Spatial cognition: the return path

Kai Hamburger (kai.hamburger@psychol.uni-giessen.de)
Lena E. Dienelt (lena-eowyn.dienelt@psychol.uni-giessen.de)
Marianne Strickrodt (marianne.strickrodt@psychol.uni-giessen.de)
Florian Röser (florian.roeser@psychol.uni-giessen.de)

Justus Liebig University Giessen, Department of Psychology,
Experimental Psychology and Cognitive Science
Otto-Behaghel-Strasse 10 F
35394 Giessen, Germany

Abstract

The cognitive representation of a return path is a rather unexplored topic including different issues, e.g., perception, mental imagery, mental spatial processing, and language. We here investigated the return path with landmarks located on different positions (optimal, suboptimal). Participants learned a total of 24 routes and had to produce the return paths (*N*=20). In a second experiment the different positions plus map learning versus verbal directions were investigated (*N*=20). Both experiments reveal that the position of a landmark at an intersection (structural salience) has an influence on wayfinding performance. However, the results are somehow ambiguous. Therefore, we also present first approaches for predicting behavior (e.g., optimal route descriptions) and for modeling the perceptual and cognitive processes involved in finding the return path, including visibility, structural salience, mental representation/ transformation, and language.

Keywords: return path; structural salience; landmarks; mental transformation

Introduction

Imagine that you are on a vacation in an unknown foreign city. After your arrival at the hotel you want to explore the surroundings and maybe visit a place of interest or a touristic feature (e.g., a famous building such as the Eiffel Tower in Paris). You may base your search on different means for successfully reaching your goal. You may want to use a verbal description that you received at the reception desk of your hotel, maybe you want to make use of a city map in your tourist guide, or, if you do not have these means at hand, you may want to ask a pedestrian on the street for giving you directions to your goal location. There is also the possibility of using a mobile navigation system. This latter example is part of the debate on "extended cognition" (e.g., Clark & Chalmers, 1998), which is beyond the scope of this project. Here, the focus is rather on the "innate" navigation system, perceptual and cognitive processes that enable humans to navigate without getting lost (most of the times). In general, wayfinders use so-called landmarks, objects or buildings that stand out of their environment, to aid navigation (e.g., Lynch, 1960; Presson & Montello, 1988; Caduff & Timpf, 2008). Let us return to our initial example. One important question is whether the verbal description is on its own sufficient for reaching the goal without being distracted or being led into a wrong direction? Or, would it be better to supplement the verbal description with a map, or maybe make only use of the map instead? This is not only a question of not getting lost (e.g., Dudchenko, 2010), but also a question about cognitive economy, namely, reaching the goal with the least cognitive or physical effort. Let us assume that we successfully reached the goal. We are now faced with a new, maybe more difficult, problem. We need to return to our hotel!

Finding a return path is an everyday problem but has rarely been investigated empirically (retrace the same route; e.g., Golledge, 1997, Büchner, Hölscher, & Strube, 2007; Papinski, Scott, & Doherty, 2009). We are able to manage this task, but we do not yet know the underlying cognitive and neural processes enabling us to find the return path.

One of the most important aspects for the return path is probably the structure of the environment (e.g., structural landmark salience; Sorrows & Hirtle, 1999; Klippel & Winter, 2005). Since we assume visual salience (or better perceptual salience) –that is how much an object stands out from its environment (e.g., Caduff & Timpf, 2008)– and semantic salience of landmarks –that is for example its name, meaning, or function (e.g., Hamburger & Knauff, 2011)– to be less important, we here try to control for these aspects and rather focus on the structural aspects as we have done in several previous experiments on structural salience (e.g., Röser, Hamburger, Krumnack, & Knauff, 2012a; Röser, Krumnack, Hamburger, & Knauff, 2012b).

