

Research report

Neural foundations of emerging route knowledge in complex spatial environments

Thomas Wolbers*, Cornelius Weiller, Christian Büchel

NeuroImage Nord, Department of Neurology, Universitätsklinikum Hamburg-Eppendorf, Martinistr. 52, 20246 Hamburg, Germany

Accepted 11 June 2004

Available online 8 August 2004

Abstract

Behavioral evidence suggests that spatial knowledge derived from ground-level navigation can consist of both route and survey knowledge. Neuroimaging and lesion studies aiming to identify the neural structures responsible for topographical learning in humans have yielded partially inconsistent results, probably due to the lack of an effective behavioral parameter allowing for a reliable distinction between different representations. Therefore, we employed a novel virtual reality environment that provides accuracy and reaction time measures precisely indicating the emergence of route vs. survey knowledge. Functional magnetic resonance imaging (fMRI) was used to localize brain regions involved in the acquisition of pure route knowledge in the form of associations between consecutive landmark views and the direction of intermediate movements. Participants were scanned during repeated encoding of the complex environment from a first-person, ground-level perspective. Behavioral data revealed emerging route knowledge in 11 out of 14 subjects. Overall comparisons between encoding and control conditions identified activation in medial frontal gyrus, retrosplenial cortex and posterior inferior parietal cortex. Most importantly, only posterior inferior parietal regions showed increasing activation across sessions, thus paralleling behavioral measures of route expertise. Given the established role of the posterior parietal cortex in spatial processing, this area is thought to provide the pivotal spatial link between two landmarks encountered in immediate temporal succession.

© 2004 Elsevier B.V. All rights reserved.

Theme: Neural basis of behaviour

Topic: Cognition

Keywords: Human spatial memory; Route learning; Virtual reality; Parietal cortex

1. Introduction

Human spatial navigation is the expression of a complex ensemble of cognitive processes based on distinct spatial representations. Survey representations refer to an enduring, goal-independent representation of the environment. These representations—also termed cognitive maps or configurational knowledge—allow the individual to infer spatial relations between any two places irrespective of his own position. In contrast, route knowledge conveys information

about a goal-directed, temporo-spatial sequence of environmental features. Salient objects (i.e. landmarks) in the environment are linked by their temporal occurrence during navigation and by spatial relations that determine movement decisions necessary to get from one landmark to the next (i.e. turn right). These spatial relations are encoded predominantly between neighboring objects [21,42], presumably with respect to a place-dependent local reference system [40]. In addition, the connections between landmarks appear to be coded only in the direction experienced during navigation [45].

Even though different representations can be the result of experiencing an environment from different perspectives [46,47], ground-level navigation by itself not only entails route knowledge [21,34,53], but can also result in survey

* Corresponding author. Tel.: +49 40 42803 5778; fax: +49 40 42803 9955.

E-mail address: wolbers@uke.uni-hamburg.de (T. Wolbers).

representations [43,51]. Hierarchical models have therefore conceptualized navigational learning as a succession of landmark, route and survey knowledge, the latter depending on a qualitative shift in the knowledge representation [48,51]. However, the importance of both the applied learning strategy and general spatial abilities for the resultant representation has been emphasized repeatedly, arguing for the possibility to develop route or survey knowledge from the very beginning of a learning experience [4,28].

Whereas many neuroimaging and lesion studies on spatial navigation have successfully investigated the retrieval of previously learned environments [1,24,26,31,32,35], the neural foundations of learning complex spatial layouts have received less attention [3,7,33,46]. Even though these studies have identified a network of areas including frontal, posterior parietal, retrosplenial and medial temporal regions the precise role of these structures remains controversial, particularly with respect to parietal function. For example, whereas Barrash et al. [7] did not find a systematic relationship between parietal lesions and route learning impairment, Shelton and Gabrieli [46] observed greater inferior parietal activity in an fMRI study for route as compared to survey encoding. One reason might be the lack of a behavioral measure allowing for a clear identification of the type of knowledge that was acquired, since many spatial memory tasks can be performed based on either route or survey representations [16]. In nonhuman primates, the importance of the posterior parietal cortex for route learning and retrieval has been demonstrated repeatedly [8,52], arguing for its involvement in processing the spatial relationship between local environmental cues instead of representing the allocentric position of a goal or a fixed sequence of movements. In addition, subdivisions of the posterior parietal cortex have consistently been associated with general spatial abilities like mental rotation [6,23,25] or perspective taking [54] in humans, presumably reflecting the occurrence of mental spatial transformations. However, it remains unclear to what extent these areas are implicated in route learning as well.

In the present study, we aimed to characterize the neural system involved in one type of navigational learning, the acquisition of pure route knowledge. Route knowledge was conceptualized as an association between landmarks encountered in immediate temporal succession and the spatial relations connecting them. A complex virtual environment was designed enabling us to link behavioral performance directly to the gradual emergence of route knowledge. Reaction times and performance measures were obtained from judgments of spatial relations between pairs of buildings from adjacent intersections. With regard to the expected brain activation patterns, we formulated the following hypotheses: One important prerequisite for route learning consists of storing the temporal sequence of landmarks. This may entail involvement of medial frontal areas, given that they repeatedly have been associated with

memory for temporal order and context [27,50]. More importantly, the spatial relations between neighboring landmarks constitute the crucial links that can effectively guide navigational behavior. These relations presumably are encoded with respect to multiple local reference systems [40] that are imposed on distinct places within the environment. Considering the egocentric nature of these reference systems, we predicted a prominent involvement of the posterior parietal cortex in coding the spatial relations; an assumption based on its well-established role in the processing of spatial positions of external objects in various egocentric reference frames [5,19,20,22].

2. Materials and methods

2.1. Participants

Fourteen healthy male volunteers (mean age: 29.3 years, S.D.: 5.4, range: 23–35) with normal or corrected-to-normal vision gave written informed consent to participate in this study. The study 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.

