The Direction Bias and the Incremental Construction of Survey Knowledge

Tobias Meilinger (tobias.meilinger@tuebingen.mpg.de)
Max Planck Institute for Biological Cybernetics
Spemannstr. 38, 72076 Tübingen, Germany

Heinrich H. Bülthoff (heinrich.buelthoff@tuebingen.mpg.de)
Max Planck Institute for Biological Cybernetics
Spemannstr. 38, 72076 Tübingen, Germany
Department of Brain and Cognitive Engineering, Korea University,
Anam-dong, Seongbuk-gu, Seoul, 136-713 Korea

Abstract

This study examines how spatial memory acquired from navigation is used to perform a survey task involving pointing. Participants learned a route through a virtual city by walking it multiple times in one direction on an omnidirectional treadmill. After learning, they were teleported to several locations along the route, self-localized and pointed to multiple other locations along the route. Pointing was done away from or towards the current location. Preliminary data show that participants were faster in pointing away. This suggests that pointing was based on an incremental process rather than an all-at-once process which is consistent with mentally walking through a cognitive map or constructing a mental model of currently non-visible areas of the city. On average participants pointed faster to targets located further down the route towards the end than to targets located route upwards towards the start. Analysis of individual performance showed that more participants than expected by chance showed such an effect of target direction also in their pointing accuracy. The direction of this effect differed between participants. These direction biases suggest that at least some participants encoded the environmental space by multiple interconnected locations and used this representation also for pointing.

Keywords: Reference frame; environmental space; spatial memory; survey knowledge; cognitive map; mental walk; mental model; pointing; virtual environment

Introduction

When navigating through an environmental space such as a city or a building we experience multiple views of parts of this environment from various perspectives (Montello, 1993). The knowledge acquired from these experiences can be used to retrace familiar routes, plan novel routes, point to distant locations, or look for shortcuts. The last two tasks are examples of survey tasks (Ishikawa & Montello, 2006; Siegel & White, 1975). To solve a survey task, one has to consider metric relations (distance, relative direction) between two locations not mutually visible. Often, these two locations are our current location and a target location we want to point to, estimate the distance to, or find a shortcut towards. In order to do so at least our current location and the target location have to be represented within a single reference frame. This could be our egocentric reference frame within which the direction and distance of the target is represented in relation to our body. It could also be an allocentric world-centered reference frame within which our current location and the target are represented. Unless we obtain our environmental knowledge from a map which already provides this information within a single reference frame we have to integrate the multiple pieces of information acquired during navigation to represent them within one single reference frame. This work aims to cast some light on how this integration process might work. We will introduce theories of survey knowledge, derive predictions from these theories, and test them in an experiment.

Theories of Survey Knowledge

Most spatial memory theories which explain survey knowledge assume that navigators form a global world-centered reference frame within which all relevant locations are represented. Such a global reference frame might be formed very quickly, with all novel locations represented within it (Mou, McNamara, Valiquette & Rump, 2004; O'Keefe, 1991; Stachniss, 2009). Alternatively, this global reference frame is eventually formed from multiple local representations (Kuipers, 2000; Mallot & Basten, 2009; McNamara, Sluzenski & Rump, 2008; Poucet, 1993; Trullier, Wiener, Berthoz & Meyer, 1997). It then either works as an additional layer embedding local representations within a metric reference frame or as the top-level in a hierarchical memory structure subsuming lower level reference frames. In the following a global world-centered reference frame will be called a cognitive map.

Survey relations can be obtained from a cognitive map in several ways. The easiest way is to simply *read out* the coordinates of the relevant locations (e.g., the current location and the target location) and compute the relative direction, the distance between the locations, etc., by subtracting these coordinates from each other. If required by the task these parameters are then transformed into an egocentric reference frame, for example, when pointing to a target.

Alternatively, navigators could *mentally walk* through a cognitive map. While mentally moving from one point to another, they integrate the metric survey relation between the start and the mental position in the map until reaching the target (Byrne, Becker & Burgess, 2007). Thus the relative direction, the distance, etc. are derived. The activation pattern of hippocampal place cells is a plausible mediator for this process – although the conscious imagery of the mental walk might take place in posterior parietal cortex.

