SPATIAL COGNITION AND COMPUTATION, *vol*(issue), start–end

Interplay between Visual and Spatial: The Effect of Landmark Descriptions on Comprehension of Route/Survey Spatial Descriptions

Paul U. Lee[1] and Barbara Tversky[2]
[1]SJSU / NASA Ames Research Center, [2]Stanford University

Successful wayfinding requires accurate encoding of two types of information: landmarks and the spatial relations between them (e.g. landmark X is left/north of Y). Although both types of information are crucial to wayfinding, behavioral and neurological evidence suggest that they have different substrates. In this paper, we consider the nature of the difference by examining comprehension times of spatial information (i.e. route and survey descriptions) and landmark descriptions. In two studies, participants learned simple environments by reading descriptions from route or survey perspectives, half with a single perspective switch. On half of the switch trials, a landmark description was introduced just prior to the perspective switch. In the first study, landmarks were embellished with descriptions of visual details, while in the second study, landmarks were embellished with descriptions of historic or other factual information. The presence of landmark descriptions did not increase the comprehension time of either route or survey descriptions, suggesting that landmark descriptions are perspective-neutral. Furthermore, visual landmark descriptions speeded comprehension time when the perspective was switched, whereas factual landmark descriptions had no effect on perspective switching costs. Taken together, the findings support separate processes for landmark and spatial information in construction of spatial mental models, and point to the importance of visual details of landmarks in facilitating mental model construction.

Keywords: visual, spatial, route, survey, perspective, landmark, wayfinding, spatial mental models

---

Correspondence concerning this article should be addressed to Paul U. Lee at Human Factors Dept, NASA Ames Research Center, Mail Stop 262-4, Moffett Field, CA 94035; email: plee@mail.arc.nasa.gov. The research was supported by Office of Naval Research Grants NOOO14-PP-1-O649, N000140110717, and N000140210534 to Stanford University

The finding that perception of a simple object activates two different neural pathways – separate processes for object and spatial information – seems puzzling, but is true. Evidence from brain injuries as well as brain imaging suggests functionally separate pathways for object location and object identification. Research on brain-damaged patients reveals this dissociation: some are impaired in object recognition; others in imagery for spatial relations (e. g., Farah and Hammond, 1988). Brain imaging research has determined that a ventral pathway, including regions of the temporal lobe, is involved in object identification, and a dorsal pathway, including regions of the parietal lobe, is involved in object location (e.g. Ungerleider and Haxby 1994). The distinct object/spatial processes along dorsal and ventral pathways extend to the prefrontal cortex, which plays an important role in visual working memory. Both regions provide neural bases for the visuospatial “sketch pad” in the working memory, as originally proposed by Baddeley and Hitch (1974).

Analogous to object vs. spatial processes, neuropsychological and neuroimaging evidence also suggests separate processes for landmark and spatial knowledge in navigation (e. g., Epstein and Kanwisher, 1998; Maguire, Frakowiak, and Frith, 1997). Lesions in the ventral pathway sometimes result in loss of landmark recognition without significant loss of spatial abilities (i.e. landmark agnosia), whereas lesions in the dorsal pathway (e.g. posterior parietal cortex) have been linked to topographical disorientation in which the main deficit is spatial without loss of landmark recognition (de Renzi, Faglioni, & Villa, 1977).

Despite the separate pathways in the brain for processing objects and for processing spatial relations, the world we apprehend seems coherent and integrated. Dissociated object and spatial processes integrate and intertwine seamlessly to allow us to perceive the world as an integrated whole rather than as disjointed pieces of objects and spatial relations. The integration of object and spatial processes is essential to wayfinding. Successful wayfinding requires accurate integration and memory of landmarks/streets in their spatial relations.

A common way to gain knowledge of a new environment is from language. Spontaneous descriptions of environments typically take one of two perspectives, route or survey, or a mixture of the two (Taylor and Tversky, 1996). Route descriptions take an imagined navigator through an environment describing landmarks relative to the navigator in terms of *left* and *right*. Survey descriptions take a bird’s eye view of the environment describing landmarks relative to each other in terms of cardinal directions. Interestingly, mental representations of well-learned environments of a manageable size appear to be perspective-free. That is, retrieval of both kinds of information, route and survey, is equally fast and accurate for environments learned from the matching perspective as for those learned from the other perspective (Taylor and Tversky, 1992). In those studies, after participants learned descriptions of a variety of environments from either a route or survey perspective, they were tested with inference true/false statements from both perspectives. Performance was as fast and accurate on the alternative perspective as on the learned one in four experiments. For verbatim statements, reaction times were faster on same perspective statements presumably because of faster recognition of the exact wording.