2.2. Experimental stimuli

In order to localize brain regions representing the emergence of route knowledge, a desktop virtual environment (Fig. 1) was constructed using Milkshape 3D (Chumbalum Soft, www.milkshape3d.com) for modeling and Blitz 3D (Blitz Research, www.blitzbasic.com) for animation purposes. The layout was derived from the “hexatown” environment developed by Gillner and Mallot [21]. This particular design provides the important advantage of allowing for a route to be planned that covers the entire town by passing through every road section only once in each direction. The roads were flanked on both sides by brick stone walls, thus restricting the view to the immediate straight ahead; remote road sections or intersections were invisible. Twelve distinct buildings (e.g. store, court house) 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 Section 2.3). In order 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 retrieval sessions were created by taking snapshots of all buildings from the same viewpoints as were encountered during navigation. These images were of the original size; in addition, smaller versions of each image

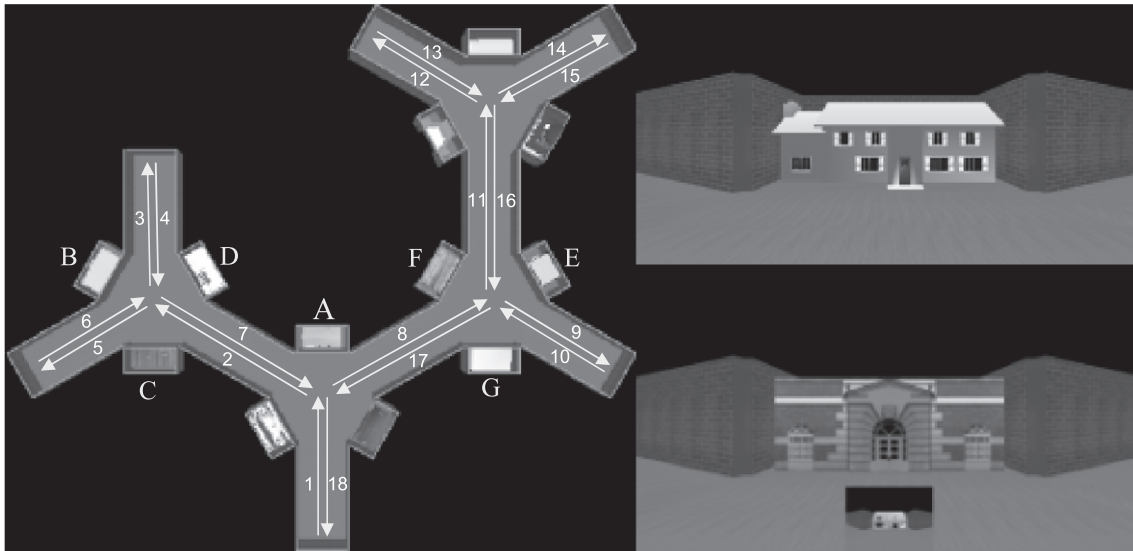


Fig. 1. Example views from the virtual environment. Left panel: Aerial view of the environment (not shown to participants). Buildings were hidden behind walls unless the midpoint of an intersection was reached. Arrows and numbers indicate one of the three routes; road sections were visited in ascending 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. Upper right panel: Ground-level view of 1 of the 12 landmarks. Lower right panel: Example of a superimposed image used for retrieval. Subjects were to indicate by button press the relative position of the small building, imagining they were standing in front of the large building. Given the route depicted in the left panel, 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 visited not immediately after the large building).

were also created. In order to test subjects’ ability to retrieve information about spatial relations between pairs of buildings, stimuli were superimposed as shown in the lower right panel of Fig. 1. 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.

2.3. Procedure

Six encoding, 6 retrieval, 3 encoding control and 3 retrieval control conditions were presented in separate sessions with a predetermined order, resulting in a total of 18 sessions for each subject. Encoding and retrieval sessions were administered in alternating order, with the control sessions taking place in the beginning, in the middle and at the end of the experiment. During encoding, participants were moved throughout the entire environment following the same route in all sessions. Subjects were instructed to memorize spatial relationships between landmarks to ensure the best possible performance during retrieval. As noted above, landmarks were hidden behind walls most of the time. However, when the midpoint of an intersection was reached, following a randomized delay between 1000 and 5000 ms, the wall in front disappeared, thereby unveiling the building behind. After 1500 ms, the subject (i.e. camera) performed a 60° turn (at $116.5^\circ/\text{s}$) and the wall was reinstated. This procedure was deliberately chosen to ensure a tight control over the exact time point at which landmarks were processed. If the buildings had been visible throughout

navigation, these time points could not be unequivocally determined. Some buildings may be consciously perceived from far away whereas others might reach awareness only shortly before arrival at an intersection. As our primary interest was to investigate learning-related processes during encoding of specific sections of the environment in an event-related design, it was very important to be able to clearly define decisive time points.

Upon reaching a dead-end, a 180° turn (at $116.5^\circ/\text{s}$) was performed and the journey continued to the previously visited intersection. Three different routes were programmed and pseudo-randomly assigned to participants. During control sessions, subjects were moved repeatedly back and forth through a single corridor. Close to each dead end, the wall would disappear following a randomized delay lasting between 1000 and 5000 ms to reveal one of the 12 buildings encountered during encoding. After 1500 ms, the camera performed a 180° turn (at $116.5^\circ/\text{s}$), the wall was reinstated and the virtual movement was continued towards the opposite wall. Each time a wall vanished, a different landmark was seen, allowing for the presentation of all 12 buildings within 6 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. As we ensured that visual stimulation was carefully matched between control and encoding conditions, activation differences could be reliably attributed to the learning of the spatial layout.

During retrieval, pairs of original and small-sized landmark images were presented for 4000 ms. Participants were

instructed to imagine they were facing the large building and to assess the relative spatial position of the small building (referred to as 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; stimulus onset asynchrony randomly varied between 7000 and 9000 ms. Six pairs depicted buildings encountered in immediate succession during encoding; these pairs will be 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 will be referred to as close pairs (Fig. 1, pairs A–C/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 will be referred to as remote pairs (Fig. 1, pairs A–E/A–F/A–G). In summary, while 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. These delays were shortest for direct pairs, followed in ascending order by close and remote pairs.

The rationale behind this approach was to establish a behavioral measure allowing for a clear cut distinction between acquisition of route knowledge and survey representations. If subjects were able to infer a cognitive map of the environment from ground-level navigation, no significant behavioral differences (response time and accuracy) between direct, close and remote pairs would be expected, since spatial distances were identical. However, if the spatial representations predominantly consisted of temporo-spatial associations between consecutive landmark views, a different outcome would have to be predicted. In this case, direct pairs should reveal a clear behavioral advantage over both close and remote pairs, since only direct pairs contained landmarks encountered in immediate temporal succession during encoding.

