

Dissociable Retrosplenial and Hippocampal Contributions to Successful Formation of Survey Representations

Thomas Wolbers and Christian Büchel

NeuroImage Nord, Department of Neurology, Universitätsklinikum Hamburg-Eppendorf, 20246 Hamburg, Germany

During everyday navigation, humans encounter complex environments predominantly from a first-person perspective. Behavioral evidence suggests that these perceptual experiences can be used not only to acquire route knowledge but also to directly assemble map-like survey representations. Most studies of human navigation focus on the retrieval of previously learned environments, and the neural foundations of integrating sequential views into a coherent representation are not yet fully understood. We therefore used our recently introduced virtual-reality paradigm, which provides accuracy and reaction-time measurements precisely indicating the emergence of survey knowledge, and functional magnetic resonance imaging while participants repeatedly encoded a complex environment from a first-person ground-level perspective. Before the experiment, we gave specific instructions to induce survey learning, which, based on the clear evidence for emerging survey knowledge in the behavioral data from 11 participants, proved successful. Neuroimaging data revealed increasing activation across sessions only in bilateral retrosplenial cortices, thus paralleling behavioral measures of map expertise. In contrast, hippocampal activation did not follow absolute performance but rather reflected the amount of knowledge acquired in a given session. In other words, hippocampal activation was most prominent during the initial learning phase and decayed after performance had approached ceiling level. We therefore conclude that, during navigational learning, retrosplenial areas mainly serve to integrate egocentric spatial information with cues about self-motion, whereas the hippocampus is needed to incorporate new information into an emerging memory representation.

Key words: survey learning; virtual reality; spatial memory; hippocampus; retrosplenial cortex; functional MRI

Introduction

The daily task of navigating a complex environment requires an ensemble of cognitive processes that rely on distinct spatial representations (Franz and Mallot, 2000). These representations can be assembled from direct experience during navigation or from indirect information such as verbal descriptions or pictorial maps. Whenever humans experience an environment from a ground-level perspective, the ensuing spatial representations may consist of both route knowledge (Gillner and Mallot, 1998; Mallot and Gillner, 2000; Wolbers et al., 2004) and survey knowledge (Thorndyke and Hayes-Roth, 1982; Ruddle et al., 1997), depending at least in part on the learning strategy applied (Kyllonen, 1993; Allen, 1999). Given that metric information about distance and direction can be incorporated early in the learning process (Montello, 1998), it seems possible to develop survey knowledge from the very beginning of a learning experience, contrary to hierarchical models of spatial learning (Siegel and White, 1975).

Although it remains unclear whether navigation in well

known environments is based on unified cognitive maps (Wang and Brockmole, 2003; Wiener and Mallot, 2003), various neuroimaging and lesion data, along with the discovery of place cells in epileptic patients, led to the belief that allocentric representations reside in the human hippocampus (Ekstrom et al., 2003; Hartley et al., 2003; Iaria et al., 2003). Moreover, the retrosplenial cortex is thought to support stimulus conversion from egocentric reference frames in parietal cortex to allocentric reference frames in medial temporal regions and vice versa (Burgess et al., 2001a,b; Ino et al., 2002).

Because most studies on human navigation have focused on the retrieval of previously learned environments, our goal was to identify the neural structures involved in acquiring survey knowledge from ground-level navigation. To form such an observer-independent representation, spatial information initially encoded in egocentric coordinates has to be transformed into an allocentric reference system. This process requires the integration of self-motion cues, presumably involving retrosplenial and medial temporal cortices. Unfortunately, recent neuroimaging studies have not provided unambiguous findings regarding the precise functions of these areas. For example, the parallel between higher error rates/longer response times and increasing hippocampal activation during maze learning (Iaria et al., 2003) contradicts the observed positive correlation between navigational accuracy and hippocampal activation during retrieval of survey knowledge (Hartley et al., 2003). Although the experimental design of the former study did not permit dissoci-

Received Nov. 17, 2004; revised Feb. 9, 2005; accepted Feb. 19, 2005.

This work was supported by the Volkswagenstiftung and the Bundesministerium für Bildung und Forschung. We thank the Physics and Methods group at NeuroImage Nord in Hamburg, Ron Paludan for providing several three-dimensional models, and Eszter Schoell for suggestions on a previous version of this report.

Correspondence should be addressed to Thomas Wolbers, NeuroImage Nord, Department of Neurology, Universitätsklinikum Hamburg-Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany. E-mail: wolbers@uke.uni-hamburg.de.

DOI:10.1523/JNEUROSCI.4705-04.2005

Copyright © 2005 Society for Neuroscience 0270-6474/05/253333-08\$15.00/0

ation of learning and retrieval, the results indicate that the hippocampus may be most important during the initial learning stages. As spatial representations become more and more accurate and performance approaches ceiling level, hippocampal contribution may decline, as has been shown for simple object-location tasks (Grön et al., 2001).

Here, we used our recently established virtual-reality paradigm (Wolbers et al., 2004), which provides behavioral parameters precisely indicating the emergence of survey knowledge. We used functional magnetic resonance imaging (fMRI) while participants repeatedly encountered a complex environment from a ground-level perspective. Based on previous findings, we expected performance-dependent activation changes in retrosplenial and medial temporal cortices to correlate with behavioral data indicative of survey learning.

Materials and Methods

Participants. Initially, 17 healthy male volunteers with normal or corrected-to-normal vision gave written informed consent to participate in this study, which was approved by the local ethics committee. All subjects understood the instructions without difficulty, and none were aware of the hypotheses at the time of testing. Because six participants did not meet our behavioral criteria for survey learning (for details, see Results), the final data set comprised 11 subjects (age range, 19–28 years).

Experimental stimuli. We used Blitz3D (Blitz Research, Auckland, New Zealand) to animate a desktop virtual environment (Fig. 1), the layout of which was derived from the “hexatown” environment developed by Gillner and Mallot (1998). Brick stone walls on both sides of the road restricted the view to the immediate straight ahead; remote road sections or intersections were invisible. Twelve distinct buildings serving as landmarks were placed at four intersections (three landmarks per intersection). These buildings were hidden behind walls unless subjects were standing directly in front of them (see below). To obtain a control condition that was carefully matched for perceptual input, we designed a single corridor with varying landmarks placed behind walls at both ends of the corridor. The corridor comprised one road section and two intersections at both ends; its appearance, which was kept constant across sessions, was identical to the environment used for encoding.