Perspective is one of those useful terms that, as a consequence, has many meanings. A partial visual analog of this paradigm has produced results that might seem to contradict previous findings on perspective-free representation of space. Participants learned table-top virtual reality environments modeled on those of Taylor and Tversky (1992). The survey analog was from a bird's eye view of the environment and the route analog was from a perspective embedded in the environment. Recognition tests were scene clips sampled from the videos and foils were scrambled scene clips. Remarkably, participants were faster but not more accurate recognizing scene clips from the viewed perspective than from the other perspective. Given the difference in visual details of the scenes viewed from above or from within, it is surprising that there were no differences in accuracy. These tests were comparable to Taylor and Tversky's (1992) verbatim tests, for which there was an advantage to the read perspective. The evidence for perspective independence in the case of learning environments from descriptions was from the inference statements, where congruent scenes were described differently. Constructing analogs to the inference statements for the virtual reality environments can be done, but not easily. Inference analogs would have to entail the same degree of change in the visual details and spatial relations as the perspective changes. This would be an interesting comparison; given other work on scene recognition, it seems that visual similarity and spatial relations should be more important than visual perspective. Indeed, Shelton and McNamara's (2004) findings are similar to related work showing viewpoint dependence of spatial memory (e.g. Rieser, 1989; Diwadkar and McNamara, 1997).

The analysis of verbal and visual analogs illustrates the complexity of the term *perspective*. In language, it entails a complex set of variables, primarily, reference terms and referent objects. Although there are claims that combinations of these features form coherent reference frames, for example, relative, intrinsic, and absolute (e. g., Levinson, 1996), the evidence from natural speech shows clearly that they are mixed (Taylor and Tversky, 1996). Neither reference terms nor referent objects have visual analogs; they are part of the way that people encode or describe scenes. Conversely, the visual analog reflects a kind of perspective that is hard to capture verbally, a specific view on the rich elements of a scene. The inherent difficulties in constructing visual and verbal analogs are instructive.

In sum, what seem like contradictory claims are not. The findings for virtual reality do not challenge the prior findings that well-learned verbal descriptions of environments result in perspective-free representations. There are no claims here or earlier that representations are always or necessarily perspective-free. Quite the contrary, what is surprising is that perspective-dependent information *can* lead to perspective-free representations. The transformation from perspective-dependent information to perspective-free representations was explored in a series of experiments by Lee and Tversky (2001). Participants learned simple environments from four sentences. The first three sentences were always from the same perspective, route or survey. On half the trials, the perspective of the fourth sentence was switched. When it was switched, reading times increased, suggesting significant perspective switching costs during the construction of spatial mental models. A follow-up study isolated two key variables of route and survey perspectives, namely reference terms (e.g. left/right vs. north/south/east/west) and embedded vs. external viewpoint to determine the locus of perspective switching costs during text comprehension. The results showed that the costs were primarily due to the switch of reference terms that accompanies the perspective change and to a lesser degree due to the change in viewpoints.

Once the environments were learned, however, perspective effects dissipated during retrieval. Surprisingly, switching reference terms seemed to exact little cost during retrieval. While small switching costs remained for some of the studies, they disappeared altogether in others (Lee and Tversky, 2001). A study by Lee (2002) explicitly demonstrated a gradual shift from perspective dependent to perspective-free spatial representation over time. In this study, small switching costs were present when the spatial information was retrieved from a different perspective immediately after learning but the costs dissipated when the information was retrieved from one perspective (i.e. route or survey) after an extensive learning and testing from another perspective.

Considering a broader set of paradigms and a broad set of interpretations, it is undeniable that there are situations in which mental representations appear to be perspective-dependent and others in which they appear to be perspective-free. Perspective-dependence is not surprising – after all, experience is from a perspective. Yet the mind transforms perspective dependent information to representations that are more general, even the mind of a rat (O'Keefe and Nadel, 1978).

In spontaneous route directions, descriptions of landmarks abound, particularly at turning points, and are key to successful wayfinding (e. g., Denis, 1997; Denis, Pazzaglia, Cornoldi, and Bertolo, 1999). Given the prevalence of landmark descriptions in normal route directions, it is likely that people integrate non-locative landmark descriptions with locative route/survey descriptions, and there is evidence that they do (McNamara, Halpin, & Hardy, 1992). Although descriptions of spatial layouts are often interrupted with landmark descriptions, integrating non-locative and locative information may increase cognitive load or decrease comprehensibility of spatial information (Ruddle, Payne, & Jones., 1997; Schneider & Taylor, 1999).