We also included a control condition for retrieval sessions that did not require the retrieval of spatial representations. In this condition, all 18 pairs used for retrieval, plus 6 additional pairs depicting the same building in regular and small size, were presented along with 6 null events. Subjects were instructed to indicate by a button press whether both buildings were identical (right button) or different (left button). The timing parameters were identical to those in the retrieval sessions. After fMRI

scanning concluded and without prior knowledge, subjects were asked to draw a map of the environment as accurately as possible.

Beforehand, we performed extensive behavioral testing with different participants to ensure the behavioral validity of this new paradigm. Whereas the distinction between acquiring route knowledge and survey representations turned out to be successful, the majority of subjects reported silent naming of landmarks during encoding for orientation purposes. We were thus concerned about the possible variance related to the presence or absence of silent naming. As a consequence, prior to fMRI scanning, participants were presented with all 12 buildings accompanied by names that had to be memorized. Subjects were then asked to retrieve the name of each building; in case of failure, the experimenter provided the correct name. This procedure continued until each participant could reliably name each landmark. We therefore were confident to have minimized the above mentioned source of interindividual variance.

2.4. MRI acquisition

MR scanning was performed on a 1.5 T MRI Scanner (Siemens Vision) with a standard headcoil. Twenty-eight axial slices (slice thickness: 3 mm, 1 mm gap) were acquired using a gradient echoplanar (EPI) T2*-sensitive sequence (TR=2.24 s, TE=40 ms, flip angle 90°, matrix 64*64, field of view 192*192 mm). A high resolution (1*1*1 mm voxel size) T1-weighted structural MRI was acquired for each volunteer using a 3-D FLASH sequence.

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 (10*15° field of view). To minimize head movement, all subjects were stabilized with tightly packed foam padding surrounding the head. Retrieval sessions were controlled by a PC running the “Presentation” package (Neurobehavioral Systems, www.neurobs.com). Encoding sessions were presented using Blitz 3D.

2.5. Image processing and statistical analysis

Image processing and statistical analysis were carried out using SPM2 (www.fil.ion.ucl.ac.uk/spm). All volumes were realigned to the first volume, spatially normalized [17] to a standard EPI template [14] and finally smoothed using a 10-mm full-width at half-maximum isotropic Gaussian kernel. To test the hypothesis of learning associations between consecutive landmark views and the direction of intermediate movements, we defined six events during each encoding session. These events lasted from the beginning of the presentation of a first building until the end of the presentation of a second building, thus encompassing the presentation of both buildings as well as the movement in between. Encoding during navigation within dead-ends was

not modeled. Data analysis was performed by modeling the onsets of the presentation of the first building as trains of delta functions convolved with a hemodynamic response function (HRF). An event-related approach was used since the stimulus onset asynchrony was randomized and not constant. Regression coefficients for all regressors were estimated using least squares within SPM2. Specific effects were tested with appropriate linear contrasts of the parameter estimates for the HRF regressor of all trial types, resulting in a t -statistic for each voxel. These t -statistics constitute a statistical parametric map (SPM). SPMs are interpreted by referring to the probabilistic behavior of Gaussian random fields. 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. Separate contrast images for each of the nine regressors were then generated. At the group level, a random effects approach [18] was applied to obtain results that could be generalized beyond the subjects taking part in this study. This was done by entering the contrast images of the first-level analysis into a repeated measures analysis of variance, thus allowing for computing regression coefficients across subjects for each session separately. The problem of non-independent data within subjects as well as error variance heterogeneity was addressed by performing a non-sphericity correction. The threshold was set to $p < 0.05$ corrected for multiple comparisons unless otherwise specified. To reveal performance-related activation decreases across encoding sessions, normalized mean reaction times obtained during retrieval were used as contrast weights. To look for activation increases across encoding sessions, these normalized reaction times were inverted and then used as contrast weights.

3. Results

3.1. Behavioral data

Participants were classified as route learners based on behavioral performance; in addition, we checked the maps drawn after fMRI scanning for global and/or local inconsistencies. According to our hypotheses, acquisition of route knowledge should manifest itself in performance improvements only for direct pairs. This behavioral pattern was observed in 11 out of 14 participants as reflected by response times and accuracy measures during retrieval (see Fig. 2). Whereas performance for both close and remote pairs remained close to chance level, accuracy rates for direct pairs showed a significant increase across sessions. These results were confirmed by a 3×6 within-subjects analysis of variance (factors: type of pair and retrieval session) with performance measures as the dependent variable. We observed a significant overall difference between direct, close and remote pairs (main effect type of pair: $F=10.647$, $p < 0.01$), in contrast, nonsignificant results were obtained for main effect retrieval session ($F=2.092$, $p=0.116$) and type of pair \times retrieval session interaction ($F=1.603$, $p=0.177$). The latter was primarily caused by the absence of significant changes across sessions for both close and remote pairs as indicated by linear regression analyses of the learning curves. Planned comparisons yielded regression coefficients (representing the slopes of the regression lines) significantly greater than zero only for direct pairs ($p < 0.001$), no significant results were obtained for close ($p < 0.29$) and remote pairs ($p < 0.37$). The same procedure was applied to response time data showing learning related decreases across sessions only for direct pairs. Even though we did not observe significant effects in the within-subjects ANOVA (main effect type of pairs: $F=1.525$, $p < 0.29$; main effect retrieval session:

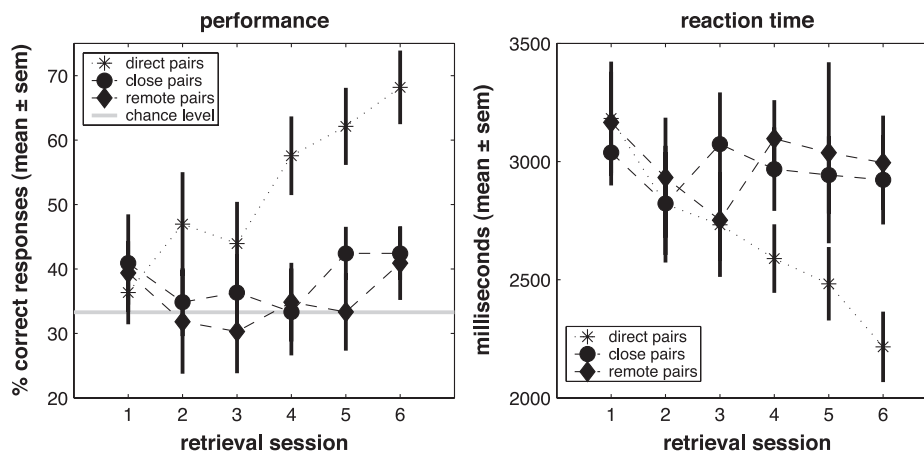


Fig. 2. Behavioral results during retrieval. Performance and reaction time data for 11 route learners. Only correct responses were included for computing mean reaction times. Significant changes across sessions were obtained for direct pairs only, indicating that learning was restricted to the acquisition of route knowledge.

