = \mathbf{x}_{\text{start}} = 0$  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  $\hat{\mathbf{x}}_{s+1}$ :

$$\begin{aligned} p(\mathbf{z}_{s+1} | \mathbf{z}_1, \dots, \mathbf{z}_s) &= \int d\hat{\mathbf{x}}_{s+1} p(\mathbf{z}_{s+1} | \hat{\mathbf{x}}_{s+1}) p(\hat{\mathbf{x}}_{s+1} | \mathbf{z}_1, \dots, \mathbf{z}_s) \\ &= \int d\hat{\mathbf{x}}_{s+1} \mathcal{N}(\mathbf{z}_{s+1} | h(\hat{\mathbf{x}}_{s+1}), \Sigma_z) \mathcal{N}(\hat{\mathbf{x}}_{s+1} | \mu_{s+1|s}, P_{s+1|s}) \\ &\approx \int d\hat{\mathbf{x}}_{s+1} \mathcal{N}(\mathbf{z}_{s+1} | h(\mu_{s+1|s}) + H_{s+1}(\hat{\mathbf{x}}_{s+1} - \mu_{s+1|s}), \Sigma_z) \\ &\quad \mathcal{N}(\hat{\mathbf{x}}_{s+1} | \mu_{s+1|s}, P_{s+1|s}) = \mathcal{N}(\mathbf{z}_{s+1} | h(\mu_{s+1|s}), S_{s+1}) \end{aligned} \quad (16)$$

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

$$LL(\Theta) = \left\langle \sum_{s=0}^3 \log p(\mathbf{z}_{s+1} | \mathbf{z}_1, \dots, \mathbf{z}_s; \Theta) \right\rangle_{\text{trials}} \quad (17)$$

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

$$\Theta_{\text{ML}} = \text{argmax}_{\Theta} LL(\Theta)$$

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

**Incorporating trials without intermediate stopping.** 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 Eq. (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:

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

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  $\Theta = \Theta_{\text{ML}}$  and the trajectory parameters  $\Delta\mathbf{x}$  for each segment. Each repetition generates stochastic observations  $\hat{d}_{\text{model}}$  and  $\hat{\varphi}_{\text{model}}$  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  $\bar{\beta}$ , multiplicative velocity gain  $\alpha$  and noise  $\bar{\sigma}_0^2$ . The additive bias  $\bar{\mathbf{b}}$  and reporting noise parameters  $\kappa = (\sigma_d^2, \sigma_{\varphi}^2)$  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  $\theta = (\bar{\beta}, \alpha, \bar{\mathbf{b}}, \bar{\sigma}_0^2)$  but the reporting noise parameters  $\kappa = (\sigma_d^2, \sigma_{\varphi}^2)$  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-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  $\kappa = (\sigma_d^2, \sigma_{\varphi}^2)$  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

$$\begin{aligned} \mu_{s+1} &= \mu_s e^{-\beta|\Delta\mathbf{x}|} + \left( \alpha \frac{\Delta\mathbf{x}}{|\Delta\mathbf{x}|} + \bar{\mathbf{b}} \right) \frac{1}{\beta} (1 - e^{-\beta|\Delta\mathbf{x}|}) \\ \sigma_{s+1}^2 &= \bar{\sigma}_0^2 \end{aligned}$$

Note that  $\mu_{s+1}$  depends on the previous predicted mean  $\mu_s$  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

**Table 1** Number of parameters for different models.

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

stopping points:

$$\begin{aligned} LL_{\kappa=0}(\theta) &= \left\langle \sum_{s=0}^3 \log p(\tilde{\mathbf{x}}_{s+1}; \theta) \right\rangle_{\text{trials}} = \left\langle \sum_{s=0}^3 \log \mathcal{N}(\tilde{\mathbf{x}}_{s+1} | \mu_{s+1}(\theta), \sigma_{s+1}^2(\theta)) \right\rangle_{\text{trials}} \\ &= \left\langle -\sum_{s=0}^3 \frac{1}{2\sigma_0^2} (\tilde{\mathbf{x}}_{s+1} - \mu_{s+1}(\theta))^2 - \frac{1}{2} \log(2\pi\sigma_0^2) \right\rangle_{\text{trials}} \end{aligned}$$

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.<sup>17,18</sup>, 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  $\tilde{\mathbf{x}}_{s+1}$ , and both predictive and updated distribution in Eq. (13) are equal to the prior distribution:

