

When do we integrate spatial information acquired by walking through environmental spaces?

Agnes Henson (agnes.henson@tuebingen.mpg.de)

Max Planck Institute for Biological Cybernetics
Spemannstr. 38, 72076 Tübingen, Germany

Hanspeter A. Mallot (hanspeter.mallot@uni-tuebingen.de)

Cognitive Neuroscience, Eberhard-Karls-University Tübingen,
Auf der Morgenstelle 28, 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

Tobias Meilinger (tobias.meilinger@tuebingen.mpg.de)

Max Planck Institute for Biological Cybernetics
Spemannstr. 38, 72076 Tübingen, Germany

Abstract

The present study examined whether spatial information of a novel environment was integrated within a reference frame during initial learning, or only later when required for pointing to other targets. Twenty-two participants repeatedly walked through a multi-corridor virtual environment, presented via a head-mounted display. At several stages within the learning process they were teleported to locations along the route and asked to self-localize and point to other locations. Pointing was faster during later tests as well as for closer targets, both of which might require less integration. Participants tested only after extended exposure (late pointers) took longer than participants who had received testing interspersed throughout the same amount of exposure (early pointers). Pointing latency did not differ between groups when comparing performance on their first pointing test, despite vastly different exposure. These results are inconsistent with the assumption that participants already integrated spatial information within a single reference frame during learning and simply accessed this information during testing. Rather, spatial integration is a time consuming process which is not necessarily undertaken if not required.

Keywords: Reference frame; environmental space; spatial integration; survey task; pointing; virtual environment

Introduction

When exploring a novel environmental space such as a city or building, navigators encounter various views of this space. Each location within the environment is experienced from an egocentric perspective, but for so called survey tasks such as shortcircuiting, pointing or straight-line distance estimations, these locations must be spatially integrated into a common reference frame (egocentric or allocentric). Spatial integration can be defined here as “the process of combining different spatial representations that have been

formed by multiple experiences within a single frame of reference or co-ordinate system” (Meilinger, Berthoz & Wiener, *in press*). For example, with regard to pointing, navigators must know where their target is relative to their current position, i.e., they must represent the target within the same reference frame as their body. The question asked here is ‘when does this integration happen?’.

Many theories concerned with spatial memory assume that when navigating a space, all spatial information is integrated within a single global reference frame, at least eventually, and can then be used for survey tasks (Byrne, Becker & Burgess, 2007; McNamara, Sluzenski & Rump, 2008; O’Keefe, 1991; Poucet, 1993). Some of these positions assume or imply that integration occurs during encoding or consolidation, regardless of whether navigators will use this knowledge for survey tasks or not. Spatial integration is thus independent from accessing this information. Alternatively, multiple representations acquired during navigation might also be kept separate in memory and only integrated when necessary, for example, when conducting a survey task (Meilinger, 2008). These two positions yield different predictions about the time it takes to perform a survey task, such as pointing as a function of (1) the amount of experience with an environment, (2) the amount of prior testing and (3) the distance towards a target.

On average, spatial knowledge increases with the amount of learning (Evans, Marrero & Butler, 1981; Gärling, Lindberg & Mantyla, 1983; Thorndyke & Hayes-Roth, 1982; but see Ishikawa & Montello, 2006). Repeated testing of navigators’ survey knowledge generally yields an increase in accuracy over learning. If spatial information is integrated during encoding, the acquired knowledge could simply be accessed during survey tasks, independently of

whether participants' survey knowledge was first tested early or late within the learning process. This would be due to the acquisition of survey knowledge being independent from its access. As such, navigators first tested after extended navigation should perform comparatively as well as navigators tested throughout the navigation process. Contrary to accuracy, which increases throughout learning, the *access* process for this knowledge should not change markedly across learning. In other words, participants' time conducting a survey task should remain more or less constant throughout the learning process, independent of early or late testing, or indeed the accuracy. There may be an increase at the beginning, if familiarizing with the task has an effect.

Alternatively, when memorising an environment, these multiple pieces of spatial information may remain separate until their integration becomes necessary, such as during a survey task (Meilinger, 2008). This predicts a difference in access of spatial information, which should be reflected in pointing latency during a survey task. This assumption is supported by evidence to suggest increased latency when integrating two reference frames with differing preferred reference directions (Greenauer & Waller, 2010). If multiple pieces of spatial information do remain disparate until required, the time needed for integration should not differ as a function of environmental experience, but with the amount of prior integration. No matter when the initial survey task is conducted- early or late during learning an environment – comparable time should be needed. After the initial survey task, later tasks should profit from this prior integration and performance time should decrease. Crucially, this approach does not predict *accuracy* as a function of completing a survey task. Increased experience of an environment allows increased accuracy of spatial information for each individual section and their pair-wise spatial relations amongst one another. In other words, both approaches predict increased accuracy with extended environmental experience, independently of when their survey knowledge is tested. However, predictions deviate when it comes to the process of accessing and compiling this information.