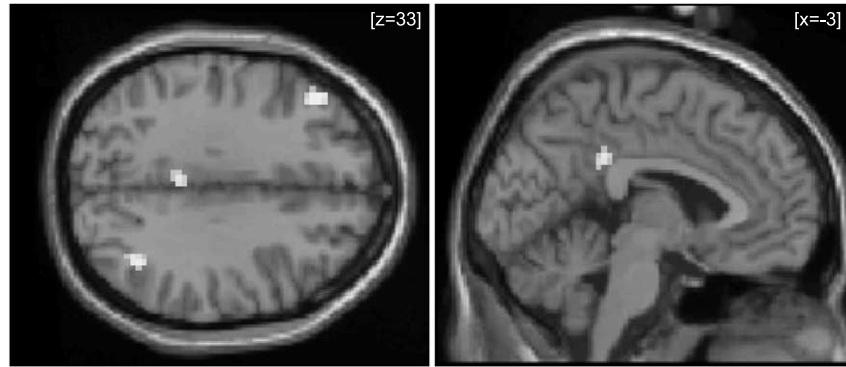


Fig. 3. Main effect of learning for route learners. Areas of significant activation for the comparison of encoding vs. encoding control (left inferior parietal cortex not shown). Results of the random effects analysis are rendered on the MNI template brain ($p < 0.05$ corrected).

$F=0.904$, $p < 0.504$; interaction: $F=0.653$, $p < 0.757$), planned comparisons of regression coefficients that were significantly smaller than zero were significant only for direct pairs ($p < 0.001$) as opposed to close ($p < 0.43$) and remote pairs ($p < 0.44$). The absence of significant effects in the ANOVA was caused by increased variance as reaction times were only calculated for correct answers. Given that several subjects did not produce any correct answers for close or remote pairs in some sessions, this led to a variance increase due to missing values in several cells.

All 11 subjects were unable to correctly reproduce the layout of the environment as reflected by the maps drawn after fMRI scanning. Both global and local inconsistencies were present; in addition, several subjects even failed to recall the number of intersections or the angles between roads. However, we did not perform a quantitative analysis of map drawing performance given that maps based on different representations can yield similar quantitative results [46]. As a consequence, it appears difficult to identify the nature of a spatial representation based on bidimensional regression coefficients or related measures.

3.2. Imaging data

3.2.1. Main effect of learning

We first contrasted encoding with encoding control conditions to look for areas showing an overall activation due to learning the temporo-spatial relations. Fig. 3 displays the group results from this contrast, Table 1 shows the locations of the activation maxima according to Montreal Neurological Institute space [14]. According to results from previous studies, multiple regions of interest were defined (retrosplenial cortex, inferior and superior parietal cortex, parahippocampal gyrus, hippocampus, medial frontal gyrus, caudate nucleus) and correction for multiple comparisons was based on these regions. A network of areas within several regions of interest was activated, including bilateral posterior inferior parietal cortices, left retrosplenial cortex and left medial frontal gyrus. In contrast, we obtained neither significant clusters of activation outside the regions of interest nor in the hippocampus or the parahippocampal gyri.

3.2.2. Learning related changes across encoding sessions

To assess whether learning related behavioral effects were paralleled by systematic changes in cortical activation during encoding, we modeled activation increases across encoding sessions. Areas showing a significant main effect of learning as revealed by the categorical comparison between encoding and encoding control (see previous section) were defined as regions of interest and correction for multiple comparisons was based on these regions. Elsewhere in the brain, correction for multiple comparisons was based on the entire search volume. In addition, to exclude areas exhibiting unspecific time effects unrelated to learning, the relevant contrast was masked exclusively by areas revealing increasing activation across encoding control sessions. Thus, we were able to identify regions showing increasing activation only across encoding sessions. The results are shown in Fig. 4 along with mean regression coefficients for the local activation maxima.

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 (t-score)
	RH	LH	
<i>Local activation maxima for the main effect of encoding</i>			
Medial frontal gyrus		-45, 33, 33	3.27
Retrosplenial cortex		-3, -39, 33	2.83
Inferior posterior parietal lobe		-36, -69, 51	2.88
	39, -60, 33		3.06
<i>Local activation maxima for performance-related increase</i>			
Medial frontal gyrus		-45, 33, 33	1.63 (n.s.)
Retrosplenial cortex		-3, -39, 33	1.44 (n.s.)
Inferior posterior parietal lobe		-51, -57, 39	3.82
		-39, -63, 33	3.62
		-42, -60, 27	3.56
	54, -48, 45		3.11
	45, -69, 42		3.10
	57, -66, 24		3.02

RH—right hemisphere, LH—left hemisphere, n.s.—non-significant.