In prior experiments by Lee and Tversky, sentences primarily described spatial relations, mentioning landmarks only by name. Given that there are separate neural pathways subserving spatial and visual information, the spatial descriptions and the visual landmark descriptions would be expected to engage the ventral and dorsal systems respectively. For both route and survey perspectives, describing the details of a landmark disrupts the continuity of the description of the spatial relations, thereby requiring a change of focus from representing spatial relations to representing visual details and potentially increasing the overall cognitive costs. There is some evidence that visual details can impede inferential reasoning when the details are irrelevant to the overall reasoning process (Knauff and Johnson-Laird, 2002). Alternatively, visual details of landmarks may reduce the overall cognitive costs despite the change of focus from spatial relations to visual details because they highlight landmarks that are crucial to the construction of the overall spatial mental model and thereby aid the integration of landmarks/streets with their spatial relations.

To investigate the interplay between visual and spatial information in wayfinding, we examined the on-line construction of spatial mental models from language, specifically, from route or survey perspectives. By comparing the reading times of route, survey, and landmark descriptions of fictitious environments, we can analyze the interactions between spatial information and visual landmark information in the on-line construction of mental models.

Experiment 1: Effect of Visual Landmark Descriptions on Construction of Mental Models from Spatial Descriptions

The first experiment extended a previous study that examined the cost of perspective switching in spatial descriptions during on-line comprehension and subsequent retrieval of spatial information (Lee & Tversky, 2001). In both the previous and the present studies, participants read a sequence of sentences describing an environment from a route or survey perspective, half with an abrupt switch of perspective. During the memory retrieval phase, participants verified true/false statements on the spatial layouts from both perspectives. There were significant costs of perspective-switching during on-line construction of mental models, indicating that construction of mental models is more efficient when perspective is constant. However, those costs diminished or disappeared during retrieval of spatial information from both perspectives, suggesting that perspective-switching encourages establishment of perspective-free mental models   Based on previous findings, the cost of switching perspective on reading times was expected to be significant in constructing new spatial mental representations but diminished during retrieval (Taylor & Tversky, 1992; Lee & Tversky, 2001).

In the current study, participants additionally read landmark descriptions just prior to the perspective switches. Although landmarks can be any distinctive reference point that can be used for navigation, such as buildings, streets, intersections, etc. (Denis, 1997), we used buildings as our primary source of landmarks and will refer to these buildings as landmarks in the following sections. Analogous to naturalistic landmark descriptions, landmark descriptions in this study highlighted salient visual features of landmarks (e.g. “The post office is a red brick building.”); some contained some information that may be construed as spatial (e.g. “It is a tall building with slanted roof on top.”). However, they were different from locative descriptions in that they were not dependent on a particular reference frame. A locative description, such as “A is left/north of B”, requires both the position and the orientation of a reference frame before determining the meaning of “left” or “north.”

To determine if there are costs related to comprehending visual landmark descriptions, the comprehension times for locative and non-locative descriptions were examined with and without non-locative landmark descriptions embedded into locative route/survey descriptions. If landmark descriptions are processed similarly to either route or survey descriptions, then switches from landmark to route/survey descriptions should incur similar switching costs as route-to-survey or survey-to-route switch. Furthermore, if landmark descriptions are consistently construed from either route or survey perspective, cognitive costs should increase when landmark descriptions are embedded into one perspective but not the other. On the other hand, if landmark descriptions are integrated into the overall spatial representation while maintaining a neutral perspective, the nature of the costs would not parallel those of perspective switches. Landmark descriptions may incur switching costs if they interfere with the comprehension of spatial descriptions, but they may not incur switching costs if they can be readily integrated into the overall spatial representation.

Method

Participants. Fifty-two undergraduates, 23 male and 29 female, from Stanford University participated individually in partial fulfillment of a course requirement. Six men and fourteen women were eliminated because their retrieval accuracy of spatial relations and/or landmark descriptions was below the criteria of 67% and 75%, respectively. Data from the remaining thirty-two participants, 17 male and 15 female, were analyzed and reported below.

Materials. Descriptions were written for sixteen fictitious environments, half with a route perspective, half survey. Each environment consisted of two intersecting roads and three adjacent landmarks. Each description began with two non-spatial sentences, followed by four sentences describing the spatial layout of the environment. In the route description, the first route direction contained a cardinal direction so that the environment could be referenced from an extrinsic reference frame, which allowed subsequent retrieval from survey perspective. Figure 1 provides examples of route and survey descriptions, respectively.