Place cells represent locations within an environment (O'Keefe & Nadel, 1978). Even in the absence of sensory stimulation (e.g., during sleep) they can fire in an ordered fashion as they would do when walking a route (Skaggs & McNaughton, 1996). Similar neural processes might happen during mental walks when performing a survey task.

A different position assumes that an environmental space is not represented within a single global reference frame (i.e., a cognitive map), but by multiple local interconnected reference frames (Meilinger, 2008). The integration within a single reference frame which is required for survey tasks happens during retrieval by constructing a *mental model* of the non-visible environment (a related model was presented for updating by Fujita, Klatzky, Loomis & Golledge, 1993). For example, navigators imagine what the environment would look like if the surrounding walls were transparent. First, they imagine the adjacent street from their current position, then they add the street branching off from it, etc. In this way all locations from the current location along a route leading towards the target location are imagined step-by-step within the current egocentric reference frame (this could also be done from a different imagined viewpoint). The mental model of the non-visible environment is constructed piecewise from a certain perspective. No one mentally walks through this constructed environment and the underlying memory structure is no cognitive map, but a network of reference frames interconnected by directed links (i.e., the links point in certain direction). The construction of the mental model is assumed to be easier when done along the direction of the links (i.e., imagine a distant location the link point towards). Otherwise these links have to be inverted which is computationally costly.

The Prediction of Performance Differences

The three positions, read out from a cognitive map, mentally walking through a cognitive map, and constructing a mental model from a network of reference frames predict specific performance differences due to incremental vs. all-at-once process of deriving survey relations and due to direction biases in the underlying memory.

All-at-once vs. Incremental Estimation of Survey Relations Reading out coordinates of two locations from a cognitive map and subtracting them is an all-at-once process in the sense that the survey relation (e.g., the relative direction of the target from a current location) is determined as whole. Contrary, mentally walking to a target or extending a mental model of the environment until it includes the target are incremental processes. The further we walk and the further the model is constructed the better we can estimate the direction and distance towards our target. Due to the incremental character locations in-between have to be represented during this process. This is not the case for reading out. One way to test this is to have navigators point to multiple locations in an ordered way. For example, they point to all locations along a route from the current location to a location B. When they do so in an order away from the current location (i.e., first point to the adjacent location, then the second closest, etc., until finally pointing to B) they can mentally walk or construct a model up to the first location, point there, extend this model or mentally walk to the second location, point there, etc., until mentally reaching location B. In the opposite case when they point in an order towards the current location (i.e., first to location B, then the second last location until finally pointing to the location closest to the current location) they first have to construct the whole model up to B, respectively mentally walk the entire distance up to location B. Then they either shift their attention to the second last target in the model, mentally walk back to the second last target or do it all over again from the current location to the second last location. No matter how navigators precisely do this, this process should last longer and/or be more error prone than pointing to targets in an order away from the current location. When reading out locations from a cognitive map navigators cannot profit from their last pointing. They have to compute the survey relation for each target individually no matter in which order they point to the locations. Order thus should not lead to different performance as in the case of a mental walk or a mental model.

The Direction Bias A cognitive map does not show direction specificity between locations (although the whole map might be oriented in a certain way such as north-up in a paper map). That means that no matter whether one points from A to B or from B to A the result should not differ in performance. This is just the same for reading out as well as for mental walk. On the contrary, a direction bias is expected in certain cases for the mental model explanation, because of the underlying memory. The mental model is based on directed interconnections between local reference frames. Constructing a model in the direction of the interconnection is easier as no inversion is required. It should yield better performance.

In order to predict performance differences one has to know where the directedness in memory originates from. According to Meilinger (2008) navigators encode local reference frames during navigation (e.g., corresponding to a street or a room). The interconnections between these local reference frames represent the metric relations (i.e., relative direction, distance, and orientation) between them. They can be derived in at least two ways. First, navigators might obtain interconnections from their visual input. They see that a street branches off to the right in 20 meters. The reference frame of this street is located 20 meters to the front and is oriented 90° to the right. This results in a forward connection, for example, expressed by vector pointing forwards. Alternatively, they could walk up to the next street while updating the origin of their current street (i.e., the origin of the memory reference frame representing the street). The interconnection to the last reference frame is the updated vector pointing back to the last street (i.e., a backwards interconnection). Here an individual navigator is expected to apply only one kind of strategy (i.e., either forward or

Figure 1: The virtual city as seen from navigation perspective (left side) and from bird's eye view with the route marked in red (right side). During learning the start, the end and each of the six intersections in-between were marked with white crosses on the floor. They worked as pointing locations and targets during the test phase.

backward encoding), at least over some time interval such as an experiment. Thus walking a route in one direction will result in directed interconnections (either forwards or backwards). Using these interconnections for constructing a mental model is easier along the direction of interconnections and should lead to better pointing performance. Depending on the encoding strategy this direction bias should be in forward or backwards direction.

