

Humans Construct Survey Estimates on the Fly from a Compartmentalised Representation of the Navigated Environment

Tobias Meilinger¹, Agnes Henson^{1,2}, Jonathan Rebane^{1,3}, Heinrich H. Bühlhoff¹ and Hanspeter A. Mallot⁴

¹ Max Planck Institute for Biological Cybernetics, Tübingen, Germany

² Leeds Beckett University, Leeds, UK

³ Stockholm University, Stockholm, Sweden

⁴ Cognitive Neuroscience, Eberhard-Karls-University Tübingen, Germany

tobias.meilinger@tuebingen.mpg.de,

hanspeter.mallot@uni-tuebingen.de

Abstract. Despite its relevance for navigation surprisingly little is known about how goal direction bearings to distant locations are computed. Behavioural and neuroscientific models proposing the path integration of previously navigated routes are supported indirectly by neural data, but behavioral evidence is lacking. We show that humans integrate navigated routes post-hoc and incrementally while conducting goal direction estimates. Participants learned a multi-corridor layout by walking through a virtual environment. Throughout learning, participants repeatedly performed pairwise pointing from the start location, end location, and each turn location between segments. Pointing latency increased with the number of corridors to the target and decreased with pointing experience rather than environmental familiarity. Bimodal pointing distributions indicate that participants made systematic errors, for example, mixing up turns or forgetting segments. Modeling these error sources suggests that pointing did not rely on one unified, but rather multiple representations of the experimental environment. We conclude that participants performed incremental on-the-fly calculations of goal direction estimates within compartmentalised representations, which was quicker for nearby goals and became faster with repeated pointing. Within navigated environments humans do not compute difference vectors from coordinates of a globally consistent integrated “map in the head”.

Keywords: Spatial learning, survey knowledge, virtual environments

1 Introduction

For survey estimates, such as novel shortcutting or distal pointing, navigators must utilise their memory to relate their current location to a distant, non-perceivable target location. A straight-forward mechanism to achieve this is read-out of current and target coordinates from an integrated memory and computation of a difference vector.

Such a computationally simple process does not depend on distance to target or familiarity with an environment, but requires that locations are represented within a coordinate system. Other behavioral [1, 2] and neuroscientific models [3–5] propose path integration of previously navigated routes, for example, in the form of a mental walk. These assumptions are supported by neural data such as successive activation of hippocampal place cells along a path to the goal [6] and medial temporal lobe activity correlating with goal distance [7, 8]. A related account does not assume mental walk, but constructing a mental model of a non-visible area by integrating successive corridors until a goal location is reached [9]. Both mental walk and mental model assume an increase in computation with route distance and both would predict speed-up with repeated survey estimation demands. In contrast to read-out from a cognitive map these two models are hereafter referred to as ‘constructive’.

In order to test these assumptions we conducted a learning experiment within an immersive virtual environment (VE) consisting of a route of eight corridors presented via a head-mounted display (HMD) (Fig. 1 and Fig. 2A). Participants repeatedly walked through the environment from start to end and back. Throughout learning, they were repeatedly asked to point from each of the nine route locations (Fig. 1B) to each other location. Participants either conducted pointing tasks interspersed throughout the whole environment familiarisation period, or later on in the navigational task. We used latency data as well as an analysis of systematic errors to draw conclusions about the alternative models explored above. Latency was measured to probe whether distance influenced location processing (i.e., read-out vs. constructive) and also whether learning was contingent on repeated survey task demands or mere exposure. The systematic error analysis characterised whether errors could be specifically modelled as non-random deviations from the correct pointing direction, such as on the basis of incorrect turns or deleting sections during integration.

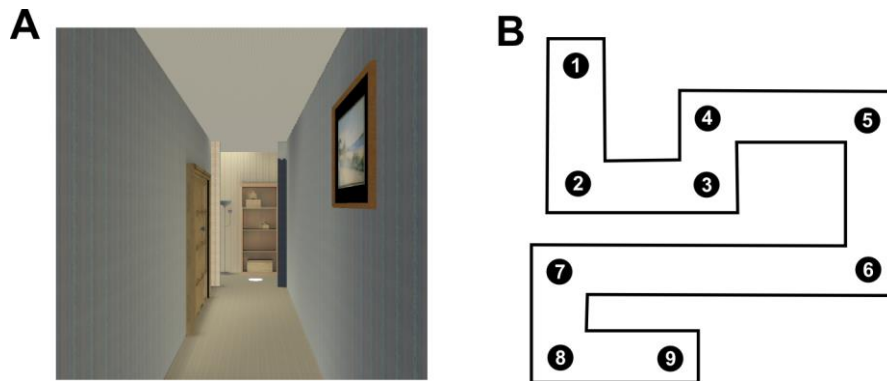


Fig. 1. Experimental environment (A) Participant's view. (B) Bird's eye view of environmental layout. Circles represent target and test locations for the pointing task numbered from start across the center of all turn to the end of the route.