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

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

$$\begin{aligned} \mu_{s+1} &= f(\mu_s) = \mu_s e^{-\beta|\Delta x|} + \left( \alpha \frac{\Delta \mathbf{x}}{|\Delta \mathbf{x}|} + \tilde{\mathbf{b}} \right) \frac{1}{\beta} (1 - e^{-\beta|\Delta x|}) \\ S_{s+1} &= H_{s+1} \Sigma_x H_{s+1}^T + \Sigma_z \end{aligned}$$

where

$$\begin{aligned} \Sigma_x &= \tilde{\sigma}_0^2 I_2 \\ \Sigma_z &= \begin{pmatrix} \sigma_d^2 & 0 \\ 0 & \sigma_\phi^2 \end{pmatrix} \\ H_{s+1} &= \left. \frac{\partial h(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mu_{s+1}} = \frac{1}{|\mu_{s+1}|^2} \begin{pmatrix} \mu_{s+1,1} & \mu_{s+1,2} \\ -\mu_{s+1,2} & \mu_{s+1,1} \end{pmatrix} \end{aligned}$$

The log-likelihood is approximated as

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

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

$$\begin{aligned} \hat{d} &= d + \sigma_d \eta_d \\ \hat{\phi} &= \phi + \sigma_\phi \eta_\phi \end{aligned} \tag{18}$$

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

$$\mathbf{z}_{s+1} = \begin{pmatrix} \hat{d} \\ \hat{\phi} \end{pmatrix} \tag{19}$$

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

$$h(\tilde{\mathbf{x}}_{s+1}) = \begin{pmatrix} d(\tilde{\mathbf{x}}_{s+1}) \\ \phi(\tilde{\mathbf{x}}_{s+1}) \end{pmatrix} = \begin{pmatrix} |\tilde{\mathbf{x}}_{s+1}| \\ \text{atan2}(|\tilde{\mathbf{x}}_{s+1}|_2, |\tilde{\mathbf{x}}_{s+1}|_1) \end{pmatrix} \tag{20}$$

and the Jacobian  $H_{s+1}$  in Eq. (15) by

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

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

$$\begin{aligned} f(\tilde{\mathbf{x}}_s) &= \mu_{s+1}(\tilde{\mathbf{x}}_s) = \tilde{\mathbf{x}}_s e^{-\beta\Delta t} + \left( \alpha \frac{\Delta \mathbf{x}}{\Delta t} + \tilde{\mathbf{b}} \right) \frac{1}{\beta} (1 - e^{-\beta\Delta t}) \\ \Sigma_x &= \sigma_{s+1}^2 I_2 = \frac{\sigma_0^2}{2\beta} (1 - e^{-2\beta\Delta t}) I_2 \end{aligned} \tag{21}$$

where  $\Delta t$  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

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

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

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

**Model comparison.** The 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<sup>62,63</sup>. 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

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

where  $n$  is the number of observations and  $k$  is the number of parameters. The number of parameters for different models is listed in Table 1. All Bayesian analyses were carried out using flat priors.

In addition to BIC, we compare models using LOOCV. Given  $T$  trajectories for each model and participant, 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  $\text{LL}_k$  is the log-likelihood corresponding to the  $k$ -th trajectory, and  $\Theta_{\text{ML}}^{-k}$  are the ML parameters on the training set excluding the  $k$ -th trajectory.

**Relative influence of model parameters.** 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:

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

$$\text{error}_{\alpha}^2 = \text{error}^2(\tilde{\beta}_{\text{ML}}, \alpha = 1, \tilde{\mathbf{b}}_{\text{ML}}, \tilde{\sigma}_{0_{\text{ML}}}^2, \sigma_{d_{\text{ML}}}^2, \sigma_{\phi_{\text{ML}}}^2)$$

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

$$\text{error}_{\sigma_0}^2 = \text{error}^2(\tilde{\beta}_{\text{ML}}, \alpha_{\text{ML}}, \tilde{\mathbf{b}}_{\text{ML}}, \tilde{\sigma}_0^2 = 0, \sigma_{d_{\text{ML}}}^2, \sigma_{\phi_{\text{ML}}}^2)$$

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

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

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  $\alpha_{\text{ML}} > 1$  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%.