There are two optimal positions for landmarks to be located on a regular/initial path: *before the intersection* (Klippel & Winter, 2005) *in direction of the turn* and *behind the intersection in direction of the turn* (Röser et al., 2012a). Most important is that the landmark is located somewhere *in direction of the turn* (Röser et al., 2012a). But, for the return path, two different positions might be the optimal ones: the positions *before the intersection in direction of the turn* and *behind the intersection opposite to the direction of the turn*. These positions are invariant for the return path (they remain unchanged). The other two positions are variant, since they have to be mentally and verbally transformed for the return path (e.g., "before the intersection opposite to the direction of turn" becomes "behind the intersection and in direction of the turn" on the way back). Further details on

this theoretical assumption are provided in the section *"Theoretical assumptions, modeling, future research"*.

Experiment 1

Method

Participants

A total of 20 Psychology students from the University of Giessen participated (16 females). They had a mean age of 23.5 years (*SD*=4.08). All participants were naive with respect to this study, provided informed written consent, and received course credits for participation. They had normal or corrected-to-normal visual acuity and were free of any pre-existing psychiatric or neurologic illness (e.g., epilepsy).

Materials

The equipment included a custom 19'' monitor (Dell), a Personal Computer (HP Compaq 6000 Pro), and a Response Pad (RB-530 Cedrus Corporation©). For presentation and data recording SuperLab 4.0 Stimulus Presentation Software (Cedrus Corporation©) was employed.

The virtual environment (maze) was set up with Google© SketchUp 8 (compare to SQUARELAND; Hamburger & Knauff, 2011), which in its original version is made of 10×10 cuboids, representing regular orthogonal intersections, and proofed very flexible in terms of experimental manipulations. Here, 24 routes, each with eight intersections in an egocentric perspective, were created. The directions left or right were used. Every intersection (24×8=192) contained one distinct landmark – one of 192 different words on a white sign (Figure 1). These distinct landmark words were used to prevent interferences of previously learned landmark and direction combinations (e.g., in Route 1 you have to turn right when you see the word "horse"; in a later route you might have to turn left when you see the word "horse"). Hence, a landmark which was shown once to the participant does not appear again later in another route. We controlled for all landmarks being comparably imaginable by using familiar, everyday words. A landmark was placed on both sides of the corresponding facades of a corner, so it was visible from both directions of travel.
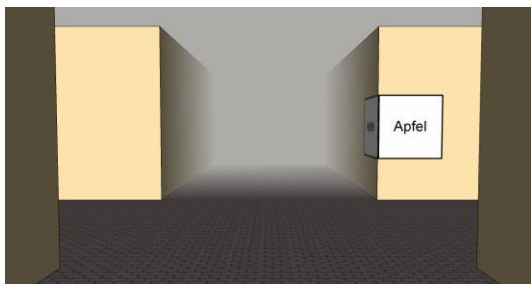

Figure 1: Screenshot of an intersection in the virtual maze (decision point). The landmark (word; Apfel = apple) is presented on both facades at one corner (position).

To control for direction or landmark position effects, the number of right/left turns and the position of landmarks (before or after the intersection, in or against moving direction) were balanced for single routes. This balancing applies to both the regular travel direction (forward) in the learning phase and to the return path in the wayfinding phase.

Procedure

Participants learned a route of eight intersections via successively presented pictures of each of the intersections (Figure 1). Every intersection was shown for duration of eight seconds (learning phase). Subsequently, participants were instructed to find the same path again either in the normal (forward from origin to destination) or the reverse travel direction (backwards from destination to origin; wayfinding phase). Every intersection was presented via pictures and served as a decision point (right or left) for which direction decisions had to be made.

After one route was navigated (eight direction decisions), the learning phase of the next route started. No feedback about the decisions was given. The total of 24 routes had to be learned by each participant. Overall, half of the routes had to be found in the forward run direction, while for the other half the return path was required. Therefore, two experimental versions were used where navigation direction in the wayfinding phase was interchanged (e.g., Route 1 had to be found again in forward direction in version 1, but in the backwards direction in version 2). The order of the routes was randomized for every participant. Correct decisions and response times served as dependent variables. At the end participants were asked to indicate any strategies they had used during the experiment.