Stimuli for the retrieval task consisted of snapshots of all buildings from the same viewpoints as those encountered during navigation. These images were of the original size; in addition, smaller versions of each image were also created. To test subjects’ ability to retrieve information about spatial relationships between pairs of buildings, images were placed above one another, as shown in Figure 1 (bottom right). These pairs always consisted of buildings that were located at adjacent intersections in the environment; none of the trials depicted landmarks from the same or from remote intersections.

Procedure. Six learning and three control conditions were presented in separate sessions with a predetermined order. The experiment started with a control session, followed by three learning sessions, the second control session, an additional three learning sessions, and the final control session. Within each session, both the encoding and the retrieval tasks were presented with intermediary fixation periods (20 s). During encoding, participants were passively moved throughout the entire environment following the same route in all sessions. Subjects were instructed to infer the spatial layout of the environment and the correct landmark locations to ensure the best possible performance during retrieval; moreover, they were told that they would have to draw a map of

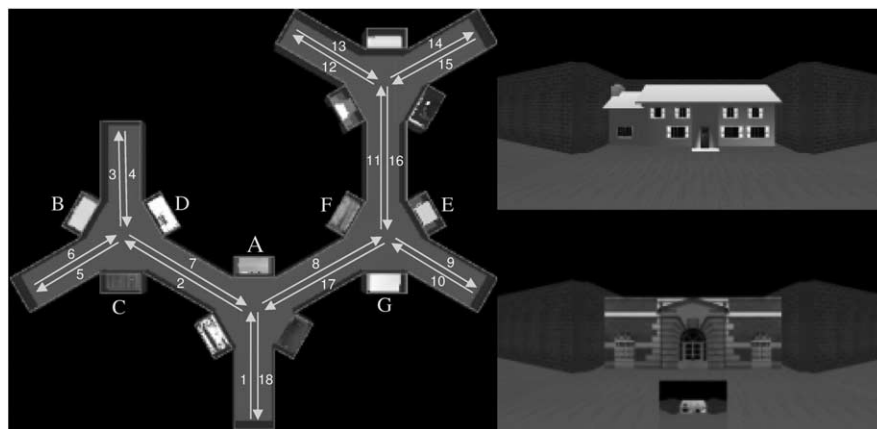


Figure 1. Example views from the virtual environment. Left, Aerial view of the environment (not shown to participants). Arrows and numbers indicate the traveled route; road sections were visited in ascending order. Letters serve to illustrate the difference between direct, close, and remote pairs. Note that, during encoding, participants were moved throughout the entire environment, thereby encountering all 12 landmarks. Top right, Ground-level view of one of the 12 landmarks. Bottom right, Example of an image used for retrieval. Subjects were to indicate by button press the relative position of the small building, imagining that they were standing in front of the large building. Given the route depicted to the left, six combinations with landmark “A” as the large building were possible. Direct pair, A–B (target building visited in immediate temporal sequence). Close pairs, A–C/A–D (target building on an adjacent intersection visited immediately after the large building). Remote pairs, A–E/A–F/A–G (target building on an adjacent intersection not visited immediately after the large building).

the virtual world after fMRI scanning. The latter seems to be a very efficient method to induce survey learning according to repeated observations during preliminary behavioral testing.

As noted above, landmarks were hidden behind walls most of the time. However, when the midpoint of an intersection was reached, after a randomized delay between 500 and 800 ms, the wall in front disappeared, thereby unveiling the building behind. After 2500 ms, the subject (i.e., camera) performed a 60° turn (at 90.23°/s), and the wall was reinstated. After reaching a dead end, a 180° turn (at 90.23°/s) was executed, and the journey continued to the previously visited intersection. During control sessions, subjects were moved repeatedly back and forth along a single corridor. Close to each dead end, the wall would disappear after a randomized delay lasting between 500 and 800 ms to reveal one of the 12 buildings encountered during encoding. After 2500 ms, the camera performed a 180° turn (at 90.23°/s), the wall was reinstated, and the virtual movement was continued toward the opposite wall. Each time a wall vanished, a different landmark was seen, allowing for the presentation of all 12 of the buildings within six visits of each wall. The presentation order was randomized across subjects. Participants were told to pay attention to all buildings; no task was to be performed. Because we ensured that visual stimulation was carefully matched between control and encoding conditions, activation differences could reliably be attributed to the learning of the spatial layout.

During retrieval, pairs of original and small-sized landmark images were presented until a response was given. Participants were instructed to imagine that they were facing the large building and to assess the relative spatial position of the small building (referred to as the target building). This was done by pressing one of three response buttons (left button, target building located to the left; right button, target building located to the right; middle button, target building located behind). Note that the exact position of the target building within the adjacent intersection was irrelevant; the task only required the subject to determine the relative spatial position of the intersection containing the target building. Subjects were encouraged to respond as quickly as possible.

Eighteen pairs of buildings and six null events were presented in randomized order; intertrial intervals randomly varied between 3000 and 5000 ms. Six pairs depicted buildings encountered in immediate succession during encoding; these pairs are referred to as direct pairs (Fig. 1, pair A–B). Six other pairs contained landmarks that were not encountered in immediate order; however, the target building was located on an adjacent intersection that was visited during encoding immediately after

the large building. These pairs are referred to as close pairs (Fig. 1, pairs A–C and A–D). The remaining six pairs differed from the close pairs in that the target building was taken from an adjacent intersection that was not visited immediately after the large building. These pairs are referred to as remote pairs (Fig. 1, pairs A–E, A–F, and A–G). To maximize the variance of the number of intermediate path segments, we included two close and two remote pairs in which the target building had been encountered before the large building. However, we ensured that none of these pairs presented the buildings from a direct pair in reverse order. In summary, although spatial distance was held constant across all pairs (the target building was always located on an adjacent intersection), the important difference lay in the varying temporal delays between encountering both landmarks during encoding. Whereas the buildings constituting direct pairs were separated by only one path segment during navigation, close and remote pairs contained buildings that were separated by varying numbers of up to 11 path segments.

The rationale behind this approach was to obtain a behavioral measure allowing for a clear identification of the nature of the acquired representation. The acquisition of route knowledge in the form of temporospatial associations between consecutive landmark views manifests itself in a significant behavioral advantage of direct pairs over both close and remote pairs (Wolbers et al., 2004), because only direct pairs contain landmarks encountered in immediate temporal succession during encoding. However, if subjects are able to infer a survey representation of the environment, no significant behavioral differences (response time and accuracy) among direct, close, and remote pairs would be expected, because spatial distances are identical.