2 Methods

2.1 Participants

Twenty-five participants (12 females and 13 males) aged between 23 and 65 ($M = 32.2$ years, $SD = 11$ years) participated in this experiment. They were randomly assigned to the *early pointing* condition (12 participants) and the *late pointing* condition (13 participants). We expected large effect sizes for the between group comparisons in addition with precise measurements due to many repetitions. Twelve participants per group should be sufficient to find them. All participants were recruited via a subject database, were paid for their participation, and signed an informed consent approved by the ethical committee of the University Clinics in Tübingen before participating in the experiment.

2.2 Material

The Virtual Environment. Figure 1 shows a snapshot of the environment as seen during walking, as well as a map of the route. The route consisted of a start and end-point, as well as seven turning locations along the route. During the first two learning trials, all nine of these ‘target locations’ were named by the experimenter as the participant arrived at the location. These target locations were named after salient landmarks at the locations. The locations were named as follows: Filing Cabinet, Bay Window, Mirror, Vase, Potted Plant, Bookcase, Painting, Grandfather Clock, Fish-bowl. The corridor design and environmental landmarks were distinct at each location, with sufficient information to identify and distinguish each location from one another.

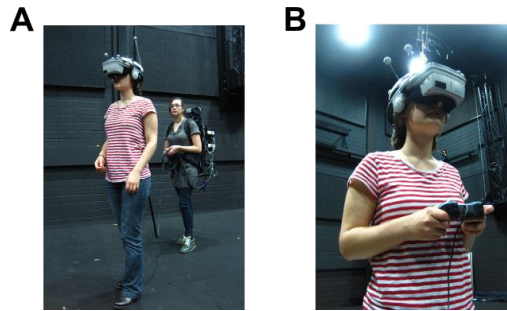


Fig. 2. Experimental Procedure. (A) A participant walking through the environment during the learning phase. (B) A participant performing the test phase using an input device. The persons on the photo consented to publication.

The Setup. Participants walked within a 12x12 meter space, of which the VE covered a 10x10 meter area. This allowed them to explore the space, without the possibility of walking into any obstacles and provided realistic proprioceptive and vestibular feedback, as well as efference copies while walking in VEs (see Fig. 2A). To obtain participants’ location in the space, their head position was tracked by 16 high-speed mo-

tion capture cameras at 120 Hz (Vicon® MX 13). This data was used to update the visualization of the VE. The visual surrounding at a location was rendered in real time (60Hz) using a NVIDIA Quadro FX 3700 graphics card with 1024 MB RAM in a standard laptop. Participants viewed the scene in stereo using an nVisor SX 60 head-mounted display that provided a field of view of 44x35 degrees at a resolution of 1280x1024 pixels for each eye with 100% overlap. This setup provided important visual depth cues such as stereo images and motion parallax.

2.3 Procedure

There were 10 walkthroughs in each direction, totaling 20 walkthroughs of the environment. Participants were instructed to follow a moving ball throughout the space in order to constrain exploration time. The virtual ball moved at an average speed of 1m per second, stopping to hover for 3 seconds over white circles on the ground at each turning location and at both ends of the environment.

During the *pointing task in the test phase*, participants were teleported to target locations. Here they were asked to successively point to all other target locations. During these trials, participants could look and rotate around, but not walk. This was enforced by placing participants in a circular handrail with 0.48 meter diameter to prevent them from leaving their location. Participants were asked to press a button on their gamepad once they recognised their location (Fig. 2B). The time required to “self-localise” was recorded for each participant. Participants were then instructed on the display to point to a named location, as if the walls were transparent. They were provided with a black midline through the display and informed to move their head until the line corresponded to the estimated target location. The name of the target location was displayed on the screen for each pointing. When participants believed they were facing the target, they pressed a button to move on to the next pointing. At each testing location all eight target locations were presented in a random order, that was newly determined for each location. No feedback on accuracy was given. After they had pointed to all targets from one location, participants were teleported to a new position. This was repeated, in random order, until participants had pointed to all target locations from all nine locations along the route. This resulted in 72 pointings per pointing task.