Results

The mean correct route decisions on the return path were about 87% in this experiment (chance level 50%).

An analysis of variance with repeated measures for the wayfinding phase was performed. Within-subject factors were navigation direction (forward/backward) and landmark position (all four possible positions). Both for correct decisions and response times a significant main effect for navigation direction (correct decisions: $F(1,19)=19.865$, $p<.001$; response times: $F(1,19)=21.571$, $p<.001$), but not for landmark position (correct decisions: $F(3,57)=1.020$, $p=.391$; response times: $F(3,57)=.871$, $p=.461$) could be found. Participants were better and faster in navigating the original route direction (forward) compared to the reverse direction (backwards), but the position of a landmark did not lead to any performance differences.

A wide range and variability of learning strategies was reported by the participants and different levels of self-confidence in performance were expressed. Thus, we were interested in possible group differences. Therefore, we divided our sample in participants with an overall better (*N*=14) and an overall worse performance (*N*=6) with respect to mean overall performance. This mean-

performance-grouping now functioned as a between-subject factor in a re-analysis of variance with repeated measures. For response times again only a main effect for direction ($F(1,18)=16.196$, $p=.001$), but not for landmark position ($F(3,54)=.508$, $p<.678$) occurred. However, for correct decisions again a significant main effect for direction ($F(1,18)=22.322$, $p<.001$) and an additional interaction of navigation direction $\times$ landmark position $\times$ group ($F(3,54)=3.895$, $p=.025$) emerged. This means that the wayfinding performance with landmarks on varying positions differs with navigation direction and depends on the participant being a high or a low performer.

Discussion

An overall effect for the wayfinding direction could be found. People were faster and better when travelling the route in the originally learned direction (forward) compared to navigating the return path, which is not very surprising but has not been investigated systematically before. No landmark position effect was found. Only the mean performance of low performers indicates that some people (maybe depending on spatial ability and learning strategy) might be affected by structural differences (positions), and that the helpfulness of a landmark might differ depending on the direction of travel (forward, backwards). Such data need to be further analyzed in future research with the focus on individual strategies (wayfinding performance vs. sense-of-direction; e.g., Kato & Takeuchi, 2003).

In Experiment 2 we used a more realistic setup: video sequence from an egocentric perspective with approximated true physical sizes on a projection screen; only one route but with the option of going straight; two learning conditions; more than just eight intersections, etc.

Experiment 2

Method

Participants

A total of 20 Psychology students from the University of Giessen participated (13 females). They had a mean age of 26.1 years ($SD=9.03$). All participants were naive with respect to this study, provided informed written consent, and received course credits for participation. They had normal or corrected-to-normal visual acuity and were free of any pre-existing psychiatric or neurologic illness (e.g., epilepsy).

Materials

The same setup was used as in Experiment 1 but this time the routes were presented on a customary projection screen (171x238 cm) with a projector (Panasonic PT-F100NT).

For the experiment two different routes through the maze, with 20 intersections each, were created. Therefore, a total of 20 different words served as landmark objects (Figure 1). The words were derived from a catalog of pictograms which made them visually similar, realistic, and easy to imagine. In the maze the landmarks were again placed on both facades of a corner (position), so they were visible from both directions of travel.

Videos of the two routes were generated from an egocentric perspective, with an eye height of 1.70m and a constant walking speed of about 2m/s. For presentation and data recording SuperLab 4.0 Stimulus Presentation Software (Cedrus Corporation©) was employed.

Procedure

Participants were assigned to two groups: One of them learned a path with 20 intersections via a map, the other one through verbal description (allocentric vs. egocentric learning condition). After a five minute break, the learned path was shown as video in reverse order through the virtual maze, which was stopped at every intersection (decision point) for participants to indicate the path directions right, left, or straight (wayfinding phase). Learning condition (map/description) and landmark position (optimal/ suboptimal) served as independent variables while correct route decisions and response times served as dependent variables.