The retrieval task during control sessions did not require the retrieval of spatial representations. In this condition, all 18 pairs used for retrieval, plus six additional pairs depicting the same building in regular and small sizes, were presented with a duration of 4000 ms along with six null events. Subjects were instructed to indicate by a button press whether both buildings were identical (right button) or different (left button). Intertrial intervals were identical to those in the learning sessions.

After fMRI scanning concluded, subjects were asked to draw a map of the environment as accurately as possible. To minimize possible interindividual variance related to the presence or absence of silent naming of the landmarks, participants had to memorize the names of all buildings before fMRI scanning. After presenting all 12 of the buildings accompanied by their names, subjects were asked to retrieve the name of each building; in cases of failure, the experimenter provided the correct name. This procedure continued until each participant could reliably name each landmark.

MRI acquisition. MR scanning was performed on a 3 T MRI scanner (Trio; Siemens AG, Munich, Germany) with a standard head coil. Thirty-seven contiguous axial slices (slice thickness, 3 mm) were acquired using a gradient-echo echo-planar T2*-sensitive sequence (repetition time, 2.12 s; echo time, 25 ms; flip angle, 70°; matrix, 64 × 64; field of view, 192 × 192 mm).

A liquid crystal display video projector back-projected the stimuli on a screen positioned on top of the head coil. Subjects lay on their backs within the bore of the magnet and viewed the stimuli comfortably via a 45° mirror that reflected the images displayed on the screen. To minimize head movement, all subjects were stabilized with tightly packed foam padding surrounding the head. Encoding stimuli were presented using Blitz3D; retrieval stimuli were presented using Presentation (Neurobehavioral Systems, Albany, CA).

Image processing and statistical analysis. Image processing and statistical analysis were performed using SPM2 (www.fil.ion.ucl.ac.uk/spm/software/spm2). All volumes were realigned to the first volume, spatially normalized (Friston et al., 1995) to an echo-planar imaging template in a standard coordinate system (Evans et al., 1993), and finally, smoothed using a 9 mm full-width at half-maximum isotropic Gaussian kernel. To test the hypothesis of deducing survey knowledge from ground-level navigation, we modeled the onset of each encoding block as a δ function convolved with a hemodynamic response function (HRF). Specific effects were tested with appropriate linear contrasts of the parameter estimates for the HRF regressor. Data were analyzed for each subject individually (first-level analysis) and for the group. At the single-subject level,

we applied a high-pass filter to remove baseline drifts. Design matrices containing nine separate sessions (six encoding and three encoding control sessions) were specified, thereby removing session-specific effects. In addition, we included realignment parameters to regress out movement-related activation. Separate contrast images for each of the nine regressors were subsequently entered into a random-effects analysis (Friston et al., 1999) to obtain results that could be generalized beyond the subjects taking part in this study.

A main effect of learning and the performance-related increase were assessed with a repeated-measures ANOVA; the problem of nonindependent data within subjects, as well as error variance heterogeneity, was addressed by performing a nonsphericity correction. To model a performance-related activation increase across encoding sessions, we first estimated the underlying learning curve for remote pairs by subjecting the mean performance data to the recently introduced state-space smoothing algorithm (Smith et al., 2004). This procedure was deliberately chosen because absolute performance values underlie random fluctuations that may not accurately represent the current state of learning. We specifically chose remote pairs because it is possible that, in addition to learning the layout of the environment, subjects acquire some knowledge about the route as well, which could influence performance, especially for direct pairs. In contrast, remote pairs contain buildings with very long temporal delays and are thus most difficult to solve by route knowledge. As a consequence, these pairs provide the purest measure of survey expertise that is available in our paradigm.

The normalized learning estimates, as obtained from the state-space smoothing algorithm, were then used as contrast weights. We also tested for a performance-related decrease in activation across encoding sessions, using the inverted normalized learning curve. Given the substantial variability across subjects regarding learning speed, we checked for activation predicting the learning-related change (i.e., the increment in performance from one session to the next) by applying the individual estimated learning curves for remote pairs at the single-subject level. Specifically, we used the normalized differences between the estimated learning parameters of two consecutive sessions as contrast weights (partial derivative of performance with respect to time) and entered the resulting contrast images (one image per subject) into a one-sample *t* test.

For all analyses, the threshold was set to $p < 0.05$, corrected for multiple comparisons. According to results from previous studies, several regions of interest were defined (medial frontal gyrus, retrosplenial cortex, inferior and superior parietal cortex, parahippocampal gyrus, hippocampus, and caudate nucleus), and correction for multiple comparisons was based on these regions. We applied spherical search volumes ranging from 2056 cm³ (hippocampus) to 4120 cm³ (retrosplenial cortex). Elsewhere in the brain, correction was based on the entire search volume.

Results

Behavioral data

According to our hypotheses, acquisition of survey knowledge should manifest itself in performance improvements for all three categories of retrieval pairs, thus being independent of the varying temporal delays. We determined whether learning had occurred for each category by checking for learning trials with the state-space smoothing algorithm (Smith et al., 2004). Learning trials are defined as the first trial in which there is reasonable certainty (>0.95) that a subject performs better than chance for the balance of the experiment. In 11 of 17 participants, we identified learning trials for direct, close, and remote pairs, thus meeting our criterion for survey learning. The remaining six participants did not show consistent learning effects in any of the retrieval categories (Fig. 2, bottom). It remains unclear whether these subjects continuously attempted to learn the environment, despite being unsuccessful. We therefore decided to exclude their data from additional analysis to ensure that the neuroimaging results could be reliably attributed to learning the spatial layout of the environment.

The behavioral data of the 11 survey learners were characterized by increasing accuracy and decreasing response times in all

three categories (Fig. 2, top). A 3×6 repeated-measures ANOVA (factors, type of pair, and retrieval session) with performance measures as the dependent variable revealed significant main effects for retrieval session ($F = 34.25$; $p < 0.001$) and type of pair ($F = 4.23$; $p = 0.035$) but no evidence for a significant interaction ($F = 1.47$; $p = 0.222$). *Post hoc* Tukey's honestly significant difference tests revealed that the main effect for type of pairs was driven by significant differences between close and remote pairs ($p = 0.039$) and by a trend toward significance for the comparison direct versus remote pairs ($p = 0.07$), whereas direct and close pairs did not differ significantly ($p = 0.953$). For the response time data, we also observed significant main effects for retrieval session ($F = 4.03$; $p = 0.045$) and type of pairs ($F = 12.92$; $p < 0.001$), whereas the interaction term did not reach statistical significance ($F = 0.50$; $p < 0.668$). The main effect for type of pair was related to significant differences between direct and remote pairs ($p = 0.001$) and close and remote pairs ($p = 0.001$), whereas direct and close pairs did not differ significantly ($p = 0.959$).