Methods

For the experiment we used an immersive virtual city environment presented via a head-mounted display (HMD). In the learning phase, participants experienced the virtual environment by walking through it on an omnidirectional treadmill. They only walked the route in one direction. In the testing phase, participants were teleported to different locations in the environment, without walking physically. They were then asked to identify their location and heading and were instructed to point towards multiple targets on the route. Pointing order could be either towards their current location or away from it. Direction biases were examined by comparing pointing performance for pointing to targets located route upwards (to the start) with pointing performance for targets route downwards (to the end).

Participants

So far eleven participants (5 females and 6 males) aged between 21 and 34 (M = 26.6 years, SD = 4.5 years) participated in the experiment. They were recruited via a subject database and were paid for their participation. All participants signed an informed consent approved by an ethical committee before participating in the experiment.

Material

The Virtual City In the learning phase, participants had to learn a route through a virtual city. Figure 1 shows a snapshot of the city as seen during walking, as well as a bird's eye view of the route. The route consisted of a start, six intersections and an end. During learning, all eight locations were marked with a white X on the floor, as were all intersections visible from this route. The type of houses changed along the route, as did street width and the heights of houses. In addition, individual houses ensured sufficient landmark information to identify each location.

The Setup Participants walked on a 4x4 meters omnidirectional treadmill (Figure 2 left side). It allowed them to walk for infinite distances in any direction by moving them back to the centre of the treadmill. This unique interface allows for realistic proprioceptive and vestibular feedback as well as efference copies while walking in virtual environments. Participants wore a climbing harness for the unlikely event of falling and hurting themselves on the moving platform. To obtain participants' location on the treadmill their head position was tracked by 16 high-speed motion capture cameras at 120 Hz (Vicon® MX 13). This data was used both to control the treadmill and to update the visualization of the virtual environment. The visual surrounding at a location was rendered in real time (60Hz) using a NVIDIA Quadro FX 4600 graphics card with 768 MB RAM in a standard PC. Cables connected the PC to the display via the ceiling. Participants viewed the scene in stereo using a nVisor SX head-mounted display that provided a field of view of 44×35 degrees at a resolution of 1280×1024 pixels for each eye with 100% overlap. The setup thus also provided important visual depth cues such as stereo images and motion parallax.

Figure 2: The virtual reality setup. The left image depicts a participant walking on the omnidirectional treadmill during the learning phase. The right image shows a participant pointing to a target during the testing phase by facing the target and pressing a button on a gamepad.

Procedure

In the *learning phase*, participants walked the route at least six times from start to end. They were instructed to first learn the route, and secondly be able to self-localize when teleported to an X along the route after the learning phase. Participants were free to look around as long as they wanted. In their first run, they walked up to an intersection, looked around, and the experimenter pointed out the street to take when the participant looked down the correct street by stating "the route is this direction" (the experimenter was in the same room and could task with the participant). No verbal turning information (e.g., "left", "straight on", etc.) was given. When reaching the end and having looked around participants were teleported back to the start. From the second run onwards participants were asked to approach an intersection, look into the direction the route was going on and say "this way". The experimenter gave feedback whether this was right or wrong, before participants proceeded. They were not allowed to leave the route. For each new run, the virtual environment was rotated 90° clockwise relative to the lab. Sound sources within the lab could thus not be used to derive global orientation. The learning phase ended when participants walked the route at least six times and at least two runs were error-free. This criterion ensured comparable levels of route knowledge for all participants. Participants briefly trained walking on the treadmill before starting the experiment.