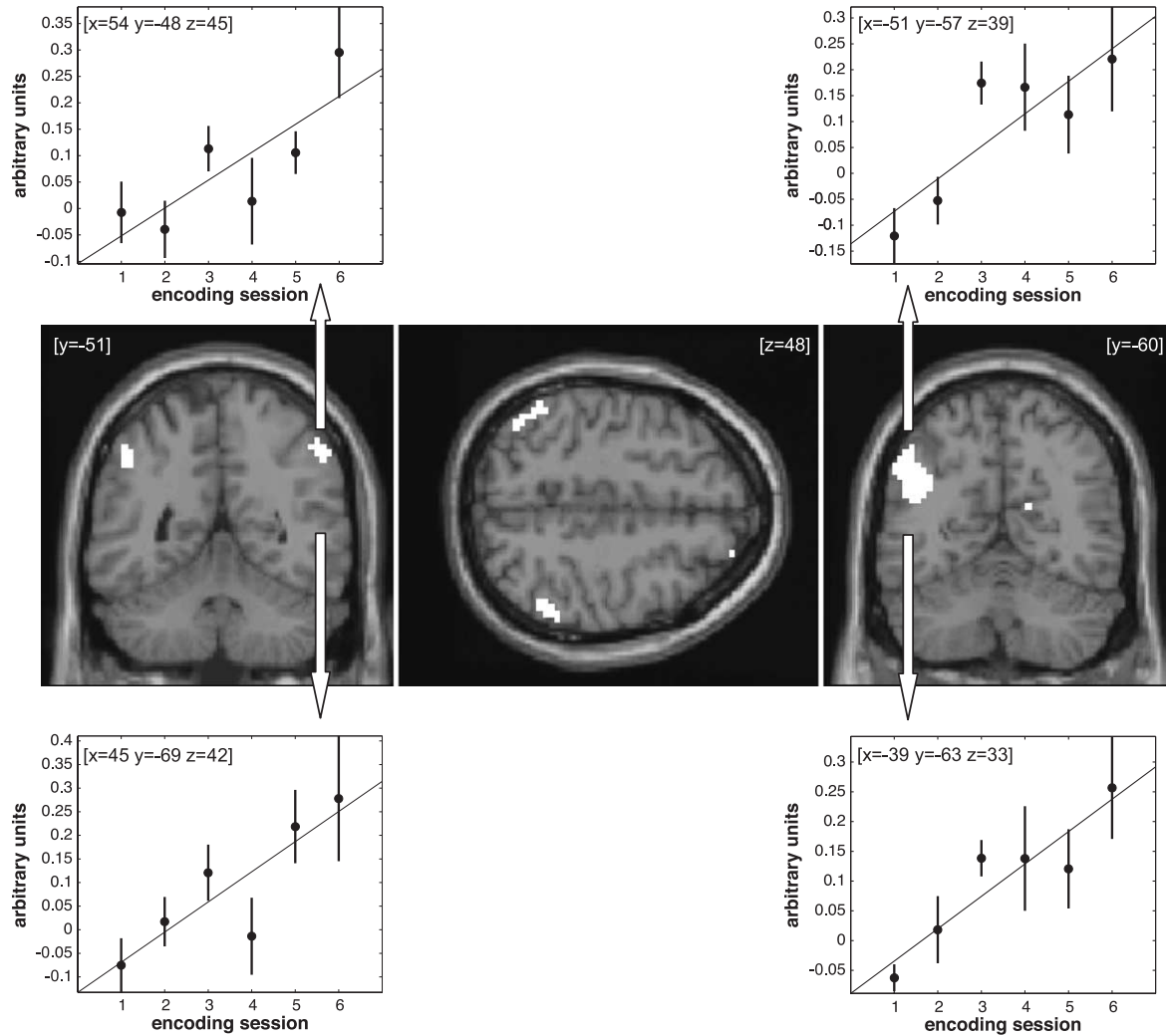


Fig. 4. Learning-related changes for route learners. Areas showing a significant activation increase across encoding sessions. Note that the specified contrast excluded regions with an increase across encoding control sessions. Upper and lower panels display mean regression coefficients of peak voxels in both parietal regions.

Areas with significantly increasing BOLD responses were found only in bilateral posterior inferior parietal cortices; none of the remaining regions of interest contained supra-threshold voxels (*t*-values of peak voxels from all regions of interest are presented in Table 1). No further activations were observed elsewhere in the brain.

To further elucidate the relationship between behavioral data and changes in BOLD signal, we subjected individual accuracy measures as predictor variables, and individual mean regression coefficients of peak voxels in both parietal regions to linear regression analyses. Regression coefficients in the left posterior inferior parietal cortex were significant ($[x=-39, y=-63, z=33]$; $p<0.05$), indicating that higher activation was paralleled by higher accuracy. In the right posterior inferior parietal cortex, regression coefficients showed a trend towards statistical significance ($[x=54, y=-48, z=45]$; $p<0.08$).

Due to the fact that learning effects are sometimes paralleled by activation decreases, we also modeled

activation decreases across encoding sessions. Again, to exclude areas exhibiting unspecific time effects unrelated to learning, this contrast was masked exclusively by areas showing decreasing activation across encoding control sessions. However, we did not obtain areas with significant activation in this analysis.

In total, whereas overall activations due to route learning were found in left medial frontal gyrus, left retrosplenial cortex and bilateral posterior inferior parietal cortices, significant learning-related activation increases across encoding sessions mimicking behavioral measures were confined to the posterior inferior parietal cortex.

4. Discussion

In the present study, our aim was to determine the neural structures involved in route learning defined as temporospatial associations between consecutive landmarks. There-

fore, we developed a behavioral paradigm that provided an objective distinction between route knowledge and survey representations. Eleven out of 14 participants showed evidence of pure route learning, reflected by increasing accuracy and decreasing response times during retrieval. Importantly, in these subjects, no behavioral evidence was found for the emergence of allocentric knowledge in terms of a cognitive map. Neuroimaging results revealed a network of structures involved in route learning, including left medial frontal gyrus, left retrosplenial cortex and bilateral posterior inferior parietal cortices. Most importantly, only posterior inferior parietal regions showed increasing activation across learning sessions, thus paralleling behavioral measures. Consequently, these areas can be regarded as crucial for establishing associations between landmarks encountered in immediate temporal succession.

Both behavioral measures revealed significant learning effects only for direct pairs, indicating that spatial learning was confined to forming temporo-spatial associations between consecutive landmark views. It is important to note that the observed learning effects cannot be explained by simple stimulus-response associations. If subjects had based their decisions during retrieval only on the direction of the movement associated with a landmark during encoding, performance improvements for close pairs should be observable as well. This follows from the fact that in close pairs, the target building was located on the adjacent intersection that was visited immediately after the large building, so choosing the movement associated with the large building would be a correct answer. Given that no performance improvements were obtained for close pairs, the formed associations obviously included not only a landmark and the direction of the following movement, but also the subsequently encountered landmark.

When contrasting encoding with the control condition, we observed activation in left retrosplenial cortex, left medial frontal gyrus and bilateral posterior inferior parietal cortices. The retrosplenial involvement in human navigation has been demonstrated both in neuroimaging and lesion studies [2,30]. Anatomically, retrosplenial areas are closely linked to various medial temporal regions, posterior parietal cortex and mid-dorsolateral prefrontal cortex in the monkey brain. It has consequently been hypothesized to serve as a transition zone between inputs in egocentric and allocentric reference frames [30], allowing for the conversion of stimulus information into the appropriate reference frame for the task at hand. Although our study did not address this conversion process, the retrosplenial activation is in accord with repeated findings from lesion studies. Patients with retrosplenial lesions have been described to suffer from an inability to form or recall links between landmark identity and directional information [2], which was one of the major tasks in our experiment.