Results

With landmarks being located in (assumed) optimal positions correct decisions on the return path were made in about 67.5% (chance level 33.3%) if the initial path was learned via a route description. When the path was encoded via a map about 65% correct route decisions were made. With landmark objects being in suboptimal positions on the return path, 59% correct decisions were made for the description condition and 65% for the map condition.

For the optimal positions the response times were lower (3900ms) in the description condition, compared to the map condition (4960ms). Responses for intersections with landmarks on suboptimal positions revealed again a shorter response time for the description condition (4175ms), in comparison to the map condition (4825ms).

An analysis of variance with the within-subject factor landmark position (optimal/suboptimal) and the between-subject factor learning condition (map/verbal description) was performed. It revealed a significantly higher performance for landmarks on optimal positions ($F(1,18)=4.99$, $p=.038$). But, the position did not reveal significant differences according to the response times ($F(1,18)=.033$, $p=.858$). The learning conditions did neither differ significantly in the wayfinding phase with respect to correct decisions ($F(1,18)=.066$, $p=.800$), nor with respect to response times ($F(1,18)=.621$, $p=.441$). The three possible options of choice on the intersections (left, right, straight on) did not lead to significant differences according to correct decisions ($F(2,38)=.818$, $p=.449$). No interactions were obtained.

Discussion

The landmark position led to significant differences in performance (correct decisions), though this was not the case for the decision times. Consistent with the expectations

better decisions were made if landmarks were located on optimal positions. Since no decision time differences could be obtained, this effect cannot be due to longer viewing times for the landmarks. We may therefore conclude that the quality of a landmark as a point of reference for finding the return path very much depends on its position, as has previously been assumed for the "initial path" (forward run; Klippel & Winter, 2005; Röser et al., 2012a,b).

The different learning conditions map and description (allocentric/egocentric) did not lead to a significant difference in the wayfinding phase, neither for correct decisions nor for the response times. This absence of an effect may be explained by the "dual coding theory of human wayfinding knowledge" (Meilinger, Knauff, & Bülthoff, 2008). It assumes that environmental information is (sometimes) encoded in a spatial format alone but sometimes additionally in a verbal format. Information learned through maps (allocentric) is encoded verbally as well as information learned through descriptions (egocentric mental imagery). The similar performances after studying a map or a verbal description may be attributed to verbal representations existing for both encoding conditions (Meilinger & Knauff, 2008).

In Experiment 2 position effects were found in comparison to Experiment 1. It is possible that Experiment 1 only tested the direction memory (memory task), while Experiment 2 represents a realistic wayfinding task. Since these results are not conclusive, more theoretically driven assumptions and empirical research are required.

Theoretical assumptions, modeling, future research

In the following we present current ideas on how landmarks, places, and directions might be cognitively processed for the return path. As we have seen so far from our first two experiments on the return path and which role landmarks and landmark positions play in this context, more systematic empirical work is required.

As can be seen in Figure 2 we need to differentiate between an allocentric and an egocentric perspective. In the allocentric perspective (forward run) the assumed optimal position (Klippel & Winter, 2005; Röser et al., 2012a) is position D, *before the intersection* and *in direction of the turn*. For the forward run optimal positions have been suggested theoretically/mathematically (Klippel & Winter, 2005) and have been evaluated empirically (Röser et al., 2012a,b). For the return path the optimal positions are not yet known. We assume that position D should still be optimal, since it is before the intersection in direction of the turn (identical to the initial path), and this position is invariant independent of direction (no right/left encoding necessary). According to the findings by Röser et al. (2012a) position C could be optimal as well in the egocentric perspective and A could be optimal in both perspectives, since A is also invariant (opposite to the direction of the turn) as is the case for D.
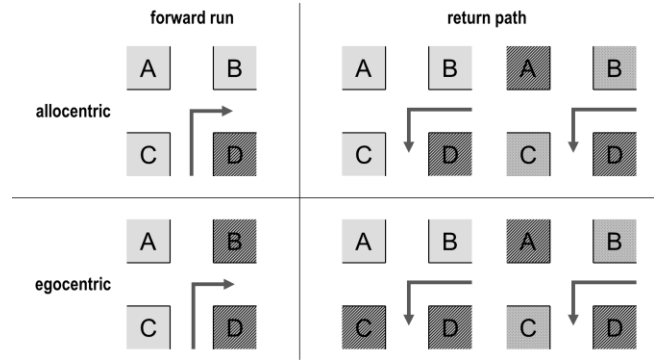