The final section of the *test phase* consisted of a *sequence task*. Participants were again transported to each location and were required to detail the turning sequence from that location to each end location by pressing the ‘left’ and ‘right’ keys on the gamepad corresponding to the turning sequence from their location to one of the end locations. This was collected for both directions from every location except the end locations themselves, and the penultimate locations before an end location. For these locations, only one sequence direction was recorded, as one end location is always visible for each penultimate location. Data from this task is not further reported here.

Participants were randomly assigned to two conditions, which dictated when they experienced the pointing task. Participants in the *early pointing group* were given the complete test phase (pointing task followed by sequence task) every four trials. Participants in the *late pointing condition* performed only the sequence section of the test

In order to identify such turning and forgetting errors, new mazes were generated in MATLAB by alternating single turn directions (e.g., at location 4 or 6 as in Fig 3) resulting in 7 alternative mazes or by eliminating legs (e.g., between location 3 and 4) resulting in 6 alternative mazes. Correct and alternative mazes were tested for best fit per subject and trial period on the basis of minimal absolute error. For each subject and period this was done three times, for all pointings together, as well as separately, for both forward and backward directions. Forward vs. backward can be considered as the simplest compartmentalization which keeps complexity of the pointing tasks constant (i.e., each pair of locations A and B occurs in each sub-set).

3 Results and Discussion

3.1 The generation of survey estimates

We used pointing latency as a function of target distance and familiarity to distinguish between read out from a cognitive map and constructive accounts (i.e., mental walk and mental model). A distance effect on pointing latency would support predictions of constructive strategies, since greater integration demands entail greater processing, either in form of a longer mental walk or a larger constructed mental model. However, a read-out strategy would assume constant cognitive effort for all estimations, and thus predict constant latency. Pointing latency indeed differed as a function of distance, $F(2.34, 161)=9.72$, $p<.001$, $\eta_p^2=.30$, and increased with distance to the target up to a distance of four corridors, $F's>8.1$, $p's<.010$, before plateauing (Fig. 4A). Error also differed, $F(2.22, 161)=39.8$, $p<.001$, $\eta_p^2=.63$, and increased up to six corridors, $F's>19.5$, $p's<.001$ (Fig. 4B). Similar error increase has been demonstrated previously [11] and might stem simply from learning, as larger travelled distance results in larger average errors [12]. While the current data do not differentiate between models, the latency increase with distance nevertheless aligns with the predictions of constructive theories of spatial processing.

Constructive theories assume an effortful estimation processes via walkthroughs or segmental integration; however, repeated estimations should result in speeded pointing due to increased familiarity with these active processes [13]. Importantly however, such training effects should occur only when such strategies are required (i.e., during pointing), not during navigation itself. Simple exposure to the environment does not make demands on these estimation practices. To test this prediction we compared two groups of navigators. The early pointing group completed a pointing task after four learning trials (i.e., walkthroughs through the environment), and were then tested after every four learning trials throughout the experiment. The late pointing group completed 16 learning trials before the first pointing task and were tested again after the 20th learning trial.

Pointing latency decreased with pointing experience (Fig. 4C), as opposed to experience navigating an environment. During their first pointing trial, early and late pointers pointed equally quickly, $F(1,23)=0.008$, $p=.929$, despite the greater environmental experience of late pointers. Afterwards, latencies decreased both for late poin-