With regard to the significant main effects for type of pairs, we were concerned that participants might have mentally traveled along the encoded route to solve the retrieval pairs, thus leading to longer reaction times and inferior performance for remote pairs in particular. We therefore correlated reaction times and performance with the number of traveled path segments between encountering buildings during navigation for each session individually. However, neither analysis revealed significant correlations, strongly indicating the emergence of survey instead of complex route knowledge.

Our interpretation that the behavioral results reflect the acquisition of map-like representations was supported by the maps drawn after fMRI scanning. All 11 of the subjects were able to correctly reproduce the layout of the environment, as well as the correct landmark locations. However, we decided not to perform a quantitative analysis of map-drawing performance, because it seems difficult to identify the nature of a spatial representation with bidimensional regression coefficients or related measures (Waterman and Gordon, 1984; Spiers et al., 2001; Friedman and Kohler, 2003). Montello et al. (2004) have argued that an accurate map can be drawn from a quantitatively scaled route representation. This assumption is supported by recent findings showing that, although participants used sequential versus hierarchical strategies when drawing maps after route versus survey encoding, the corresponding distortion indices did not differ significantly (Shelton and Gabrieli, 2002).

Imaging data

Main effect of learning

We first contrasted encoding with the control to identify overall activation attributable to learning the spatial layout of the environment. This analysis yielded several significantly activated

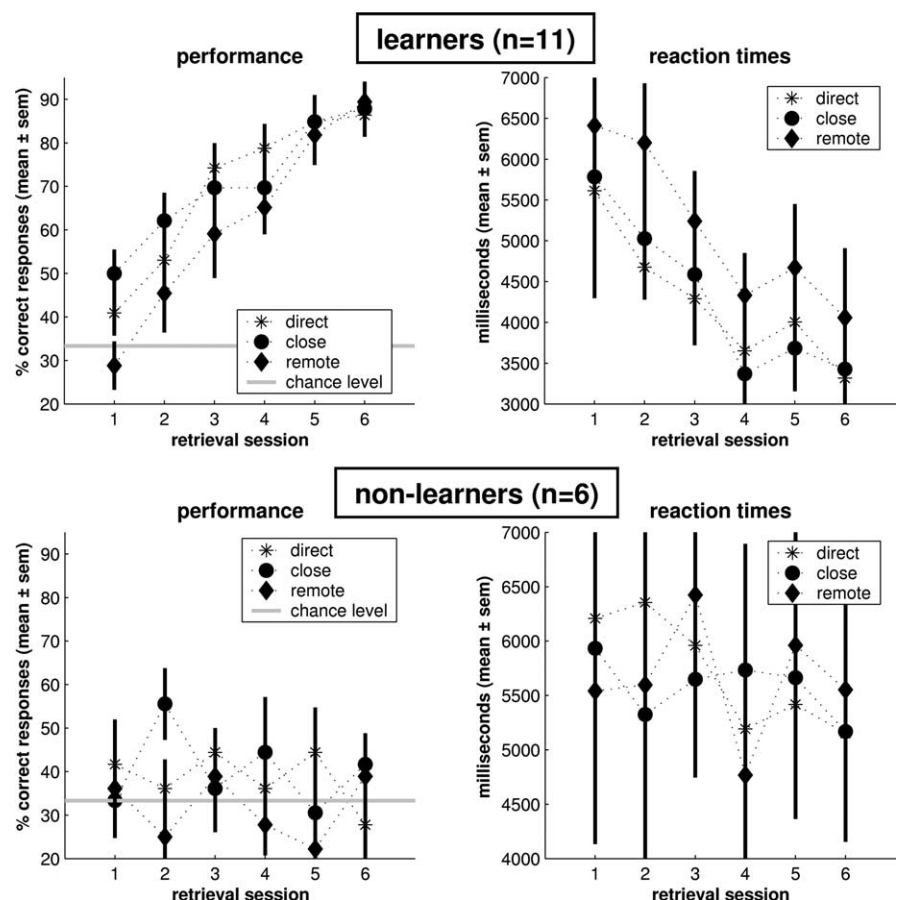


Figure 2. Behavioral results during retrieval. Top, Performance and reaction time data for the 11 participants who met our criteria for survey learning. Significant changes across sessions were observed for direct, close, and remote pairs, indicating the emergence of survey knowledge. Bottom, Performance and reaction time data for the six participants who did not meet our criteria for survey learning. Because of the absence of learning for all three of the categories, these subjects were excluded from additional analyses.

clusters, including the right medial frontal gyrus, the inferior parietal lobe, and the cuneus. Table 1 lists the significantly activated areas according to Montreal Neurological Institute (MNI) space (Evans et al., 1993), along with coordinates and statistical results for the respective peak voxels.

Performance-related increase across encoding sessions

We modeled learning across sessions with the fitted learning curve to assess whether performance improvements were paralleled by systematic changes in cortical activation during encoding. Areas exhibiting unspecific time effects were excluded by masking this contrast exclusively with areas showing increasing activation across control sessions. Thus, we were able to identify regions showing performance-related activation only across encoding sessions. The results are shown in Figure 3, along with mean regression coefficients for the local activation maximum; Table 1 displays the z -value of the peak voxel. Increasing blood oxygenation level-dependent (BOLD) responses were observed bilaterally in a region ranging from the posterior part of the retrosplenial cortex to the anterior bank of the parieto-occipital sulcus; no additional activations emerged elsewhere in the brain.