Figure 2: Possible optimal (dark gray) and suboptimal (light gray) landmark positions for the forward run and the return path in the allocentric and egocentric perspective. See text for details.

Another important issue in the egocentric perspective is the so-called "visibility" (Winter, 2003; Röser et al., 2012b). This means that different locations have different visibilities depending on the observers own position (Figure 3). Visual attention is generally paid to the direction of turn. It seems that in an egocentric perspective it is important that a landmark is at least located *in direction of the turn* and that *before* and *behind* become less important.
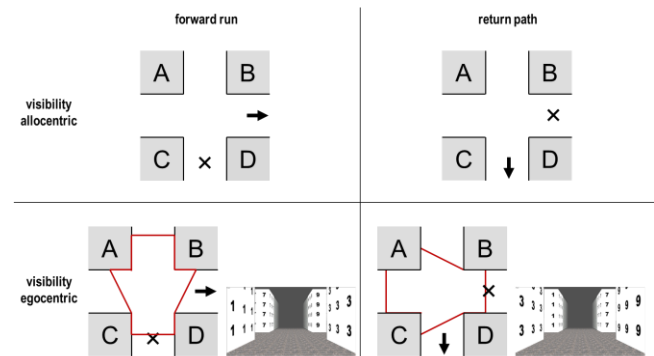

Figure 3: Visibility from two different positions: initial path (left) and return path (right). X = position of individual; → = walking direction. In the allocentric perspective each position is equally visible for both directions, not so for the return path. The small images on the bottom visualize the sight in the egocentric perspective. See text for details.

For the return path it is important to take this into account. This means that for the return path the optimal position in the allocentric perspective remains the same (D), since this location is still *before the intersection* and *in direction of the turn* (invariant; see central section of Figure 2). According to the above findings and the previous logic, in the egocentric perspective the optimal positions should now be C and D. However, position C was a suboptimal one on the forward run and therefore it may now be doubted that it becomes optimal on the return path, since it is a variant

position (forward run: *before the intersection* and *opposite to the direction of the turn*; return path: *behind the intersection* and *in direction of the turn*). This would require some additional mental transformation for the observer in order to correctly find the return path.

Now it is interesting to see that positions D and A are invariant for the initial and the return path, while B and C are variant locations (see right section of Figure 2). But, this is only the case if the spatial information is *unspecific*; that is *right* has to be transformed into *left* on the return path (*direction specific*), while *turn into direction of D* or *turn in the opposite direction of A* remain the same for the return path (*direction unspecific*).

According to the concept of "advanced visibility" (Winter, 2003) it is furthermore important in the egocentric perspective, whether both facades at one location at the intersection are visually identical/similar (e.g., same color, texture) or totally different (e.g., one facade is brown and the other white). This may change the recognizability on the return path in a dramatic way (Figure 3). For instance, if both facades are similar, then this information can be used for the return path, but if they differ significantly, then position D becomes useless on the return path, since it cannot be recognized anymore (only if the observer turns the head on the initial path at the intersection). From a perceptual point of view the object must be recognizable. If this condition is not fulfilled, the former optimal position D might become totally worthless (see Tables 1 and 2 for theoretical predictions; please note the lower right value, which has the most dramatic effects depending on visibility and equal appearance).

Not only the visibility represents an important issue but also language and how it is used when giving instructions, learning new pathways, and transforming them mentally (for the return path). As mentioned above, there are at least two ways of spatial directions: *direction specific* and *direction unspecific* information (Figure 4).