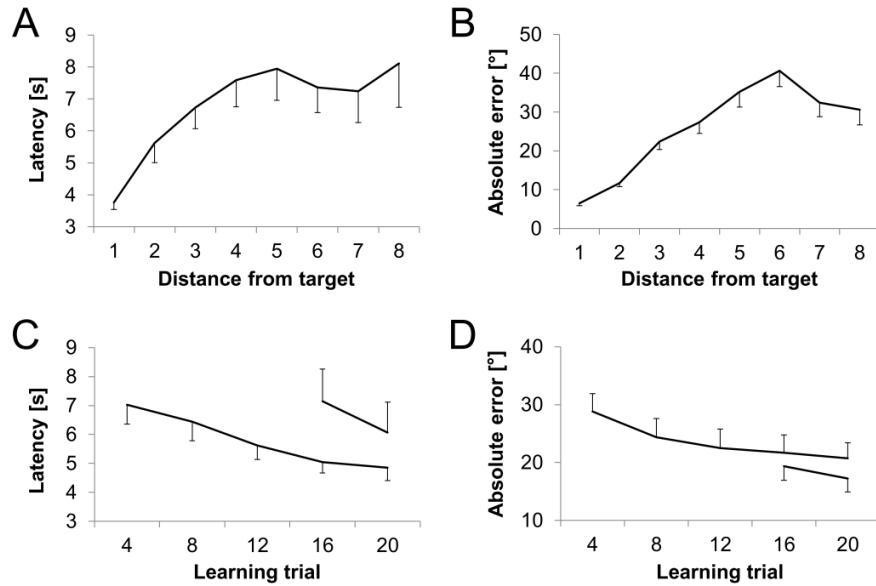


Fig. 4. Pointing results. (A) Latency and (B) pointing error as a function of distance (number of corridors). (C) Latency and (D) absolute pointing error as a function of learning trials for early and late pointers. Error bars indicate 1 standard error as estimated from the marginal means.

ters, $F(1,12)=38$, $p<.001$, and for early pointers up to the 16th learning trial, $F(1,11)'s>4.24$, $p's<.065$. We also observed improvement from the first to the second half of the first pointing test, $F(1,23)=5.51$, $p=.028$. Comparing the first halves, after 16 learning trials (i.e., same familiarity with the environment), the early learners pointed more quickly, $F(1,23)=5.68$, $p=.026$, profiting from their prior pointing experience. Pointing error (Fig. 4D), however, was inversely related to familiarity with the environment, rather than experience in pointing estimates. The error of learning groups did not differ at the same level of environmental familiarity, $F(1,23)=0.363$, $p=.553$, although late pointers were more accurate at their first pointing than early pointers, $F(1,23)=6.01$, $p=.022$. This indicates that the quality of the representation improved with navigation, though the estimation process itself only improved with repeated opportunities to point. This dissociation of pointing errors and latencies also excludes an explanation that larger errors result in lower confidence, which results in longer latencies. Late pointers show low error, but high latencies which then cannot be explained by low confidence.

While these results strongly support the predictions of constructive theories of survey estimation, we also propose that this procedure is consistent with accounts of the biological mechanism of place cell pre-play. Pre-play is a process that has been documented in place cells which identify locations, for example, along a route. During pre-play place cells along a route successively fire as if the animal was walking along the route towards the goal while being physically located at one spot [6]. Pre-play is

more consistent with an account of integration via mental walk. Mental walk would assume that before a navigator conducts a survey task, they mentally walk through an environmental representation, where place cells encode self-location [3]. This would explain the distance effect in time, as mentally walking longer distances should also result in longer estimation times. However, in order to account for latency reduction as a function of testing, the mental walk account would be forced to conceptually decouple mental and physical walking.

The cognitive demand initially burdening latency could conceivably be due to either spatial processing or handling the pointing task. Handling is unlikely to have significantly increased cognitive load, as the task of aligning a black line to a target was basic, and no subject experienced issues performing the task. Handling demands would also predict a parallel decrease in both error and latency over learning, while both dissociate for late pointers. Finally, the plateau that occurs in pointing latency as a function of distance could reflect estimation based on a truncated number of segments, since further estimation will not have a large impact on pointing accuracy. Alternatively, there could be cognitive capacity limitations that prevent more distant estimations.

Pointing latency improvement with experience could originate from an enhanced estimation process or from recalling previous pointing responses, thereby skipping the estimation process. While memorising vectors read out from a map can similarly explain such a latency reduction, the difference vectors predicted by this account cannot explain latency increase with target distance

3.2 Compartmentalization in spatial memory

In addition to the underlying process of survey navigation as indicated in the latency data, we wished to probe the types of representations formed during such tasks via systematic errors.