In the following *test phase*, participants were teleported to locations on the route formerly marked by an X (i.e., the start, the end or one of the six intersections in between). They were now asked to self-localize and then successively point to multiple target locations which had all been formerly marked by an X. For self-localization, participants could look and rotate around, but not walk around – a circular handrail around them with 0.48 meters diameter prevented them leaving their location during the test phase (Figure 2 right side). As soon as they subjectively knew their location and orientation, they were asked to press a button on a gamepad they were holding. Then they pointed to multiple targets. Pointing was done by turning on the spot until a vertical black line in the middle of the display matched the direction in which the participant thought the target was located. They thus would look directly at the target location if the surrounding houses were transparent. When participants thought they faced the target, they pressed a button and then pointed to the next target. No feedback was provided. After they had pointed to all targets, participants pressed a second button on the gamepad and were teleported to a new position.

Four conditions determined the targets and the order in which participants were asked to point towards them (Table 1). They should point either (1) first to the start and then to all locations between start and the current location in the order of walking (i.e, start, 1st intersection, 2nd intersection, etc.). (2) They should point to the same locations, but in reverse order (i.e., first the intersection before the current location, then the second last, etc. until finally pointing to the start). (3) They should point to the next intersection along the route after the current location, then the second next, etc. until pointing to the end. Or they should (4) point first to the end, then the last intersection, the second last intersection, etc. until pointing to the intersection after the current location. Consequently, we varied the two factors 'target direction' (route upwards to start vs. route downwards to end) and 'pointing order' (away vs. towards the current location; see Table 1). Please note that the adjacent intersections to point towards were always visible during pointing (although the Xs were removed). From the eight locations on the route (including start and end) participants pointed to every other location twice (in the order away and towards the current location). All 28 pointing sets were presented in random order for each participant (pointing downwards from seven locations, upwards from seven locations, both in two orders). This whole procedure was repeated resulting in 56 pointing sets altogether. After finishing a pointing set participants received feedback about the number of pointing targets they pointed towards: whether they pointed towards the right number of targets, how many targets they missed; or how many superfluous targets they pointed towards. No feedback about pointing accuracy was provided. Pointing sets with too few or too many pointings were not analyzed as the target locations could not be assigned to pointings. We recorded self-localization time (not further reported), pointing time and pointing direction for each pointing in a complete pointing set. After pointing participants drew a sketch map and we asked for subjective strategies with a questionnaire. The whole experiment lasted approximately two hours.

For the analysis we used pointing time and computed the absolute pointing error (i.e., the deviation between correct and estimated pointing direction irrespective of the direction of the error). Values deviating more than two standard deviations from a participant's mean were not analyzed. Only if a participant's mean absolute pointing error significantly differed from 90°, indicating that some survey knowledge was acquired, data were analyzed (90° error is the average error you get when randomly pointing in all directions). For analyses within participants we used t-tests. For analyses across participants we computed mean values per participant and condition and used within-participants ANOVA and t-tests. Cohens d and partial eta-square (η_p^2) are presented for the estimation of effect sizes.

Table 1: The Four Pointing Conditions

Target direction on the route	*Pointing order (relative to the current location)*	*Instruction: Point from…*
Upwards	Away	current location to start
Upwards	Towards	start to current location
Downwards	Away	current location to end
Downwards	Towards	end to current location

Results

For all but one participant pointing accuracy differed significantly from chance (t's > 10.9, p's < .001). They did acquire survey knowledge and were thus further analyzed. Their average absolute pointing error was 19.6°; mean pointing time was 2.8 seconds per pointing.

Mental walk and mental model theories of survey knowledge predicted performance differences for pointing order.

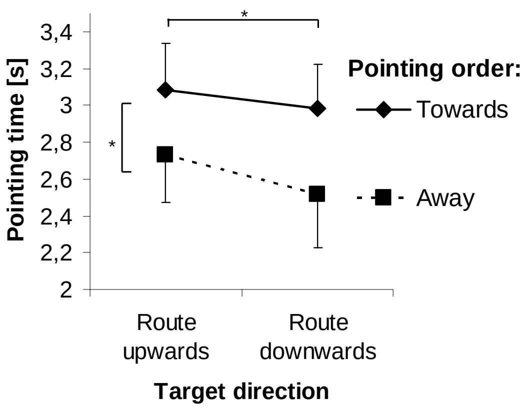

Figure 3: Average pointing time as a function of target direction and pointing order. Both main effects were significant as indicated by the asterisks. Means and (between participants) standard errors are displayed.