It should be noted that the current study restricted the length of time spent learning the environment. Integration costs during retrieval are only known to occur when controlling for learning time (Hanley & Levine, 1983, Yamamoto & Shelton, 2008); however, if given unrestricted learning time, this effect may not hold. This is especially true in an experimental environment, where often participants are motivated to perform as well as possible. Another issue is the amount of information to be integrated. When walking larger distances it is often the case that more spatial information must be integrated. More information to be integrated should result in larger errors (Thorndyke & Hayes-Roth, 1982). When integrating spatial information during testing, more spatial information to be integrated should also result in longer integration times and thus time to conduct a survey task. However, if integration happens during learning, no increase in pointing latency with

distance is expected. The spatial information would be integrated beforehand and would only have to be accessed.

In summary, if we integrate spatial information during learning, participants' pointing latencies should be more or less constant throughout the familiarization with an environment, whereas accuracy should increase with familiarity. Navigators with the same amount of learning experience should show comparative survey performance both for accuracy and latency. Pointing to targets further away should yield larger errors, but should take a comparable amount of time. However, if integration occurs while conducting a survey task, latency should be comparable for the first time a survey task was conducted. This should be independent of the learning experience in an environment and should decrease afterwards, as less spatial information would have to be integrated. Accuracy should increase with the amount of experience and should, at the same level of experience, not differ substantially between navigators with or without prior testing. Both pointing latency and error should increase with the distance to the target. None of the earlier mentioned studies measured latencies of survey tasks and were thus not able to distinguish between these two positions. The present study intends to close this gap, giving an opportunity to investigate when and why this integration occurs.

Method

In order to test the assumptions introduced in the introduction we conducted a learning experiment within an immersive virtual environment (VE), consisting of a set of corridors presented via a head-mounted display (HMD). In the *learning trials*, participants experienced the route-shaped VE by following a virtual ball through the 'corridors'. In the *test phase*, participants were required to complete a pointing task. Learning trials and a test phase repeatedly followed one another in order to measure the acquisition process. We also compared two groups of navigators. The *early pointing group* started the test phase after four learning trials and continued testing after every four throughout the experiment. Participants from the *late pointing group* completed 16 learning trials before the first pointing task and were tested again after the twentieth trial. Participants walked both directions but a single walkthrough, forwards or backwards, constituted a *learning trial* in its own right.

Participants

Twenty two participants (9 females and 13 males) aged between 23 and 65 ($M = 32.2$ years, $SD = 11$ years) participated in this experiment. 12 participants took part in the *early pointing* condition (5 males and 7 females) and 10 participants took part in the *late pointing* condition (7 males and 3 females). 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.



Figure 1: View from inside (left) and bird's eye view of environment. Target and test locations for the pointing task were identical and located at the center of a turn as well as at the start and the end of the route (nine locations altogether)

Material

The Virtual Environment. Figure 1 shows a snapshot of the environment as seen during walking, as well as a bird's eye view of the route. The route consisted of a start and endpoint, 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, Fishbowl. The corridor design and environmental landmarks were distinct at each location, with sufficient information to identify and distinguish each location from one another. For methodological issues related to learning a virtual environment by walking through it, please refer to Meilinger and Bülthoff (2010).



Figure 2: A participant walking though the environment during the learning phase (left) and the test phase (right).

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 Figure 2). To obtain participants' location in the space, their head position was tracked by 16

high-speed motion 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 head-mounted display that provided a field of view of 44x35 degrees at a resolution of 1280x1024 pixels for each eye with 100% overlap. The setup thus also provided important visual depth cues such as stereo images and motion parallax.

Procedure

There were 20 total walkthroughs of the environment, totaling ten walkthroughs for each direction. Participants' exploration time was constrained by a moving ball which they were instructed to follow through the environment. The virtual ball moved at an average speed of 1m per second, stopping only 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 on the route. They were then 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. After looking around and as soon as they subjectively knew their location, they were asked to press a button on their gamepad (Figure 2 right side). 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 and then pointed to the next target location. At each testing location all eight target locations were presented in a random order. 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 every time the pointing task was completed.