Activity in the medial frontal gyrus has not only been observed during encoding or retrieval of large-scale environments [35,46], but also in spatial-working memory

tasks [29] as well as during retrieval of information about temporal order and context [27,50]. The latter is especially important in the present study, as the order in which landmarks were encountered provides the important temporal information necessary to form associative links between subsequent landmarks. Hence, the recruitment of the medial frontal gyrus in our paradigm might reflect working memory processes necessary for successful long-term encoding not only of spatial, but in particular of the temporal relations between landmarks.

In regard to the performance-independent medial frontal and retrosplenial activations, these regions appear to be necessary but not crucial sites for route learning. The cognitive processes subserved by these areas will presumably also be associated with various kinds of spatial and nonspatial learning. In contrast, the observed parallel between posterior inferior parietal activations during encoding and behavioral measures indicates a crucial role of the posterior inferior parietal cortex for successful learning of spatial relations between consecutive landmarks. In this context, it is important to attempt a more detailed delineation of the nature of the subserved function. Even though route knowledge has often been reduced to a mere sequence of stimulus-response pairs that, in effect, is ‘empty’ between landmarks, recent evidence suggests the inclusion of metric information about distance and direction from early learning stages on (for a review, see Ref. [36]). Given that landmarks appear to be stored in viewpoint-dependent snapshots during route learning [44], the spatial relation between two landmarks presumably is defined relative to the observer’s heading when facing the first landmark. Constituting the basis for spatial concepts such as left/right this information might later be transformed into more abstract or verbal descriptions, but initial encoding appears to occur in an egocentric reference frame.

There is ample evidence that the posterior parietal cortex provides information about location and distance of external objects in several egocentric reference frames independent of sensory modality [5,19,20,22]. This information is used to construct a conscious representation of the extrapersonal space, to update spatial relationships and to direct movements in space [9,15,22]. Consequently, lesions of the posterior parietal cortex entail neuropsychological deficits that can impair space perception and action [10,22,37]. Regarding the parietal involvement in route learning, we believe that our results may shed light on previous findings in humans [7,46]. By isolating route learning and by observing a congruency of parietal activation with behavioral performance, it is fair to assume that the increasing activation reflects increasingly robust egocentric encoding of the spatial relation between two landmarks. The better this spatial link is established in memory, the faster and more accurate this information can be accessed during retrieval. This interpretation would be in line with findings from nonhuman primates, demonstrating that parietal

lesions entail difficulties in the processing of local spatial relationships during route retrieval [8,52]. Interestingly, inferior parietal lesions do not produce a retrograde amnesia for the allocentric position of a goal, but selectively impair local visual processing [8]. However, it should be pointed out that in the context of learning, the posterior parietal cortex presumably is not exclusively involved in establishing route representations, but may also be important for extracting survey information. For example, whenever a landmark is encountered during encoding, computing its egocentric location may be a necessary step to infer its position within an allocentric reference system. Further studies comparing route vs. survey learning from the same encoding perspective are needed to determine whether the posterior parietal cortex supports the emergence of other types of space representation as well.

It is interesting to note that the increasing parietal activation during encoding associated with better recall performance can be conceptualized as a subsequent memory effect. This effect has been demonstrated in various paradigms [11,38,39], and it is often suggested that higher activity during encoding in task-relevant areas reflects deeper levels of information processing, thereby inducing a more robust representation of a stimulus. Whereas our paradigm only allowed us to assess item-related effects given that we employed an event-related design, further studies are needed to clarify whether state-related effects also contribute to successful encoding in the context of topographical learning.

In the present study, we did not observe activation in either the hippocampus or the parahippocampal gyrus. Although the null hypothesis (no involvement of these structures in route learning) cannot be verified, our data are in line with the idea that the hippocampus may provide an allocentric representation of the environment. Since our study exclusively addressed route learning, hippocampal involvement would not be expected upon the base of the cognitive map theory. The lack of parahippocampal activation could be due to the fact that its primary role in spatial navigation consists of viewpoint-dependent processing of complex spatial layouts [3,12,13]. During control conditions, all 12 landmarks were presented in an identical fashion as compared to encoding. Consequently, parahippocampal involvement would be expected in both conditions, with the effect that these activations should cancel each other out in direct comparisons. In addition, learning related changes across sessions would also not be predicted if the above mentioned hypothesis were true.

Even though virtual reality displays provide the unique opportunity to study topographical orientation in complex environments during fMRI scanning, caution is appropriate when transferring the results to real-world navigation. Given the lack of vestibular and sensorimotor feedback, spatial updating seems to be hampered. For example, greater alignment effects and distance estimation errors have been observed after virtual reality as compared to real-world

navigation, probably due to difficulties with integrating and scaling information about rotation and locomotion [41,43]. However, these differences appear to decline with extended navigation experience, leaving virtual reality a valuable tool for investigating the neural correlates of topographical orientation in manipulable environments.

A further problem consists of the potential confounds due to either attentional effects or to eye movements that are likely to occur in any navigation task. However, it is unlikely that the amount of eye movements increases across sessions, as subjects were engaged in acquiring route knowledge from the very beginning of the experiment. With regard to attention, one would expect that attentional effects should be strongest in the early phase of the experiment, given the virtual environment's initial high degree of novelty. With repeated exposure to the identical environment, these effects should attenuate, thereby leading to an activation decrease. As a consequence, it seems reasonable to assume that the most important result of our study—the learning-dependent, increasing activation observed in inferior parietal cortex—can be attributed rather to the spatial learning component of our task than to eye movements or attention.

Human spatial navigation is based on a dynamic system of coexisting egocentric and allocentric representations of the environment. It is highly flexible, as it enables us to switch between different representations depending on what is optimal in each situation. However, between-subject variations have been demonstrated in behavioral experiments, questioning the assumption that identical information is accessed across participants during navigation in familiar environments [49]. In addition, even when navigation is based on a single representation, different sources of directional information can be used in different situations [49]. As a consequence, neuroimaging results in studies of topographical learning and spatial navigation should be closely tied to behavioral measures that permit a reliable assignment of cortical activation to cognitive processes. We believe to have established a paradigm that enables us to draw a general distinction between route and survey representations. When subjects showed evidence for pure route learning bilateral posterior inferior parietal regions could be identified as crucial for constructing a spatial link between consecutive landmarks. Within the framework of route representations, further studies might apply more finessed tasks to elucidate differences between various sources of spatial information that can be used to infer route knowledge.