**Statistics and reproducibility.** Statistical comparisons were carried out using a one-sided permutation test with 10,000 permutations, unless otherwise noted. The experiment was conducted one time (i.e., the experiment was not repeated with an independent sample and consequently the obtained results were not replicated). However, the behavioral results of our experiment (path integration performance in young versus older adults) are widely in line with previous studies, and several different cross-validation approaches have led to similar performance of our computational model in “unseen data”, suggesting that the results of this work can be reproduced also with independent datasets.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The data underlying the results of this manuscript are available for download: <https://osf.io/ufk4x/>. A reporting summary for this Article is available as a Supplementary Information file.

## Code availability

The custom computer code and algorithms used to generate results that are reported in the paper are available from the corresponding authors upon reasonable request.

Received: 20 December 2018; Accepted: 20 March 2020;

Published online: 26 May 2020

## References

- Wang, R. F. Building a cognitive map by assembling multiple path integration systems. *Psychon. Bull. Rev.* **23**, 692–702 (2016).
- Buzsáki, G. & Moser, E. I. Memory, navigation and theta rhythm in the hippocampal–entorhinal system. *Nat. Neurosci.* **16**, 130–138 (2013).
- He, Q., McNamara, T. P., Bodenheimer, B. & Klippel, A. Acquisition and transfer of spatial knowledge during wayfinding. *J. Exp. Psychol. Learn. Mem. Cogn.* **45**, 1364–1386 (2019).
- Lester, A. W., Moffat, S. D., Wiener, J. M., Barnes, C. A. & Wolbers, T. The aging navigational system. *Neuron* **95**, 1019–1035 (2017).
- Etienne, A. S. & Jeffery, K. J. Path integration in mammals. *Hippocampus* **14**, 180–192 (2004).
- Britten, K. H. Mechanisms of self-motion perception. *Annu. Rev. Neurosci.* **31**, 389–410 (2008).
- Bassett, J. P. & Taube, J. S. Neural correlates for angular head velocity in the rat dorsal tegmental nucleus. *J. Neurosci.* **21**, 5740–5751 (2001).
- Biazoli, C. E., Goto, M., Campos, A. M. P. & Canteras, N. S. The supragenual nucleus: a putative relay station for ascending vestibular signs to head direction cells. *Brain Res.* **1094**, 138–148 (2006).
- Clark, B. J., Brown, J. E. & Taube, J. S. Head direction cell activity in the anterodorsal thalamus requires intact supragenual nuclei. *J. Neurophysiol.* **108**, 2767–2784 (2012).
- Butler, W. N. & Taube, J. S. The nucleus prepositus hypoglossi contributes to head direction cell stability in rats. *J. Neurosci.* **35**, 2547–2558 (2015).
- Cullen, K. E. The vestibular system: multimodal integration and encoding of self-motion for motor control. *Trends Neurosci.* **35**, 185–196 (2012).
- Angelaki, D. E. & Cullen, K. E. Vestibular system: the many facets of a multimodal sense. *Annu. Rev. Neurosci.* **31**, 125–150 (2008).
- Loomis, J. M. et al. Nonvisual navigation by blind and sighted: assessment of path integration ability. *J. Exp. Psychol. Gen.* **122**, 73–91 (1993).
- Klatzky, R. L., Beall, A. C., Loomis, J. M., Golledge, R. G. & Philbeck, J. W. Human navigation ability: tests of the encoding-error model of path integration. *Spat. Cogn. Comput.* **1**, 31–65 (1999).