We also looked for performance-related activation decreases across encoding sessions using the inverted normalized learning curve. Again, to exclude areas exhibiting unspecific time effects, this contrast was masked exclusively by areas showing decreasing

Table 1. Spatial coordinates of the local maxima in the group analysis ($p < 0.05$, corrected)

| Region | Coordinate (x, y, z; in mm) | | Voxel level (z-score) |
|---|-----------------------------|-------------|-----------------------|
| | LH | RH | |
| Local activation maxima for the main effect of learning | | | |
| Medial frontal gyrus | | 30, 54, 0 | 3.95 |
| | | 22, 52, 2 | 3.84 |
| Premotor cortex | | 28, -2, 44 | 5.54 |
| | | 42, 8, 32 | 5.18 |
| | -46, 6, 36 | | 5.49 |
| Inferior parietal cortex | | 34, -50, 46 | 5.50 |
| | | 54, -38, 52 | 4.80 |
| | -32, -56, 56 | | 4.62 |
| | -30, -48, 44 | | 4.47 |
| Superior parietal cortex | | 18, -68, 60 | 5.81 |
| | -14, -54, 58 | | 5.03 |
| Cuneus | | 28, -76, 46 | 4.68 |
| | -16, -74, 36 | | 4.91 |
| Local activation maximum for performance-related increase | | | |
| Retrosplenial cortex | | 4, -62, 24 | 3.59 |
| Local activation maxima for learning-related change | | | |
| Hippocampus | -22, -12, -14 | | 3.20 |
| | -20, -12, -18 | | 3.19 |
| | | 24, -4, -18 | 2.39 (n.s.) |

RH, Right hemisphere; LH, left hemisphere; n.s., nonsignificant.

activation across encoding control sessions. However, we did not obtain areas with significant activation in this analysis.

Learning-related change across encoding sessions

Subsequently, we checked for areas with activation reflecting the amount of knowledge acquired in a given session. This analysis would identify regions that are responsible for integrating new information into an emerging memory representation. We observed significant results in the left hippocampus, whereas the right hippocampus only showed a trend toward statistical significance (Fig. 4, Table 1). None of the remaining regions of interest and no areas elsewhere in the brain contained suprathreshold voxels. To further illustrate the relationship between hippocampal activation and behavioral performance, Figure 4 also contains fitted learning curves, their partial derivatives, and the regression coefficients from left hippocampal peak voxels for two single subjects. Subject 11 showed very rapid learning, with performance approaching ceiling level by session 3. Hippocampal activation was the most prominent in the first two sessions and then decayed substantially. In contrast, subject 03 mainly showed performance improvements in the second half of the experiment, thus approaching ceiling level by session 5. Consequently, hippocampal activation was most prominent in sessions 4 and 5, when the strongest learning effects were observed.

Altogether, whereas significant activation increases across encoding sessions mimicking the estimated learning curves were confined to the retrosplenial cortex, activation in the hippocampus reflected the amount of knowledge acquired in a given session.

Discussion

The present study attempted to identify the neural foundations of successfully forming a survey representation derived from ground-level navigation. We used our recently introduced

virtual-reality paradigm, because it provides an objective distinction between route and survey knowledge (Wolbers et al., 2004). Eleven of 17 participants showed clear evidence for survey learning, reflected by increasing accuracy and decreasing response times independent of the varying temporal delays. Neuroimaging results revealed a dissociation between the contributions of the retrosplenial cortex and the hippocampus. Whereas retrosplenial activation paralleled behavioral measures of map expertise, BOLD responses in the left hippocampus indicated the extent of learning in a given session. We therefore conclude that, in the context of navigational learning, the hippocampus mainly serves to incorporate additional information into an emerging memory representation.

Significant learning effects for direct, close, and remote pairs were observed in both behavioral measures. This clearly demonstrates that spatial learning was not confined to the formation of temporospatial associations between consecutive landmark views, as shown during route learning (Wolbers et al., 2004). Instead, participants were able to infer a configurational representation from ground-level navigation, because performance improvements occurred in all retrieval categories, despite the substantially varying temporal delays. The excellent map-drawing performance lends additional support to this interpretation. However, we cannot exclude the possibility that participants may have acquired some route knowledge as well, which could explain the overall differences revealed by the *post hoc* tests. Most interestingly, giving specific instructions that point to an allocentric learning strategy seems necessary to induce survey learning, because most participants in our previous study, when only unspecific instructions were given, showed evidence for route learning only (Wolbers et al., 2004).

We can only speculate why six participants failed to meet our criteria for survey learning. Mental abilities, such as spatial visualization and spatial orientation, contribute to interindividual differences in acquiring spatial information from a virtual environment (Waller, 2000). These abilities might be less pronounced in the excluded participants, which makes it impossible to infer a substantial amount of survey knowledge within six exposures to the complex environment. Reliable insights into the neural correlates of survey learning can therefore only be obtained from those participants who showed clear evidence for inferring an allocentric representation.

What processes are needed to transform a virtual environment experienced from a first-person perspective into a survey representation? During navigation, spatial positions of landmarks are perceived in egocentric coordinates. Because the coordinates change with observer movements, optic flow needs to be taken into account to assess current heading direction and the extent of translational movements. This information allows for dynamic spatial updating and for computing displacement vectors (Muehl and Sholl, 2004). Spatial relationships between landmarks can then be estimated, based on an intrinsic reference system that presumably depends on the initial orientation of the observer (Shelton and McNamara, 2004).

Contrasting encoding with the control condition revealed a network of areas that have been repeatedly observed during spatial learning in virtual environments (Aguirre et al., 1996; Shelton and Gabrieli, 2002; Iaria et al., 2003). Whereas the recruitment of the medial frontal gyrus might reflect working memory processes necessary for long-term encoding (Leung et al., 2002; Suzuki et al., 2002), inferior and superior parietal regions process spatial information in egocentric reference frames (Andersen et al., 1997; Galati et al., 2001; Halligan et al., 2003). The lack of signif-

icant effects in the parahippocampal cortex may be attributable to its primary role in viewpoint-dependent processing of spatial scenes (Aguirre et al., 1996; Epstein et al., 2003). Given that identical landmarks were presented in a very similar virtual world in the control compared with the encoding condition, parahippocampal involvement would be expected in both conditions. Consequently, these activations should cancel each other out in direct comparisons.