The final section of the *test phase* consisted of a *sequence task*. The *sequence task* involved participants being transported to each location again, but instead of pointing, they were required to detail the turning sequence from that location to each end location. This was achieved 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.

Participant trials were split into two conditions, which dictated when they experienced the pointing task. Twelve participants in the *early pointing group* were given the complete *test phase* (*pointing task* followed by *sequence task*) every four trials. Ten participants in the *late pointing condition* performed only the sequence section of the *test phase* after *learning trials* 4, 8 and 12. They eventually experienced the full *test phase* after 16 and then 20 trials. Half of participants in the *early pointing group* were informed prior to the learning phase that there would be a pointing task. The other half was not informed. This variation did not yield any performance differences and is not further reported here. Participants from the *late pointing group* were not informed that they would have to complete a pointing trial. Additional post hoc tasks and questionnaires are not reported here. The whole experiment lasted approximately 3.5 hours in the *early pointing* and 2.5 hours in the *late pointing group*. Participants were assigned randomly to conditions.

Analysis

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 three standard deviations from a participant's mean were not analyzed. Accuracy and latency were analyzed with a linear mixed model analysis (e.g. Snijder & Bosker, 1999) with the fixed factors learning trial (4, 8, 12, 16 & 20), distance to the target expressed as the number of corridors (1-8) and learning group (early vs. late pointing) where appropriate within a full factorial design (i.e. modeling all possible interactions). Compared to an ANOVA this analysis is less restrictive with regard to distribution assumptions and allows for varying effect sizes within different participants. Commonly accepted effect sizes for linear mixed models are not yet available. We thus report partial eta square η_p^2 derived from data aggregated per participant and the respective condition.

Results

The first analysis was concerned with an overall improvement over learning trials and with the distance to the target. In order to map the whole learning process only data from the early pointing group was analyzed within a linear mixed model analysis with the fixed factors of learning trial and target distance. As predicted by both positions regarding spatial integration during learning vs. testing, we found an effect in pointing accuracy for learning trials ($F(4, 4140) = 18.66, p < .01, \eta_p^2 = .55$) and for pointing distance ($F(7, 4137) = 178.58, p < .01, \eta_p^2 = .69$; interaction: $F(28, 4137) = 1.38, p = .09, \eta_p^2 = .13$) suggesting higher accuracy for shorter distances to the target and for more learning experience (Figure 3 top). Spatial integrating during testing, but not spatial integration during

learning predicted the same effects for pointing latency. Indeed there was a general effect of distance to the target ($F(7, 4137) = 87.51, p < .01, \eta_p^2 = .54$) and learning trial ($F(4, 4141) = 45.77, p < .01, \eta_p^2 = .60$) suggesting quicker responses after more learning trials and for closer targets. A significant interaction indicated that for larger pointing distances the order of successive learning trials sometimes reversed probably due to a higher variability ($F(28, 4137) = 1.84, p = .01, \eta_p^2 = .10$). The effect of distance on both pointing error and latency was also found in the further analyses, but will no longer be referred to.

In order to examine the effect of prior testing we compared performance for 'early' and 'late' pointing participants between trials 16 and 20 (fixed factors learning trial, target distance, and group). There was a significant interaction between learning trial and group on the pointing time ($F(1, 3057) = 9.92, p < .01, \eta_p^2 = .24$; main effect learning trial: $F(1, 3057) = 9.92, p < .01, \eta_p^2 = .26$; main effect group: $F(1, 134) = 3.35, p = .07, \eta_p^2 = .10$). In order to investigate this interaction, the differences between *early* and *late* groups were calculated at trials 16 and 20 respectively. As predicted by spatial integration during testing, participants in the early pointing group who could rely on prior integration pointed more rapidly at trial 16 than the late pointing participants who pointed for the first time ($F(1, 120) = 6.33, p = .01, \eta_p^2 = .14$). This difference