Table 1: Visibilities for the different landmark positions (A-D) in Figures 2 and 3 for the initial path, the theoretical return path, and for the real return path; 0 indicates that no facade is visible, 0.5 indicates that one facade in visible, and 1 means that both possible facades of a building at an intersection are visible. Here, both facades of a single building have the same characteristics/appearance. Thus, position D has a visibility of 0.5 on the return path, since the visible facade is similar to the one seen on the initial path.

Path Position	Initial Path	Return Path (hypothetical)	Return Path (real)
A	1	1	1
B	1	0.5	0.5
C	0.5	1	1
D	0.5	0.5	**0.5**

Table 2: Visibilities for the different landmark positions. In comparison to Table 1 we now assume that the two facades of each building are different in their appearance. This leads to a visibility of 0 for position D, since here the new facade on the return path does not contain any information about this position compared to the initial path.

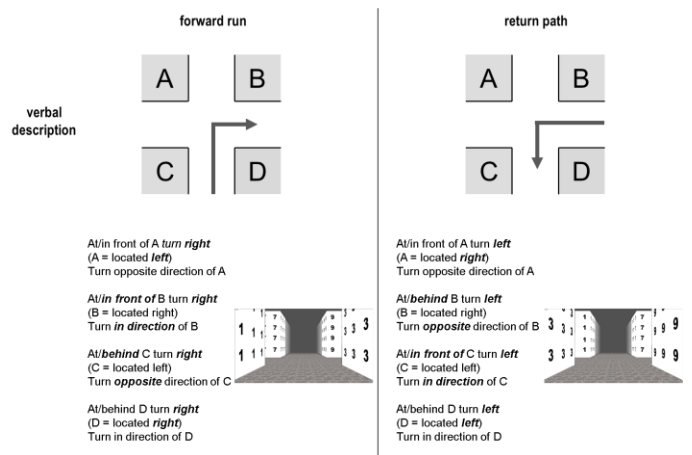| Path Position	Initial Path	Return Path (hypothetical)	Return Path (real)
A	1	1	1
B	1	0.5	0.5
C	0.5	1	0.5
D	0.5	0.5	**0.0**

Figure 4: Examples for verbal directions in the forward run and the return path. Note that the descriptions for positions D and A vary only slightly (if at all), while larger changes occur for positions C and B.

Direction specific here means that a precise direction with a single spatial word is provided, e.g., *left* or *right*. At first glance this information is easy to understand and simple to use. But, it becomes complicated if the return path has to be constructed, since then a *left turn* needs to become a *right turn* and vice versa (note that *straight* remains *straight* on the return path). Thus, an additional mental transformation is required. *Additional* in this sense means that it is also possible to encode directions in an unspecific way (without directions but rather based on landmark locations). In other words, the verbal direction *turn in the direction of the gas station* does not need to be verbally or mentally transformed if it is located on position D (the same is true for position A with the instruction *turn in opposite direction of A*). On the return path, both locations and unspecific directions would remain the same: in the mental representation the gas station would still either be *in direction of the turn* (D) or *opposite to the direction of the turn* (A). This would require one mental processing step less, since no transformation would

be required (left $\rightarrow$ right) resulting in less cognitive load. But, is this how wayfinders encode spatial information and spatial directions? Theoretically, *direction unspecific* information would be less effortful and therefore preferable over a *direction specific* strategy that results in higher cognitive load.

Therefore, it is important in a first step to systematically investigate how wayfinders encode given (unfamiliar) routes and how they transform them into a return path; and in a second step it is necessary to model the optimal strategies (also with respect to individual abilities) to make predictions about spatial performance.

General Discussion and Conclusion

In this position paper we presented first empirical data on return path research and how this information is processed to aid wayfinding (unfortunately, we could not present all empirical research within this study). As can be seen from our theoretical assumptions, much more research is required within this context. We offered a few interesting issues, e.g., structural importance, visibility, language, mental transformation, which should be investigated further. So far we did not focus on brain imaging and neural correlates of wayfinding. But, investigating the cognitive processes of how we learn and encode initial pathways and how we later transform them into new routes (especially return paths) is also of relevance for the neuroscientific branch of this research. Thus, our findings and assumptions about the return path make up an interesting project for interdisciplinary future cognitive research.