We observed performance-dependent activation increases only in the anterior bank of the parieto-occipital sulcus and in the posterior part of the retrosplenial cortex. In the monkey brain, retrosplenial areas are densely connected with various medial temporal regions, the posterior parietal cortex, and the mid-dorsolateral prefrontal cortex (Maguire, 2001; Kobayashi and Amaral, 2003). Moreover, the retrosplenial cortex of the rat contains head-direction cells that indicate the current heading of the animal (Cho and Sharp, 2001). It is, consequently, in a perfect position to integrate egocentric spatial information from the posterior parietal cortex with information about self-motion and landmark identity originating, at least in part, from medial temporal areas. When this process is disrupted by retrosplenial lesions, severe problems with forming or recalling links between landmark identity and directional information arise (Aguirre and D'Esposito, 1999). Therefore, we believe that the learning-related activation increase observed in bilateral retrosplenial cortices reflects this increasingly robust integration of different sources of spatial information. Constituting a necessary prerequisite for forming survey representations, a more efficient integration allows for faster and more accurate access to the spatial relationships between landmarks during retrieval.

Unlike the performance-dependent retrosplenial activation, BOLD responses in the hippocampus predicted the extent of learning in a given session. In the context of learning a spatial layout, geometric determinants of simple environments (i.e., proximity of local walls) can shape hippocampal place-cell firing in the rat over repeated exposure (Lever et al., 2002). A coherent representation of a complex environment, such as the one used in the present study, however, requires the integration of multiple local environments (i.e., intersections) that are not surrounded by overall boundaries. Such a representation also contains spatial relationships between superordinate entities, such as intersections or regions, which have been shown to substantially influence navigational behavior (Wiener and Mallot, 2003). We therefore suggest that the hippocampus serves to integrate both the layout of local parts and the spatial relationships between superordinate entities into a coherent representation. In early learning stages, the hippocampal contribution is very strong, because the existing representation is only rudimentary, but, as learning progresses, the amount of new associations to be assembled gradually decreases. As a consequence, hippocampal activation becomes weaker in later stages of learning, as has been shown for simple geometric patterns (Grön et al., 2001). This interpretation is supported by the positive correlation between the BOLD re-

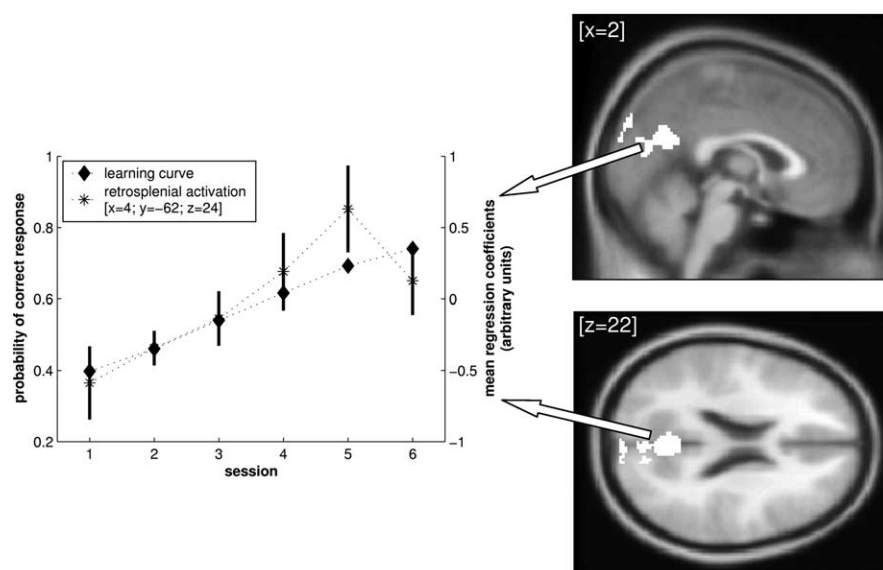


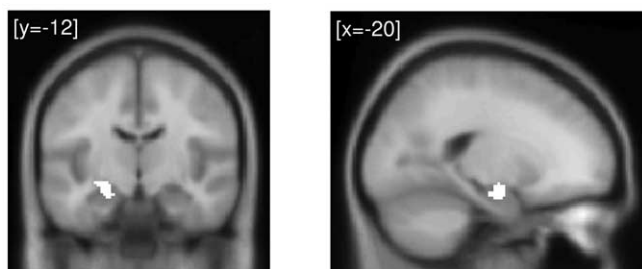
Figure 3. Performance-related increase ($n = 11$). Areas showing a significant activation increase across encoding sessions that paralleled behavioral performance are indicated. Left, Mean \pm SEM regression coefficients of the peak voxel in the retrosplenial cortex, along with the fitted learning curve for remote pairs. Results of the random-effects analysis are displayed with a threshold of $p < 0.05$ (corrected) on the averaged MNI template brain. Regions with increasing activation across encoding control sessions were excluded (for an uncorrected activation map omitting this masking procedure, see supplemental material, available at www.jneurosci.org). Note that, because of previous anatomical hypotheses, correction for multiple comparisons in the retrosplenial cortex was based on a reduced search volume.

sponse in the hippocampus and the error rate as well as the reaction time that was observed for participants using a spatial strategy during maze learning (Iaria et al., 2003). When these participants made many errors and took a long time to perform the task, hippocampal activation was high, interpreted as indicating the importance of the hippocampus for the learning process.

Activation in the posterior parietal cortex has been shown to parallel behavioral performance during route learning, reflecting the successful construction of spatial links between consecutive landmarks (Wolbers et al., 2004). Although subjects may have acquired some knowledge about the route in the present study as well, the parietal response presumably was attenuated as participants actively focused on inferring survey knowledge. The crucial task of transforming spatial relationships, initially perceived in egocentric coordinates, into an observer-independent reference frame rather depends on retrosplenial and hippocampal functioning (Maguire, 2001; Burgess et al., 2002). Although the importance of the hippocampus for retrieving allocentric representations has been questioned (Teng and Squire, 1999; McNamara and Shelton, 2003; Squire et al., 2004), converging evidence highlights its role in the learning process. This densely connected area may best be described as a modality-independent, additional working memory buffer that serves to associate complex unimodal or multimodal information (Hölscher, 2003). It seems especially important whenever associations need to be formed in a way that allows their flexible use to guide future behavior, as opposed to rigid stimulus–response associations (O’Keefe and Nadel, 1978; Packard et al., 1989; Packard and White, 1990; McNamara and Shelton, 2003). In the current study, high complexity was imposed by the need not only to associate local and superordinate spatial entities but also to transform egocentric coordinates into an observer-independent representation. Forming such a flexible representation may just require the additional recruitment of the hippocampus.