Most pointings showed a distribution centered around the correct target direction as indicated by significant circular v -tests (Fig. 3 and 5). However, there were clear deviations to that pattern at least at the first testing in 14% of location-target pairs (i.e., 2→7, 3→7, 2→8, 3→8, 1→9, 2→9, 3→9 and 4→9). Rayleigh tests demonstrated that in all cases this was not due to uniform data, but instead a result of an additional Von Mises distribution peak at another location (p 's < 0.01). It is unlikely that such bimodal distributions were merely the result of quantitative error. Please note that for bimodal distributions with close-by peaks (e.g., Fig. 3 right side) this analysis will not identify deviations from the predicted orientation.

To investigate whether these systematic errors originated from one representation or multiple underlying representations we compared forward vs. backward navigation. Figure 6 shows such mazes which were fitted per subject and learning trial to the pointing data separately for forward and backward pointing. We then compared distributions of best fitting mazes using Fisher tests. Only the pointings across leg segments of three and longer were included in the maze fittings, as no systematic errors were observed for the two leg cases.

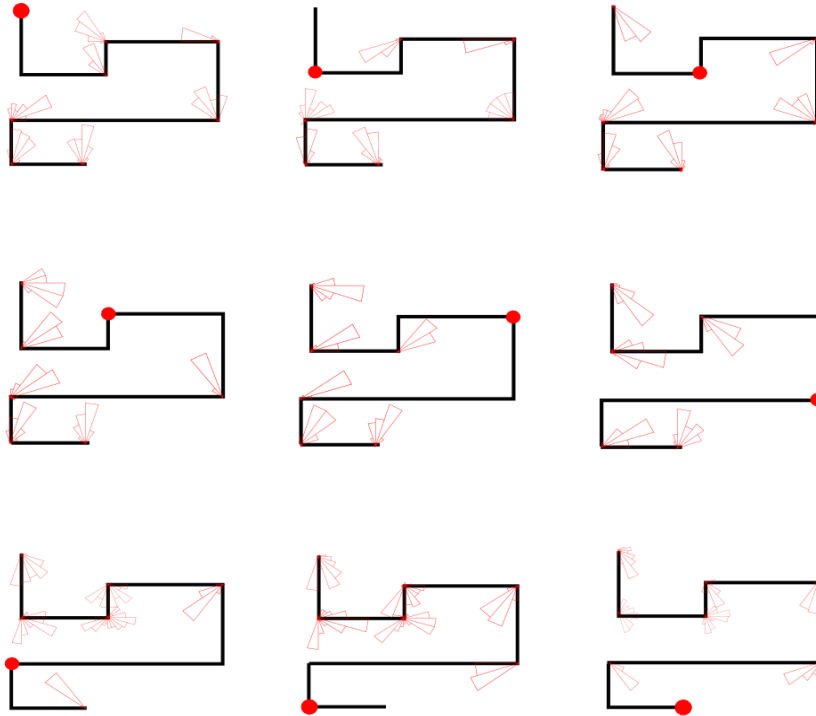


Fig. 5. Pairwise pointings. Solid circles denote correct point location. Circular histograms plots from early and late pointing groups are shown at pairwise pointing locations between selected landmark locations and the desired goal.

We observed deviations from the correct map in 40% of the mazes for forgetting one corridor leg and 29% for single turn errors. In both cases forward and backwards maze fits differed (both $p < .001$, Fig. 6). This suggests that forward and backward pointing estimates relied on representations that contained different systematic errors and were thus qualitatively different. This representational separation was stable throughout the experiment; it was observed after the 8th, 16th and 20th learning trials for forgetting error mazes and after the 4th and 20th learning trial for turn error mazes, $p < .05$. This indicates that navigators require more experience than the 20 walkthroughs (20 minutes learning) to equalise systematic errors in representations for forward and backward walking.

Mazes were fitted for a combined forward backward representation when making a single turn error and forgetting one leg. Both forward and backwards pointings for each subject and trial period were included for analysis. Fits with separate mazes reduced the error in average by 24% for forgetting and 18% for turns ($p < .001$). Fit was greater than for randomly shuffled data ($p < .001$) as well as fitting one single map only ($p < .009$). In general, fitting the data to new forward and backwards mazes considerably reduced the error. However a significant amount of error still existed after maze fitting, which could perhaps best be explained by allowing for more than