- Chrastil, E. R., Sherrill, K. R., Aselcioglu, I., Hasselmo, M. E. & Stern, C. E. Individual differences in human path integration abilities correlate with gray matter volume in retrosplenial cortex, hippocampus, and medial prefrontal cortex. *eNeuro* **4**, 0346–16 (2017).
- Brunton, B. W., Botvinick, M. M. & Brody, C. D. Rats and humans can optimally accumulate evidence for decision-making. *Science (80-)* **340**, 95–98 (2013).
- Lappe, M., Jenkin, M. & Harris, L. R. Travel distance estimation from visual motion by leaky path integration. *Exp. Brain Res.* **180**, 35–48 (2007).
- Lappe, M., Stiels, M., Frenz, H. & Loomis, J. M. Keeping track of the distance from home by leaky integration along veering paths. *Exp. Brain Res.* **212**, 81–89 (2011).
- Lakshminarasimhan, K. J. et al. A dynamic Bayesian observer model reveals origins of bias in visual path integration. *Neuron* **99**, 194–206 (2018).
- Allen, G. L., Kirasic, K. C., Rashotte, M. A. & Haun, D. B. M. Aging and path integration skill: kinesthetic and vestibular contributions to wayfinding. *Percept. Psychophys.* **66**, 170–179 (2004).
- Mahmood, O., Adamo, D., Briceno, E. & Moffat, S. D. Age differences in visual path integration. *Behav. Brain Res.* **205**, 88–95 (2009).
- Adamo, D. E., Briceno, E. M., Sindone, J. A., Alexander, N. B. & Moffat, S. D. Age differences in virtual environment and real world path integration. *Front. Aging Neurosci.* **4**, 26 (2012).
- Harris, M. A. & Wolbers, T. Ageing effects on path integration and landmark navigation. *Hippocampus* **22**, 1770–1780 (2012).
- Bates, S. L. & Wolbers, T. How cognitive aging affects multisensory integration of navigational cues. *Neurobiol. Aging* **35**, 2761–2769 (2014).
- Izard, V. & Dehaene, S. Calibrating the mental number line. *Cognition* **106**, 1221–1247 (2008).
- Schmidt, R. A., Zelaznik, H., Hawkins, B., Frank, J. S. & Quinn, J. T. Motor-output variability: a theory for the accuracy of rapid motor acts. *Psychol. Rev.* **47**, 415–451 (1979).
- Jones, K. E., Hamilton, A. F. & Wolpert, D. M. Sources of signal-dependent noise during isometric force production. *J. Neurophysiol.* **88**, 1533–1544 (2002).
- Faisal, A. A., Selen, L. P. J. & Wolpert, D. M. Noise in the nervous system. *Nat. Rev. Neurosci.* **9**, 292–303 (2008).
- Burak, Y. & Fiete, I. R. Fundamental limits on persistent activity in networks of noisy neurons. *Proc. Natl Acad. Sci. USA* **109**, 17645–17650 (2012).
- Taube, J. S., Muller, R. U. & Ranck, J. B. Head-direction cells recorded from the postsubiculum in freely moving rats. II. Effects of environmental manipulations. *J. Neurosci.* **10**, 436–447 (1990).
- Taube, J. S. Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. *J. Neurosci.* **15**, 86 (1995).
- Knierim, J. J., Kudrimoti, H. S. & McNaughton, B. L. Place cells, head direction cells, and the learning of landmark stability. *J. Neurosci.* **15**, 1648–1659 (1995).
- Terrazas, A. et al. Self-motion and the hippocampal spatial metric. *J. Neurosci.* **25**, 8085–8096 (2005).
- Aghajan, Z. M. et al. Impaired spatial selectivity and intact phase precession in two-dimensional virtual reality. *Nat. Neurosci.* **18**, 121–128 (2015).