With regard to a possible hemispheric asymmetry, the right

group analysis



individual subjects

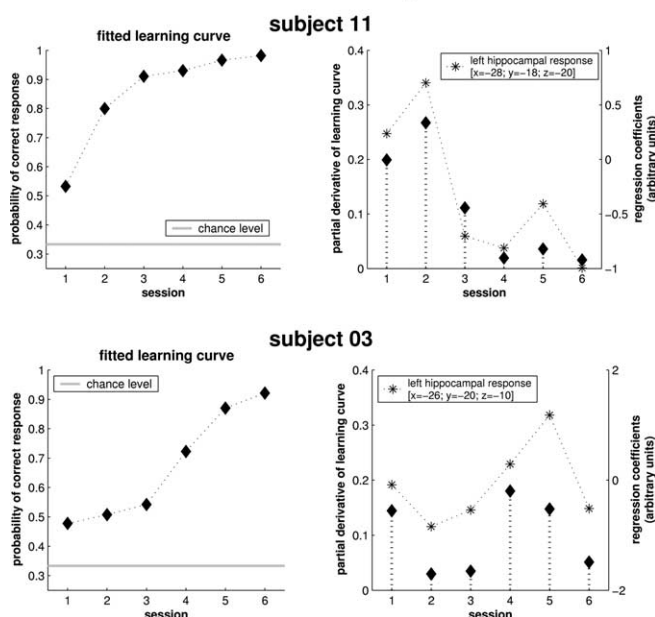


Figure 4. Learning-related change ($n = 11$). Top, Left hippocampal region with activation reflecting the amount of knowledge acquired in any given encoding session. Results of the random-effects analysis are displayed with a threshold of $p < 0.05$ (corrected) on the averaged MNI template brain (for an uncorrected activation map, see supplemental material, available at www.jneurosci.org). Note that, because of previous anatomical hypotheses, correction for multiple comparisons in the hippocampus was based on a reduced search volume. Bottom, Fitted learning curves for remote pairs in two representative subjects (left). Right, The corresponding partial derivatives, indicating the amount of learning in a given session, and the regression coefficients of left hippocampal peak voxels. Whereas hippocampal activation in subject 11 is strongest in the initial learning stage and decays rapidly after performance has approached ceiling level in session 3, the slower learning process in subject 03 is paralleled by stronger hippocampal activation in the second half of the experiment. As a consequence, hippocampal activation seems to be the most prominent whenever substantial performance improvements are observable.

hippocampus is often assumed to be especially important for spatial memory, whereas the left hippocampus may be more involved in context-dependent episodic memory (Maguire et al., 1998; Burgess et al., 2002). Although we did not observe significant learning-related changes in the right hippocampus, this does not rule out its potential involvement, because the null hypothesis cannot be verified. However, it is worth noting that there are conflicting findings regarding a hemispheric specialization in the human hippocampus. Several neuroimaging and lesion studies have demonstrated a bilateral or even left-sided hippocampal involvement in spatial memory (Grön et al., 2000; Iaria et al., 2003; Incisa della Rocchetta et al., 2004), which is not in agreement with a strict dissociation of hippocampal functions.

The results of the current study point to different roles of the retrosplenial cortex and the hippocampus during the construction of a survey representation out of ground-level navigation. Having established such a functional dissociation, subsequent studies may now attempt to assess the dynamic interactions between these and other regions that are necessary for the transfer and integration of the described sources of (non)spatial information.

References

- Aguirre GK, D'Esposito M (1999) Topographical disorientation: a synthesis and taxonomy. *Brain* 122:1613–1628.
- Aguirre GK, Detre JA, Alsop DC, D'Esposito M (1996) The parahippocampus subserves topographical learning in man. *Cereb Cortex* 6:823–829.
- Allen GL (1999) Spatial abilities, cognitive maps and wayfinding: bases for individual differences in spatial cognition and behavior. In: *Wayfinding behavior* (Golledge R, ed), pp 46–80. Baltimore: Johns Hopkins.
- Andersen RA, Snyder LH, Bradley DC, Xing J (1997) Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annu Rev Neurosci* 20:303–330.
- Burgess N, Maguire EA, Spiers HJ, O'Keefe J (2001a) A temporoparietal and prefrontal network for retrieving the spatial context of lifelike events. *NeuroImage* 14:439–453.
- Burgess N, Becker S, King JA, O'Keefe J (2001b) Memory for events and their spatial context: models and experiments. *Philos Trans R Soc Lond B Biol Sci* 356:1493–1503.
- Burgess N, Maguire EA, O'Keefe J (2002) The human hippocampus and spatial and episodic memory. *Neuron* 35:625–641.
- Cho J, Sharp PE (2001) Head direction, place, and movement correlates for cells in the rat retrosplenial cortex. *Behav Neurosci* 115:3–25.
- Ekstrom AD, Kahana MJ, Caplan JB, Fields TA, Isham EA, Newman EL, Fried I (2003) Cellular networks underlying human spatial navigation. *Nature* 425:184–188.
- Epstein R, Graham KS, Downing PE (2003) Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron* 37:865–876.
- Evans AC, Collins DL, Mills DR, Brown ED, Kelly RL, Peters TM (1993) 3D statistical neuroanatomical models from 305 MRI volumes. *Proc IEEE Nucl Sci Symp Med Imaging* 1–3:1813–1817.
- Franz MO, Mallot HA (2000) Biomimetic robot navigation. *Rob Auton Syst* 30:133–153.
- Friedman A, Kohler B (2003) Bidimensional regression: assessing the configurational similarity and accuracy of cognitive maps and other two-dimensional data sets. *Psychol Methods* 8:468–491.
- Friston KJ, Ashburner J, Frith CD, Poline JB, Heather JD, Frackowiak RSJ (1995) Spatial registration and normalization of images. *Hum Brain Mapp* 2:1–25.
- Friston KJ, Holmes AP, Worsley KJ (1999) How many subjects constitute a study? *NeuroImage* 10:1–5.
- Galati G, Committeri G, Sanes JN, Pizzamiglio L (2001) Spatial coding of visual and somatic sensory information in body-centred coordinates. *Eur J Neurosci* 14:737–746.
- Gillner S, Mallot HA (1998) Navigation and acquisition of spatial knowledge in a virtual maze. *J Cogn Neurosci* 10:445–463.
- Grön G, Wunderlich AP, Spitzer M, Tomczak R, Riepe MW (2000) Brain activation during human navigation: gender-different neural networks as substrate of performance. *Nat Neurosci* 3:404–408.
- Grön G, Bittner D, Schmitz B, Wunderlich AP, Tomczak R, Riepe MW (2001) Hippocampal activations during repetitive learning and recall of geometric patterns. *Learn Mem* 8:336–345.
- Halligan PW, Fink GR, Marshall JC, Vallar G (2003) Spatial cognition: evidence from visual neglect. *Trends Cogn Sci* 7:125–133.
- Hartley T, Maguire EA, Spiers HJ, Burgess N (2003) The well-worn route and the path less traveled. Distinct neural bases of route following and wayfinding in humans. *Neuron* 37:877–888.
- Hölscher C (2003) Time, space and hippocampal functions. *Rev Neurosci* 14:253–284.
- Iaria G, Petrides M, Dagher A, Pike B, Bohbot VD (2003) Cognitive strategies dependent on the hippocampus and caudate nucleus in human navigation: variability and change with practice. *J Neurosci* 23:5945–5952.
- Incisa della Rocchetta A, Samson S, Ehrle N, Denos M, Hasboun D, Baulac M (2004) Memory for visuospatial location following selective hippocam-