Indeed, participants pointed faster away (M = 2.6s) than towards the current location (M 3.0s; see Figure 3; $F(1, 9) = 9.50$, $p = .013$, $\eta_p^2 = .51$; accuracy: towards M = 22°, away M = 18°, $F(1, 9) = 2.21$, $p = .171$, $\eta_p^2 = .20$). This difference was not predicted by a process of reading out from a cognitive map.

According to the mental model of survey knowledge, participants' performance should differ as a function of target direction – although the direction of the effect might differ between participants. When averaging across all participants they pointed faster to targets located further down the route to the end (M = 2.7s) than to upward targets (M = *2.9s;* $F(1, 9) = 8.22$, $p = .019$, $\eta_p^2 = .48$; accuracy: upwards M = 17°, downwards M = 23°, $F(1, 9) = 2.21$, $p = .151$, $\eta_p^2 = .22$). Looking at the effect of target direction on pointing accuracy for each participant individually a more differentiated picture emerges: Six out of the ten participants showed an effect of target direction in their pointing accuracy (i.e., their pointing accuracy differed between pointings upwards to the start vs. downwards to the end t's > 1.99, p's < .049, d's > 0.19). Three of them pointed more accurately downwards the route (*M = 7.0° vs. M = 10.5°)*, three participants pointed more accurately upwards (*M = 23.6° vs. M = 45.1°)*. Four participants did not show a significant effect of target direction (upwards M = 16.2°, downwards M =,18.7°; t's < 1.88, p's > .063, d's < 0.21). If there was no target direction effect each participant has a chance of 5% to (erroneously) be classified as being direction biased in any direction by a t-test. The observed proportion of 6 out of 10 participants showing such an effect is highly unlikely to originate from such a 5% chance rate (binomial test $N = 10$, π = 5%: $p < .001$). Consequently, the null-hypothesis that there is no effect of target direction on accuracy is rejected. Since individual participants showed opposite direction biases, we observed no average global bias in pointing accuracy in one specific direction. When looking at differences in pointing time on the level of individual participants only one participant significantly pointed faster downwards the route ($t(207) = 2.29$, $p = .023$, $d = 0.22$). This proportion (one out of 10) does not significantly differ from a 5% chance rate (binomial test $N = 10$, π = 5%: $p = .401$).

We found no effect of pointing in walking order which is expressed by the interaction between target direction and pointing order (time and accuracy both $F(1, 9) < 1$). Pointing to multiple targets in walking order (i.e., from start to current location or from the current location to the end) did not differ significantly from pointing in opposite walking order (i.e., from end to current location or from current location to start).

Discussion

The present study examined predictions from three different theories about how survey relations are derived from spatial memory. The three positions (read out from a cognitive map, mentally walking through a cognitive map, and constructing a mental model from a network of reference frames) predict specific performance differences for target directions and pointing order.

We found an effect of pointing order. Participants pointed faster to targets in the order away from the current location than towards the current location. This result suggests that pointing is based on an incremental rather than an all-at-once process. Navigators might mentally walk through a cognitive map and integrate the walked distance (Byrne et al., 2007) or they could stepwise construct a mental model of the non-visible environment until this model includes the target (Meilinger, 2008).

There was also an effect of target direction. On average, participants pointed faster to targets further down the route, than to targets route upwards to the start. When looking at target direction effects for each individual, more participants than expected by chance showed a significant effect of target direction in their pointing accuracy. Half of these pointed more accurately towards locations further down the road, the other half pointed more accurately towards targets upwards the route. These results in pointing accuracy suggest different strategies in the encoding of an environment. Some participants might have encoded multiple local environments (e.g., rooms, streets, etc.), updated the last environment while walking to the next environment and stored the updated vector pointing backwards to the last environment. Deriving survey relations from this string of backwards connected locations should be easier in a backwards direction. For locations route downwards the connection would have to be inverted which is an additional process and thus an additional source of errors. Another group of participants seems to have encoded multiple local environments in the opposite direction (i.e., in the direction they walked the route). They could have derived the interconnections from their visual input: they saw how the route was going on (e.g., 30 meters straight on, then turn to the right) and used this information for connecting encoded locations, thus resulting in a forward connection. For them, constructing a mental model in forward direction did not involve inversion of the interconnection and thus resulted in more accurate pointing. The third group of participants did not show a significant effect of target direction on the level of the individual. They might have formed a cognitive map and used this representation for pointing (likely by mental walk). Alternatively, their orientation bias was not strong enough to reach the significance level. The time advantage for pointing route downwards when averaging across participants might simply be an effect of averaging across the groups and could suggest that forward encoding was more likely than backward encoding.