The results of the current study do not preclude the involvement of posterior parietal, retrosplenial and medial frontal regions in other types of navigational learning. The emergence of survey or map like representations out of ground level navigation requires precise egocentric encoding of spatial relations in posterior parietal cortex as well; however, the subsequent transformation into an allocentric reference frame may crucially depend on retrosplenial and

medial temporal functioning. One potential prediction would therefore consist of learning related activation in these regions, especially against the background of performance dependent medial temporal activation observed during retrieval of allocentric representations [24]. In addition, the complexity of integrating sequential views of an environment into a unified, coherent representation calls for dynamic interactions between parietal, retrosplenial and medial temporal regions.

Acknowledgements

We thank the Physics and Methods group at NeuroImage Nord in Hamburg, Ron Paludan (www.railwaystation.com) for providing several 3D-models, and D. Waller, K.F. Richter, F. Binkofski, E. Schoell, A. McNamara and D. Gonzalo for suggestions on an earlier draft of this paper. This work was supported by the Volkswagenstiftung and the Bundesministerium für Bildung und Forschung (BMBF).

References

- [1] G.K. Aguirre, M. D'Esposito, Environmental knowledge is subserved by separable dorsal/ventral neural areas, *J. Neurosci.* 17 (1997) 2512–2518.
- [2] G.K. Aguirre, M. D'Esposito, Topographical disorientation: a synthesis and taxonomy, *Brain* 122 (1999) 1613–1628.
- [3] G.K. Aguirre, J.A. Detre, D.C. Alsop, M. D'Esposito, The parahippocampus subserves topographical learning in man, *Cereb. Cortex* 6 (1996) 823–829.
- [4] G.L. Allen, Spatial abilities, cognitive maps and wayfinding: bases for individual differences in spatial cognition and behavior, in: R. Golledge (Ed.), *Wayfinding Behavior*, Johns Hopkins, Baltimore, 1999, pp. 46–80.
- [5] R.A. Andersen, L.H. Snyder, D.C. Bradley, J. Xing, Multimodal representation of space in the posterior parietal cortex and its use in planning movements, *Annu. Rev. Neurosci.* 20 (1997) 303–330.
- [6] J. Barnes, R.J. Howard, C. Senior, M. Brammer, E.T. Bullmore, A. Simmons, P. Woodruff, A.S. David, Cortical activity during rotational and linear transformations, *Neuropsychologia* 38 (2000) 1148–1156.
- [7] J. Barrash, H. Damasio, R. Adolphs, D. Tranel, The neuroanatomical correlates of route learning impairment, *Neuropsychologia* 38 (2000) 820–836.
- [8] C.J. Barrow, R. Latt, The role of inferior parietal cortex and fornix in route following and topographic orientation in cynomolgus monkeys, *Behav. Brain Res.* 75 (1996) 99–112.
- [9] Y.E. Cohen, R.A. Andersen, A common reference frame for movement plans in the posterior parietal cortex, *Nat. Rev., Neurosci.* 3 (2002) 553–562.
- [10] M. Critchley, *The Parietal Lobes*, Hafner Press, New York, 1953.
- [11] L. Davachi, A. Maril, A.D. Wagner, When keeping in mind supports later bringing to mind: neural markers of phonological rehearsal predict subsequent remembering, *J. Cogn. Neurosci.* 13 (2001) 1059–1070.
- [12] R. Epstein, K.S. Graham, P.E. Downing, Viewpoint-specific scene representations in human parahippocampal cortex, *Neuron* 37 (2003) 865–876.
- [13] R. Epstein, A. Harris, D. Stanley, N. Kanwisher, The parahippocampal place area: recognition, navigation, or encoding? *Neuron* 23 (1999) 115–125.
- [14] A.C. Evans, D.L. Collins, D.R. Mills, E.D. Brown, R.L. Kelly, T.M. Peters, 3D statistical neuroanatomical models from 305 MRI volumes, *Proc. IEEE Nucl. Sci. Symp. Med. Imaging* 1–3 (1993) 1813–1817.
- [15] M.J. Farrell, I.H. Robertson, The automatic updating of egocentric spatial relationships and its impairment due to right posterior cortical lesions, *Neuropsychologia* 38 (2000) 585–595.
- [16] M.O. Franz, H.A. Mallot, Biomimetic robot navigation, *Robot. Auton. Syst.* 30 (2000) 133–153.
- [17] K.J. Friston, J. Ashburner, C.D. Frith, J.B. Poline, J.D. Heather, R.S.J. Frackowiak, Spatial registration and normalization of images, *Hum. Brain Mapp.* 2 (1995) 1–25.
- [18] K.J. Friston, A.P. Holmes, K.J. Worsley, How many subjects constitute a study? *NeuroImage* 10 (1999) 1–5.
- [19] G. Galati, G. Committeri, J.N. Sanes, L. Pizzamiglio, Spatial coding of visual and somatic sensory information in body-centred coordinates, *Eur. J. Neurosci.* 14 (2001) 737–746.
- [20] G. Galati, E. Lobel, G. Vallar, A. Berthoz, L. Pizzamiglio, D. Le Bihan, The neural basis of egocentric and allocentric coding of space in humans: a functional magnetic resonance study, *Exp. Brain Res.* 133 (2000) 156–164.
- [21] S. Gillner, H.A. Mallot, Navigation and acquisition of spatial knowledge in a virtual maze, *J. Cogn. Neurosci.* 10 (1998) 445–463.
- [22] P.W. Halligan, G.R. Fink, J.C. Marshall, G. Vallar, Spatial cognition: evidence from visual neglect, *Trends Cogn. Sci.* 7 (2003) 125–133.
- [23] I.M. Harris, G.F. Egan, C. Sonkkila, H.J. Tochon-Danguy, G. Paxinos, J.D. Watson, Selective right parietal lobe activation during mental rotation: a parametric PET study, *Brain* 123 (2000) 65–73.
- [24] T. Hartley, E.A. Maguire, H.J. Spiers, N. Burgess, The well-worn route and the path less traveled. Distinct neural bases of route following and wayfinding in humans, *Neuron* 37 (2003) 877–888.
- [25] K. Jordan, H.J. Heinze, K. Lutz, M. Kanowski, L. Jancke, Cortical activations during the mental rotation of different visual objects, *NeuroImage* 13 (2001) 143–152.
- [26] K. Jordan, J. Schadow, T. Wuestenberg, H.-J. Heinze, L.C.A. Jancke, Different cortical activations for subjects using allocentric or egocentric strategies in a virtual navigation task, *NeuroReport* 15 (2004) 135–140.
- [27] S. Konishi, I. Uchida, T. Okuaki, T. Machida, I. Shirouzu, Y. Miyashita, Neural correlates of recency judgment, *J. Neurosci.* 22 (2002) 9549–9555.
- [28] P.C. Kyllonen, Aptitude testing inspired by information processing: a test of the four-source model, *J. Gen. Psych.* 120 (1993) 375–405.
- [29] H.C. Leung, J.C. Gore, P.S. Goldman-Rakic, Sustained mnemonic response in the human middle frontal gyrus during on-line storage of spatial memoranda, *J. Cogn. Neurosci.* 14 (2002) 659–671.
- [30] E.A. Maguire, The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings, *Scand. J. Psychol.* 42 (2001) 225–238.
- [31] E.A. Maguire, N. Burgess, J.G. Donnett, R.S. Frackowiak, C.D. Frith, J. O'Keefe, Knowing where and getting there: a human navigation network, *Science* 280 (1998) 921–924.
- [32] E.A. Maguire, R.S. Frackowiak, C.D. Frith, Recalling routes around London: activation of the right hippocampus in taxi drivers, *J. Neurosci.* 17 (1997) 7103–7110.
- [33] E.A. Maguire, C.D. Frith, N. Burgess, J.G. Donnett, J. O'Keefe, Knowing where things are parahippocampal involvement in encoding object locations in virtual large-scale space, *J. Cogn. Neurosci.* 10 (1998) 61–76.
- [34] H.A. Mallot, S. Gillner, Route navigating without place recognition: what is recognised in recognition-triggered responses? *Perception* 29 (2000) 43–55.
- [35] E. Mellet, S. Briscogne, N. Tzourio-Mazoyer, O. Ghaem, L. Petit, L. Zago, O. Etard, A. Berthoz, B. Mazoyer, M. Denis, Neural correlates of topographic mental exploration: the impact of route versus survey perspective learning, *NeuroImage* 12 (2000) 588–600.