- pal sclerosis: the use of different coordinate systems. *Neuropsychology* 18:15–28.
- Ino T, Inoue Y, Kage M, Hirose S, Kimura T, Fukuyama H (2002) Mental navigation in humans is processed in the anterior bank of the parieto-occipital sulcus. *Neurosci Lett* 322:182–186.
- Kobayashi Y, Amaral DG (2003) Macaque monkey retrosplenial cortex. II. Cortical afferents. *J Comp Neurol* 466:48–79.
- Kyllonen PC (1993) Aptitude testing inspired by information processing: a test of the four-source model. *J Gen Psychol* 120:375–405.
- Leung HC, Gore JC, Goldman-Rakic PS (2002) Sustained mnemonic response in the human middle frontal gyrus during on-line storage of spatial memoranda. *J Cogn Neurosci* 14:659–671.
- Lever C, Wills T, Cacucci F, Burgess N, O'Keefe J (2002) Long-term plasticity in hippocampal place-cell representation of environmental geometry. *Nature* 416:90–94.
- Maguire EA (2001) The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings. *Scand J Psychol* 42:225–238.
- Maguire EA, Burgess N, Donnett JG, Frackowiak RS, Frith CD, O'Keefe J (1998) Knowing where and getting there: a human navigation network. *Science* 280:921–924.
- Mallot HA, Gillner S (2000) Route navigating without place recognition: what is recognised in recognition-triggered responses? *Perception* 29:43–55.
- McNamara TP, Shelton AL (2003) Cognitive maps and the hippocampus. *Trends Cogn Sci* 7:333–335.
- Montello DR (1998) A new framework for understanding the acquisition of spatial knowledge in large-scale environments. In: *Spatial and temporal reasoning in geographic information systems* (Egenhofer MJ, Golledge RG, eds), pp 143–154. New York: Oxford UP.
- Montello DR, Hegarty M, Richardson AE, Waller D (2004) Spatial memory of real environments, virtual environments, and maps. In: *Human spatial memory: remembering where* (Allen GA, ed), pp 251–285. Mahwah, NJ: Erlbaum.
- Muehl KA, Sholl MJ (2004) The acquisition of vector knowledge and its relation to self-rated direction sense. *J Exp Psychol Learn Mem Cogn* 30:129–141.
- O'Keefe J, Nadel L (1978) *The hippocampus as a cognitive map*. Oxford: Oxford UP.
- Packard MG, White NM (1990) Lesions of the caudate nucleus selectively impair "reference memory" acquisition in the radial maze. *Behav Neural Biol* 53:39–50.
- Packard MG, Hirsh R, White NM (1989) Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: evidence for multiple memory systems. *J Neurosci* 9:1465–1472.
- Ruddle RA, Payne SJ, Jones DM (1997) Navigating buildings in "desk-top" virtual environments: Experimental investigations using extended navigational experience. *J Exp Psychol Appl* 3:143–159.
- Shelton AL, Gabrieli JD (2002) Neural correlates of encoding space from route and survey perspectives. *J Neurosci* 22:2711–2717.
- Shelton AL, McNamara TP (2004) Orientation and perspective dependence in route and survey learning. *J Exp Psychol Learn Mem Cogn* 30:158–170.
- Siegel AW, White SH (1975) The development of spatial representations of large-scale environments. In: *Advances in child development and behaviour* (Reese HW, ed), pp 9–55. New York: Academic.
- Smith AC, Frank LM, Wirth S, Yanike M, Hu D, Kubota Y, Graybiel AM, Suzuki WA, Brown EN (2004) Dynamic analysis of learning in behavioral experiments. *J Neurosci* 24:447–461.
- Spiers HJ, Burgess N, Maguire EA, Baxendale SA, Hartley T, Thompson PJ, O'Keefe J (2001) Unilateral temporal lobectomy patients show lateralized topographical and episodic memory deficits in a virtual town. *Brain* 124:2476–2489.
- Squire LR, Stark CE, Clark RE (2004) The medial temporal lobe. *Annu Rev Neurosci* 27:279–306.
- Suzuki M, Fujii T, Tsukiura T, Okuda J, Umetsu A, Nagasaka T, Mugikura S, Yanagawa I, Takahashi S, Yamadori A (2002) Neural basis of temporal context memory: a functional MRI study. *NeuroImage* 17:1790–1796.
- Teng E, Squire LR (1999) Memory for places learned long ago is intact after hippocampal damage. *Nature* 400:675–677.
- Thorndyke PW, Hayes-Roth B (1982) Differences in spatial knowledge acquired from maps and navigation. *Cognit Psychol* 14:560–589.
- Waller D (2000) Individual differences in spatial learning from computer-simulated environments. *J Exp Psychol Appl* 6:307–321.
- Wang RF, Brockmole JR (2003) Human navigation in nested environments. *J Exp Psychol Learn Mem Cogn* 29:398–404.
- Waterman S, Gordon D (1984) A quantitative-comparative approach to analysis of distortion in mental maps. *Prof Geogr* 36:326–337.
- Wiener JM, Mallot HA (2003) 'Fine-to-coarse' route planning and navigation in regionalized environments. *Spat Cogn Comput* 3:331–358.
- Wolbers T, Weiller C, Büchel C (2004) Neural foundations of emerging route knowledge in complex spatial environments. *Brain Res Cogn Brain Res* 21:401–411.