The results reported here were found in a virtual reality setup. Therefore, we cannot exclude the possibility that participants might behave differently in real environments. However, the present setup provided most of the bodily and visual cues also available when walking through a real environment (proprioceptive feedback, efference copy, vestibular stimulation, motion parallax, stereo vision, texture gradient, familiar size cues, etc.). Also, on average pointings were quite precise. A generalization to real environments does, thus, not seem implausible.

One major limitation is the small sample size. More participants are needed to see whether the effects observed are really stable. With more participants we will also be able to directly compare the different subgroups in target direction and have a closer look at their strategies.

This study examined how navigators derive survey relations used for pointing or short cutting from memory of an environmental space which they have to navigate through in order to experience it. Our results suggest that pointing is based on an incremental process as predicted by mentally walking a cognitive map or by constructing a mental model of the non-visual environment. At least for some participants we found indications for a direction specific encoding of such an environment (i.e., a string of location representations connected via directed links). Their pattern of performance is consistent with a mental model construction based on such a memory. Future experiments will have to clarify the exact circumstances which yield which kind of memory.

Acknowledgments

This research was supported by the DFG grant "The functional, computational and neural basis of human survey knowledge – comparing mental maps and mental graphs", the Max Planck Society and by the WCU (World Class University) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (R31-2008-000-10008-0). The authors thank Jan Souman for help in planning the experiment, discussing the results and writing the paper, Nadine Simon for help in data collection, Joachim Tesch, Michael Kerger, and Harald Teufel for intensive technical support, as well as Johanna Steffen for proof reading.

References

Byrne, P., Becker, S. & Burgess, N. (2007). Remembering the past and imagining the future: a neural model of spatial memory and imagery. *Psychological Review, 114,* 340-375.

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

Ishikawa, T. & Montello, D.R. (2006). Spatial knowledge acquisition from direct experience in the environment: individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology, 52,* 93-129.

Kuipers, B. (2000). The spatial semantic hierarchy. *Artificial Intelligence, 119,* 191-233.

Mallot, H.A., & Basten, K. (2009). Embodied spatial cognition: biological and artificial systems. *Image and Vision Computing, 27,* 1658-1670.

Meilinger, T. (2008). The network of reference frames theory: a synthesis of graphs and cognitive maps. In C. Freksa, N.S. Newcombe, P. Gärdenfors, & S. Wölfl (Eds.), *Spatial Cognition VI* (pp.344-360). Berlin: Springer.

McNamara, T.P., Sluzenski, J. & Rump, B. (2008). *Human Spatial Memory and Navigation.* In H.L. Roediger, III (Ed.), Cognitive Psychology of Memory. Vol. 2 of Learning and Memory: A Comprehensive Reference (pp.157-178). Oxford: Elsevier.

Montello, D. R. (1993). Scale and multiple psychologies of space. In A.U. Frank & I. Campari (Eds.), *Spatial information theory: A theoretical basis for GIS* (pp. 312-321). Berlin: Springer.

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

O'Keefe (1991). An allocentric spatial model for the hippocampal cognitive map. *Hippocampus, 1,* 230-235.

O'Keefe, J., & L. Nadel (1978). *The hippocampus as a cognitive map.* Oxford: Clarendon Press.

Poucet, B. (1993). Spatial cognitive maps in animals: Newhypotheses on their structure and neural mechanisms. *Psychological Review, 100,* 163-182.

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. In H. Reese, (Ed.), *Advances in Child Development and Behavior, Vol. 10,* (pp 10–55). New York: Academic Press.

Skaggs, W.E. & McNaughton, B.L. (1996). Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science, 271*, 1870-1871.

Stachniss, C. (2009). *Robot Mapping and Exloration.* Berlin: Springer.

Trullier, O., Wiener, S.I., Berthoz, A. & Meyer, J.-A. (1997). Biologically based artificial navigation systems: Review and prospects. *Progress in Neurobiology, 51,* 483-544.