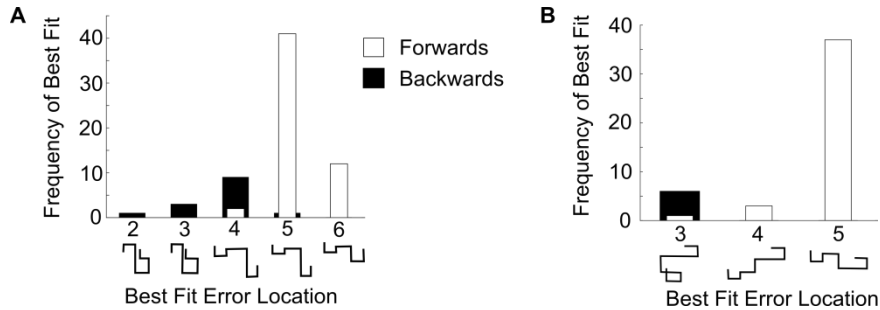


Fig. 6. Frequencies of best fit maze models. (A) Maze fits for forgetting error. (B) Maze fits for single turn errors. Below the x-axis are the types of mazes formed assuming an error of the corresponding type was made at the designated location (i.e., 3 would denote the 3rd landmark location from the start of the maze as indicated in Fig. 1B).

one single error in the maze, and combining forgetting and turning errors as well as incorrect leg-length estimation or handling errors, and working memory limitation in generating a survey estimate.

It may be possible that subjects made systematic errors as a result of being unable to self-localise. However, the pointing distributions for two leg-length segments were all unimodal with relatively low variance (Fig. 5), meaning that subjects were accurate in assessing the current and adjacent target locations. It is unlikely that having to point further resulted in subjects incorporating positional error. This suggests that self-localization error was relatively low compared to other sources of error.

We used forward vs. backward pointing mainly because comparing both representations results in an overall similar route complexity as each pair A-B occurs in each representation. However, the observed direction dependent division was also observed in previous priming, route choice and place cell firing [14–16]. Present results extend such a separation also to survey estimations.

Our results are based on one specific environment tested. It is conceivable that, for example, the observed error dissociation between forward and backward pointing is only found in environments with an overall U-shape (in which the examined systematic errors have large effects), but not so much in rather linear routes. Future experiments have to evidence whether the obtained results depend on a specific layout of our environment or also generalises to others.

3.3 Implications for models of spatial memory

Returning to the three proposed models of survey estimation, the evidence would firstly not point to a read-out model, due to clear contradictions with our results. Read-out does not predict the observed distance effect, nor the familiarity or compartmentalization effects. Mental walk model [3] can be best connected with animal literature such as pre-play [6]. Mental walk model would predict the distance effect and can account for the observed familiarity, compartmentalization and directional

effects, though would require an explicit dissociation between learning (physical navigation) and estimation processes (mental walk), as latency only improves with the latter rather than the former. This is in line with findings from single cell recording, which show that in similar corridor environments, different place cells are active depending on the direction of walking [16]. Clusters of direction-specific place cells might form separate maps for forward and backward pointing. Finally, the mental model [9] is consistent with all data. It is accurate in describing the distance effect through mental model construction, and can explain performance improvement with repeated model construction through pointing that is independent from walking itself. It also assumes that the underlying representation is compartmentalised (e.g. one representation for each corridor [17–19]), thus permitting different errors resulting from unique forward and backward connections between parts. The largest downside of the mental model proposition is that it cannot be connected to animal literature, although such survey estimates might be specifically human [20, 21], with other species not possessing the cognitive capacity to build a mental model of distant locations.

In conclusion, we claim that the formation of survey estimates in humans occurs incrementally at response, rather than occurring via read-out from a cognitive map. This is suggested, first, by the increase in pointing latency with distance, and second, by decreasing latency as a function of pointing training, as opposed to familiarity with the environment. Furthermore, the bimodal distribution of pointings, suggests that human survey knowledge incorporates both quantitative errors (e.g., acquired through path integration during learning), and systematic errors (i.e., those which originate from mixing up or forgetting elements of the route walked). The error modelling provided evidence to suggest there might be no unified underlying representations for survey navigation tasks in humans. This was indicated by different model fits for forwards and backwards representations. Results are clearly inconsistent with reading out coordinates from a globally consistent survey representation. Instead, they are consistent with constructive accounts, such as mental walk and mental model.