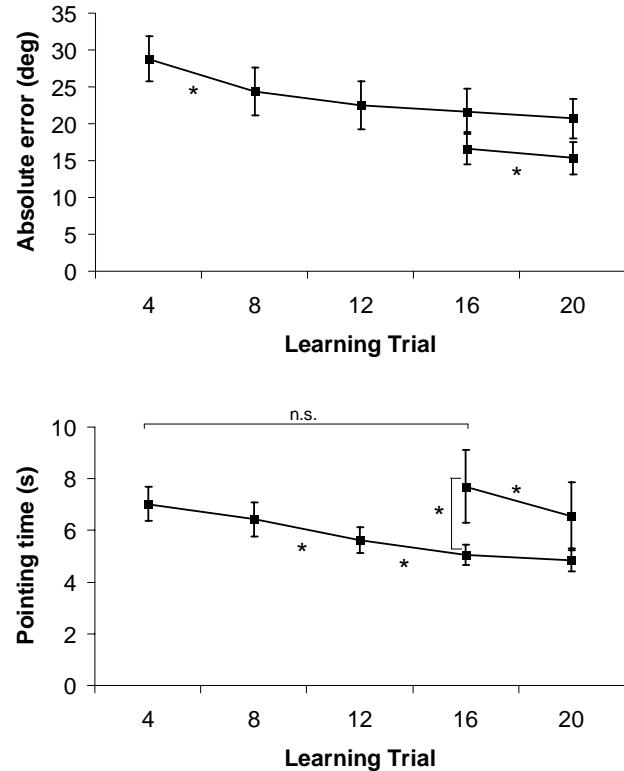


Figure 3: Mean absolute pointing error (top) and latency (bottom) across learning trials for early and late pointers. Error bars indicate ± 1 standard error. Asterisks * indicate significant differences in pair-wise comparisons between successive learning trials and other planned comparisons.

vanished with more extended pointing experience after learning trial 20 ($F(1, 21) = 2.93, p = .10, \eta_p^2 = .09$). Phrased differently, late pointers' pointing time decreased between their first and their second pointing test ($F(1, 1342) = 8.65, p < .01, \eta_p^2 = .44$), but not so for early pointers who showed no significant decrease between their fourth and fifth pointing test ($F(1, 1706) = .189, p = .66, \eta_p^2 < .01$). Please note that both groups had exactly the same amount of navigational experience with the environment.

Integration during learning as well as integration during testing both predicted that pointing accuracy was predicted by learning trial, but not so by early or late testing. Indeed we found no significant effect of group on pointing error (main effect group: $F(1, 134) = 3.35, p = .07, \eta_p^2 = .10$; main effect learning trial: $F(1, 3057) = 6.30, p = .01, \eta_p^2 = .27$; interaction: $F(1, 3053) = 0.57, p = .45, \eta_p^2 = .02$).

We also compared performance within the first pointing test which 'early' pointers conducted at trial 4 and 'late' pointers at trial 16 (fixed factors group and target distance). Neither perspective predicts a difference between latency in these cases. If access stays constant over learning, there should be no change in latency from first pointing to fourth pointing task. On the other hand, if we predict that both of these pointing tasks are the first instances at which integration occurs, no difference in latency would be expected either. Consistent with both positions, no such difference was observed ($F(1, 21) = 0.638, p = .43, \eta_p^2 = .01$). Similarly, both predicted a significant difference in error, which was found ($F(1, 23) = 15.4, p < .00, \eta_p^2 = .31$).

Discussion

The core issue addressed within this work was when navigators integrate the different pieces of spatial information acquired while exploring an environment. Our results suggest that in the current setup participants integrated when required to do so, i.e., when conducting a survey task (Meilinger, 2008) rather than integrating during encoding or evaluation, as assumed or suggested by some theories of spatial memory (Byrne, et al., 2007; O'Keefe, 1991; McNamara, et al., 2008; Poucet, 1993; Trullier, Wiener, Berthoz & Meyer, 1997). Several effects support this conclusion.

Pointing became quicker with repeated survey testing. To the knowledge of the authors, this effect has not been described before. This is consistent with the assumption that participants increasingly relied on information integrated during prior testing, thus shortening the average integration process. This would not be expected if participants were simply accessing already integrated spatial information. Mere improvement in task handling independent of the spatial content of the task is rather unlikely as, first, the pointing task is simple (turning ones head into the target direction and pressing a button) and second, improvement continued throughout the experiment – mere task improvement should have saturated much earlier.

Participants tested for the first time after 16 runs took longer than those with the same learning experience, who

experienced prior testing. Participants tested earlier could rely on already integrated information and thus their latency was shorter than those tested for the first time. This difference diminished when tested after 20 training runs, as the late pointing group improved more strongly than the early pointing group between trial 16 and 20. Both groups took approximately the same time when tested first, although their experience with the environment differed considerably (4 vs. 16 runs). This suggests that the amount of experience with an environment is not responsible for differences in pointing latencies, rather the amount of integration achieved before (i.e., the number of pointing trials). This again is consistent with integrating during testing rather than integration during learning.