- [36] D.R. Montello, A new framework for understanding the acquisition of spatial knowledge in large-scale environments, in: R.G. Golledge (Ed.), *Spatial and Temporal Reasoning in Geographic Information Systems*, Oxford University Press, New York, 1998, pp. 143–154.
- [37] D.J. Mort, P. Malhotra, S.K. Mannan, C. Rorden, A. Pambakian, C. Kennard, M. Husain, The anatomy of visual neglect, *Brain* 126 (2003).
- [38] L.J. Otten, R.N. Henson, M.D. Rugg, State-related and item-related neural correlates of successful memory encoding, *Nat. Neurosci.* 5 (2002) 1339–1344.
- [39] K.A. Paller, A.D. Wagner, Observing the transformation of experience into memory, *Trends Cogn. Sci.* 6 (2002) 93–102.
- [40] B. Poucet, Spatial cognitive maps in animals: new hypotheses on their structure and neural mechanisms, *Psychol. Rev.* 100 (1993) 163–182.
- [41] A.E. Richardson, D.R. Montello, M. Hegarty, Spatial knowledge acquisition from maps and from navigation in real and virtual environments, *Mem. Cogn.* 27 (1999) 741–750.
- [42] R. Rothkegel, K.F. Wender, S. Schumacher, Judging spatial relations from memory, in: K.F. Wender (Ed.), *Spatial Cognition*, Springer, Berlin, 1998, pp. 79–106.
- [43] R.A. Ruddle, S.J. Payne, D.M. Jones, Navigating buildings in “desktop” virtual environments: experimental investigations using extended navigational experience, *J. Exp. Psychol. Appl.* 3 (1997) 143–159.
- [44] B. Schölkopf, H.A. Mallot, View-based cognitive mapping and path planning, *Adapt. Behav.* 3 (1995) 311–348.
- [45] K. Schweizer, T. Herrmann, G. Janzen, S. Katz, The route direction effect and its constraints, in: K.F. Wender (Ed.), *Spatial Cognition*, Springer, Berlin, 1998, pp. 19–38.
- [46] A.L. Shelton, J.D. Gabrieli, Neural correlates of encoding space from route and survey perspectives, *J. Neurosci.* 22 (2002) 2711–2717.
- [47] A.L. Shelton, T.P. McNamara, Orientation and perspective dependence in route and survey learning, *J. Exp. Psychol. Learn. Mem. Cogn.* 30 (2004) 158–170.
- [48] A.W. Siegel, S.H. White, The development of spatial representations of large-scale environments, in: H.W. Reese (Ed.), *Advances in Child Development and Behaviour*, vol. 10, Academic, New York, 1975, pp. 9–55.
- [49] S.D. Steck, H.A. Mallot, The role of global and local landmarks in virtual environment navigation, *Presence* 9 (2000) 69–83.
- [50] M. Suzuki, T. Fujii, T. Tsukiura, J. Okuda, A. Umetsu, T. Nagasaka, S. Mugikura, I. Yanagawa, S. Takahashi, A. Yamadori, Neural basis of temporal context memory: a functional MRI study, *NeuroImage* 17 (2002) 1790–1796.
- [51] P.W. Thorndyke, B. Hayes-Roth, Differences in spatial knowledge acquired from maps and navigation, *Cogn. Psychol.* 14 (1982) 560–589.
- [52] J. Traverse, R. Latto, Impairments in route negotiation through a maze after dorsolateral frontal, inferior parietal or premotor lesions in cynomolgus monkeys, *Behav. Brain Res.* 20 (1986) 203–215.
- [53] R.F. Wang, E. Spelke, Human spatial representation: insights from animals, *Trends Cogn. Sci.* 6 (2002) 376–382.
- [54] J.M. Zacks, J.M. Vettel, P. Michelon, Imagined viewer and object rotations dissociated with event-related fMRI, *J. Cogn. Neurosci.* 15 (2003) 1002–1018.