4 Acknowledgements

This work was supported by the DFG grant ME 3476/2-1. We would like to thank Joachim Tesch for help with the virtual environment setup.

5 References

1. Poucet, B.: Spatial cognitive maps in animals: new hypotheses on their structure and neural mechanisms. *Psychol. Rev.* 100, 163–182 (1993).
2. Wang, R.F.: Building a cognitive map by assembling multiple path integration systems. *Psychon. Bull. Rev.* 23, 692–702 (2016).
3. Byrne, P., Becker, S., Burgess, N.: Remembering the past and imagining the future: a neural model of spatial memory and imagery. *Psychol. Rev.* 114, 340–75 (2007).

4. McNaughton, B.L., Battaglia, F.P., Jensen, O., Moser, E.I., Moser, M.B.: Path integration and the neural basis of the “cognitive map.” *Nat Rev Neurosci.* 7, 663–678 (2006).
5. Sanders, H., Rennó-Costa, C., Idiart, M., Lisman, J.E.: Grid cells and place cells: an integrated view of their navigational/memory function. *Trends Neurosci.* 38, 1–13 (2015).
6. Pfeiffer, B.E., Foster, D.J.: Hippocampal place-cell sequences depict future paths to remembered goals. *Nature.* 497, 74–9 (2013).
7. Morgan, L.K., Macevoy, S.P., Aguirre, G.K., Epstein, R.A.: Distances between real-world locations are represented in the human hippocampus. *J Neurosci.* 31, 1238–1245 (2011).
8. Howard, L.R., Javadi, A.H., Yu, Y., Mill, R.D., Morrison, L.C., Knight, R., Loftus, M.M., Staskute, L., Spiers, H.J.: The hippocampus and entorhinal cortex encode the path and euclidean distances to goals during navigation. *Curr. Biol.* 24, 1331–1340 (2014).
9. Meilinger, T.: The network of reference frames theory: a synthesis of graphs and cognitive maps. In: Freksa, C., Newcombe, N.S., Gärdenfors, P., and Wöfl, S. (eds.) *Spatial Cognition VI*. pp. 344–360. Springer, Berlin (2008).
10. Zar, J.H.: *Biostatistical Analysis*. Prentice Hall, Upper Saddle River (2010).
11. Thorndyke, P.W., Hayes-Roth, B.: Differences in spatial knowledge acquired from maps and navigation. *Cogn. Psychol.* 14, 560–589 (1982).
12. Cheung, A., Ball, D., Milford, M., Wyeth, G., Wiles, J.: Maintaining a cognitive map in darkness: the need to fuse boundary knowledge with path integration. *PLoS Comput. Biol.* 8, e1002651 (2012).
13. Tse, D., Langston, R.F., Kakeyama, M., Bethus, I., Spooner, P.A., Wood, E.R., Witter, M.P., Morris, R.G.M.: Schemas and memory consolidation. *Science.* 316, 76–82 (2007).
14. Janzen, G.: Memory for object location and route direction in virtual large-scale space. *Q. J. Exp. Psychol.* 59, 493–508 (2006).
15. Stern, E., Leiser, D.: Levels of Spatial Knowledge and Urban Travel Modeling. *Geogr. Anal.* 20, 140–155 (1988).
16. Muller, R.U., Bostock, E., Taube, J.S., Kubie, J.L.: On the directional firing properties of hippocampal place cells. *J Neurosci.* 14, 7235–7251 (1994).
17. Marchette, S.A., Ryan, J., Epstein, R.A.: Schematic representations of local environmental space guide goal-directed navigation. *Cognition.* 158, 68–80 (2017).
18. Meilinger, T., Riecke, B.E., Bühlhoff, H.H.: Local and global reference frames for environmental spaces. *Q. J. Exp. Psychol.* 67, 542–569 (2014).
19. Meilinger, T., Strickrodt, M., Bühlhoff, H.H.: Qualitative differences between environmental and vista space memory are caused by the separation of space, not by movement or successive perception. *Cognition.* 155, 77–95 (2016).
20. Bennett, A.T.: Do animals have cognitive maps? *J. Exp. Biol.* 199, 219–24 (1996).
21. Grieves, R.M., Dudchenko, P.A.: Cognitive maps and spatial inference in animals: Rats fail to take a novel shortcut, but can take a previously experienced one. *Learn. Motiv.* 44, 81–92 (2013).