We also found a distance effect indicating that larger distances to the pointing target (i.e., number of corridors) resulted in longer pointing times and larger errors. Such a distance effect was described before for accuracy, but to the knowledge of the authors not for latency (Thorndyke & Hayes-Roth, 1982). It is consistent with the assumption that more pieces of spatial information to be integrated take longer and also result in larger errors. Increased pointing time with larger distance is not predicted if the spatial information was accessed from an integrated representation stored in memory. However, higher error rate for larger distances is consistent with both positions about the time point of integration.

Similar to pointing time, error decreased with experience. This effect has been described before and suggests that people indeed acquire more precise knowledge about their environment with increased experience (Evans, et al., 1981; Gärling, et al., 1983; Thorndyke & Hayes-Roth, 1982). The specific new methodological contribution of the current study was first, looking at time differences for all the mentioned effects and second, comparing early and late testing thus disentangling learning and testing.

Use of a virtual environment also allowed for high control and accuracy in measurement, while providing an immersive experience within which to conduct a survey task, with realistic proprioceptive and vestibular feedback, as well as efference copies.

Altogether, the present results suggest that the integration of spatial information is an effortful process which requires time and which does not necessarily happen during learning an environment, but can also be conducted when accessing spatial information later. It is consistent with the evidence that suggests latency in such tasks works as a function of aligning reference axes of multiple reference frames into a common reference frame (Greenauer & Waller, 2010). The alternative of building up an integrated representation of an environment during learning would *not* have predicted substantial time differences in accessing this information neither for the amount of learning, nor for larger distances to the target, or as a function of prior testing.

One assumption for integrating during learning is that integrated survey information only has to be accessed. An alternative position assumes a navigator mentally walks

though an integrated environmental representation before conducting a survey task (Byrne et al., 2007). Mentally walking would explain the distance effect in time, as mentally walking longer distances to the target should also result in longer estimation times. However, mentally walking cannot explain latency reduction as a function of testing rather than experience with an environment.

Navigators' performances usually differ largely between individuals (Ishikawa & Montello, 2006). This is also true for the present study. Consequently, it is possible that some participants already integrated during encoding, however, for the vast majority this was not the case.

Our effects were obtained within a setting with clearly restricted learning time. The integration of location information within the immediate surrounding is known to result in integration costs during retrieval, only if integration was restricted to retrieval by controlling for learning time (Hanley & Levine, 1983; Yamamoto & Shelton, 2008). Pointing performance might thus look different during self-paced learning where participants have the time to integrate (see Kelly & McNamara, 2010). However, we think that this is an option rather than a necessity. Some navigators may adopt strategies of already thinking about distant locations during learning (i.e., integrate spatial information) while others may not do so. Within the present setting, most navigators integrated during pointing, otherwise our effects would not have been observed.

The present results suggest that the integration of spatial information required for survey tasks is a time consuming process. When learning an environment, integration does not necessarily occur during learning, but navigators can memorize multiple pieces of information and integrate when required to do so.

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 funded by the Ministry of Education, Science and Technology through the National Research Foundation of Korea (R31-10008). The authors thank Ivelina Alexandrova, Betty Mohler and Joachim Tesch for their help.

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.

Evans, G. W., Marrero, D. G. & Butler, P. A. (1981). Environmental learning and cognitive mapping. *Environment and Behaviour, 13*(1), 83-104.

Gärling, T., Lindberg, E. & Mantyla, T. (1983). Orientation in buildings: Effects of familiarity, visual access and orientation aids. *Journal of Applied Psychology, 68*, 177-186.

Greenauer, N. & Waller, D. (2010). Micro- and macro-reference frames: Specifying the relations between spatial categories in memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 36* 938-957.

Hanley, G. L. & Levine, M. (1983). Spatial problem solving: the integration of independently learned cognitive maps. *Memory & Cognition, 11*, 415-422.

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.

Kelly, J. W. & McNamara, T. P. (2010). Reference frames during the acquisition and development of spatial memories. *Cognition, 116*, 409-420.

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.

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.

Meilinger, T., Berthoz, A., & Wiener, J. M. (in press). The integration of spatial information across different perspectives. *Memory & Cognition*.

Meilinger, T. & Bülthoff, H. H. (2010). The Direction Bias and the Incremental Construction of Survey Knowledge. In S. Ohnsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 2500-2505). Austin: Cognitive Science Society.

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

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

Snijder, T. A. B. & Bosker, R. J. (1999). *Multilevel Analysis: An introduction to basic and advanced multilevel modeling*. Thousand Oaks, CA: Sage Publishers.

Thorndyke, P. W. & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology, 14*, 560-589.

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.

Yamamoto, N. & Shelton, A. L. (2008). Integrating object locations in the memory representation of a spatial layout. *Visual Cognition, 16*, 140-143.