35. Petzschner, F. H. & Glasauer, S. Iterative Bayesian estimation as an explanation for range and regression effects: a study on human path integration. *J. Neurosci.* **31**, 17220–17229 (2011).
36. Petzschner, F. H., Glasauer, S. & Stephan, K. E. A Bayesian perspective on magnitude estimation. *Trends Cogn. Sci.* **19**, 285–293 (2015).
37. Compte, A., Brunel, N., Goldman-Rakic, P. S. & Wang, X. J. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb. Cortex* **10**, 910–923 (2000).
38. Brody, C. D., Romo, R. & Kepecs, A. Basic mechanisms for graded persistent activity: discrete attractors, continuous attractors, and dynamic representations. *Curr. Opin. Neurobiol.* **13**, 204–211 (2003).
39. Boucheny, C., Brunel, N. & Arleo, A. A continuous attractor network model without recurrent excitation: maintenance and integration in the head direction cell system. *J. Comput. Neurosci.* **18**, 205–227 (2005).
40. Wu, S., Hamaguchi, K. & Amari, S. Dynamics and computation of continuous attractors. *Neural Comput.* **20**, 994–1025 (2008).
41. Burak, Y. & Fiete, I. R. Accurate path integration in continuous attractor network models of grid cells. *PLoS Comput. Biol.* **5**, e1000291 (2009).
42. Angelaki, D. E. & Hess, B. J. M. Self-motion-induced eye movements: effects on visual acuity and navigation. *Nat. Rev. Neurosci.* **6**, 966–976 (2005).
43. Laurens, J. & Angelaki, D. E. The functional significance of velocity storage and its dependence on gravity. *Exp. Brain Res.* **210**, 407–422 (2011).
44. Kiani, R., Churchland, A. K. & Shadlen, M. N. Integration of direction cues is invariant to the temporal gap between them. *J. Neurosci.* **33**, 16483–16489 (2013).
45. Pinto, L. et al. An accumulation-of-evidence task using visual pulses for mice navigating in virtual reality. *Front. Behav. Neurosci.* **12**, 36 (2018).
46. Chaudhuri, R., Gerçek, B., Pandey, B., Peyrache, A. & Fiete, I. The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep. *Nat. Neurosci.* **22**, 1512–1520 (2019).
47. Stangl, M. et al. Compromised grid-cell-like representations in old age as a key mechanism to explain age-related navigational deficits. *Curr. Biol.* **28**, 1108–1115 (2018).
48. Gil, M. et al. Impaired path integration in mice with disrupted grid cell firing. *Nat. Neurosci.* **21**, 81–91 (2018).
49. Nasreddine, Z. S. et al. The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* **53**, 695–699 (2005).
50. Luis, C. A., Keegan, A. P. & Mullan, M. Cross validation of the Montreal Cognitive Assessment in community dwelling older adults residing in the Southeastern US. *Int. J. Geriatr. Psychiatry* **24**, 197–201 (2009).
51. Fujita, N., Klatzky, R. L., Loomis, J. M. & Golledge, R. G. The encoding-error model of pathway completion without vision. *Geogr. Anal.* **25**, 295–314 (1993).
52. Wiener, J. M., Berthoz, A. & Wolbers, T. Dissociable cognitive mechanisms underlying human path integration. *Exp. Brain Res.* **208**, 61–71 (2011).
53. Sadalla, E. K. & Montello, D. R. Remembering changes in direction. *Environ. Behav.* **21**, 346–363 (1989).
54. Oberlin, K. W. Variation in intensive sensitivity to lifted weights. *J. Exp. Psychol.* **19**, 438–455 (1936).
55. Gaydos, H. F. Sensitivity in the judgment of size by finger-span. *Am. J. Psychol.* **71**, 557 (1958).
56. Cornsweet, T. N. & Teller, D. Y. Relation of increment thresholds to brightness and luminance. *J. Opt. Soc. Am.* **55**, 1303–1308 (1965).
57. Fechner, G. *Elements of Psychophysics* (Holt, Rinehart and Winston, 1966).
58. Indow, T. & Stevens, S. S. Scaling of saturation and hue. *Percept. Psychophys.* **1**, 253–271 (1966).
59. Uhlenbeck, G. E. & Ornstein, L. S. On the theory of the Brownian motion. *Phys. Rev.* **36**, 823–841 (1930).
60. Pavliotis, G. A. *Stochastic Processes and Applications: Diffusion Processes, the Fokker–Planck and Langevin Equations* (Springer, 2014).
61. Thrun, S., Burgard, W. & Fox, D. *Probabilistic Robotics* (MIT Press, 2005).
62. Konishi, S. & Kitagawa, G. *Information Criteria and Statistical Modeling* (Springer, 2008).
63. Raftery, A. E. Bayesian model selection in social research. *Sociol. Methodol.* **25**, 111–163 (1995).

## Acknowledgements

MS, MR, and TW express their gratitude to Swantje Petersen, Falko Eckardt, Mareen Hanelt, Patrick Hauff, Anita Hökelmann, Marko Kirbach, Claudia Marx, Mona Reißberg, and Uwe Sobieray for their help with data acquisition, project administration, and technical assistance, and to Jonathan Shine for helpful discussions. 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.

## Author contributions

M.S., I.K., I.F., and T.W. conceptualized the work. M.S. and M.R. programmed and implemented the virtual reality path integration task. M.S. acquired the data. I.K. and I.F. developed the computational model for dissecting path integration errors. M.S. and I.K. analyzed the data, visualized the results, and drafted the manuscript. All authors edited the manuscript. I.F. and T.W. supervised the work.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-020-15805-9>.

**Correspondence** and requests for materials should be addressed to M.S. or I.K.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020