Acknowledgement

References

Büchner, S. J., Hölscher, C., & Strube, G. (2007). Path choice heuristics for navigation related to mental representations of a building. In S. Vosniadou, D. Kayser, A. Protopapas (Eds.), *Proceedings of the 2nd European Cognitive Science Conference* (pp. 504-509). Hillsdale, NJ: Lawrence Erlbaum Associates.

Caduff, D., & Timpf, S. (2008). On the assessment of landmark salience for human wayfinding. *Cognitive Processing, 9*(4), 249-267.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58*(1), 7-19.

Dudchenko, P. A. (2010). *Why people get lost – The psychology and neuroscience of spatial cognition.* Oxford: University Press.

Golledge, R. G. (1997). Defining criteria in path selection. In D. F. Ettema & H. J. P. Timmermans (Eds.), *Activity-based approaches to travel analysis* (pp. 151-169). New York: Elsevier.

Hamburger, K., & Knauff, M. (2011). SQUARELAND: A virtual environment for investigating cognitive processes in human wayfinding. *PsychNology Journal, 9*(2), 137-163.

Hamburger, K., & Röser, F. (2011). The meaning of Gestalt for human wayfinding – How much does it cost to switch modalities? *Gestalt Theory, 33*(3/4), 363-382.

Kato, Y., & Takeuchi, Y. (2003). Individual differences in wayfinding strategies. *Journal of Environmental Psychology, 23*(2), 171-188.

Klippel, A., & Winter, S. (2005). Structural salience of landmarks for route discrimination. In A. G. Cohn & D. Mark (Eds.), *Spatial Information Theory. International Conference COSIT* (pp. 347-362). Berlin: Springer.

Lynch, K. (1960). *The image of the city.* Cambridge, MA: MIT Press.

Meilinger, T., Knauff, M., & Bülthoff, H. H. (2008). Working memory in wayfinding- a dual task experiment in a virtual city. *Cognitive Science, 32*(4), 755-770.

Meilinger, T., & Knauff, M. (2008). Ask for your way or use a map: A field experiment on spatial orientation and wayfinding in an urban environment. *Journal of Spatial Science, 53*(2), 13-23.

Papinski, D., Scott, D. M., & Doherty, S. T. (2009). Exploring the route choice decision-making process: A comparison of planned and observed routes obtained using person-based GPS. *Transportation Research Part F, 12*(4), 347-358.

Presson, C. C., & Montello, D. R. (1988). Points of reference in spatial cognition: Stalking the elusive landmark. *British Journal of Developmental Psychology, 6*(4), 378-381.

Röser, F., Hamburger, K., & Knauff, M. (2011). The Giessen virtual environment laboratory: Human wayfinding and landmark salience. *Cognitive Processing, 12*(1), 209-214.

Röser, F., Hamburger, K., Krumnack, A., & Knauff, M. (2012a). The structural salience of landmarks: Results from an on-line study and a virtual environment experiment. *Journal of Spatial Science, 57*(1), 37-50.

Röser, F., Krumnack, A., Hamburger, K., & Knauff, M. (2012b). A four factor model of landmark salience – A new approach. In N. Rußwinkel, U. Drewitz & H van Rijn (Eds.), *Proceedings of the 11th International Conference on Cognitive Modeling (ICCM)* (pp. 82-87). Berlin.

Sorrows, M. E., & Hirtle, S. C. (1999). The nature of landmarks for real and electronic spaces. In C. Freksa & D. M. Mark (Eds). *Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science, International Conference COSIT 1999* (pp. 37-50). Stade: Springer.

Winter, S. (2003). Route adaptive selection of salient features. In W. Kuhn, M.F. Worboys & S. Timpf (Eds.), *Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science, International Conference COSIT* (pp. 37-50). Berlin: Springer.