Route Description	*Survey Description*
The financial district is in the heart of River City. When you enter the main street of the district, High St, you will see its restored cobblestone surface.	The financial district is in the heart of River City. The main street of the district, High St, has a restored cobblestone surface.
Go east on High St and you will intersect with a much narrower Green Ave.	High St runs east-west, intersecting a much narrower Green Ave, which runs north-south.
Turn right on Green Ave and on your right, you will see the stock market.	South of High St on the west side of Green Ave is the stock market.
Past the stock market, on your right on Green Ave, you will see the mortgage bank.	South of the stock market on the west side of Green Ave is the mortgage bank.
On your right on Green Ave, past the mortgage bank is the legal firm.	On the west side of Green Ave, south of the mortgage bank is the legal firm.

Figure 1. Route and Survey Descriptions

In order to examine the costs of perspective switching during on-line comprehension, the last sentence of the study phase, termed the target sentence, was presented either in the same or a new perspective relative to the preceding descriptions. Figure 2 shows the perspective switches in both directions.

Perspective Switch: Route to Survey	*Perspective Switch: Survey to Route*
Past the stock market, on your right on Green Ave, you will see the mortgage bank.	South of the stock market on the west side of Green Ave is the mortgage bank.
On the west side of Green Ave, south of the mortgage bank is the legal firm.	On your right on Green Ave, past the mortgage bank is the legal firm.

Figure 2. Perspective Switch during On-line Comprehension

For half of the environments, a landmark description was introduced just after the introduction of the second landmark and just prior to the target statement. Landmark descriptions gave visual details of salient features of the landmarks (see Figure 3). The number of propositions and the sentence length of the descriptions were roughly equated to those of the target sentences. The landmark descriptions were designed to evoke vivid visual imagery while minimizing spatial information, although some of the visual information may have evoked some spatial imagery (e.g. “high ceiling”). In the control condition, landmark descriptions were omitted prior to the target statements. Instead, street descriptions with comparable lengths were inserted after the introduction statements to equate the overall description lengths between the two conditions. The street descriptions were introduced at the beginning so that maximum number of spatial descriptions could be introduced uninterrupted prior to the perspective switch. The street descriptions gave visual details of a street introduced in the opening paragraph.

Landmark Descriptions	*Street Descriptions*
The Picasso gallery is a sparse room that has a high ceiling with skylights.	The Central Corridor looks warm and inviting when plenty of light shines through the skylights.
The mortgage bank is built in Santa Fe style that uses lots of terra cotta clay.	The cobblestones on High St are large smooth stones that are arranged in neat orderly fashion.

Figure 3. Examples of Landmark and Street Descriptions

Five true/false verification statements – i.e. two statements each for route and survey perspective and one landmark/street description – followed the comprehension task (see Figure 4 and 5).

Route T/F Statements	*Survey T/F Statements*
The stock market is on your left as you walk from the legal firm towards High St on Green Ave.	The stock market is on the southwest corner of Green Ave and High St.
The legal firm is on your right when you walk past the stock market on your right.	The mortgage bank is on the west side of Green Ave and north of the legal firm.
The mortgage bank is on your right when you face the legal firm from Green Ave.	The mortgage bank is south of the stock market and north of the legal firm.

Figure 4. Route and Survey T/F Verification Statements

Route and survey verification statements provided assurance that participants formed accurate mental models of the spatial layouts. The statements were presented in random order. Half of the statements were from each perspective, so that half will require perspective switch with respect to the study perspective. All of the statements were inference statements, querying the spatial relations that could be inferred but were not directly specified in the descriptions (see Figure 4). Landmark/street verification statements were also presented to assure that they processed the landmark/street descriptions presented during the comprehension phase (see Figure 5). There was one landmark/street verification statement per environment.

Landmark T/F Statements	*Street T/F Statements*
The Picasso gallery has skylights that illuminate a sparse room.	The Central Corridor receives plenty of light that makes it look inviting.
The mortgage bank is an Italian style building that is made of terra cotta clay.	The cobblestones on High St are smooth and organized in a haphazard fashion.

Figure 5. Landmark and Street T/F Verification Statements

Design and Procedure: After a couple of practice trials, each participant read sixteen text descriptions, eight route and eight survey. The texts appeared on the screen one sentence at a time. Each sentence remained on the screen until the participant pressed a key to initiate the next sentence. The reading time for each sentence was recorded. In the landmark description condition, a landmark description was introduced after the introduction of second landmark and prior to the target statement. In the control condition, a street description was inserted after the opening paragraph. After reading a description, each participant answered five true/false statements by pressing assigned keys for true and false. Both response time and accuracy were recorded for each statement. After completing the experiment, participants were cued with a short name for each environment (“modern museum” or “financial district”) and asked to recall as many landmarks and streets as possible. They were given five minutes to complete this task. Figure 6 illustrates the overall design for both control and landmark description conditions.

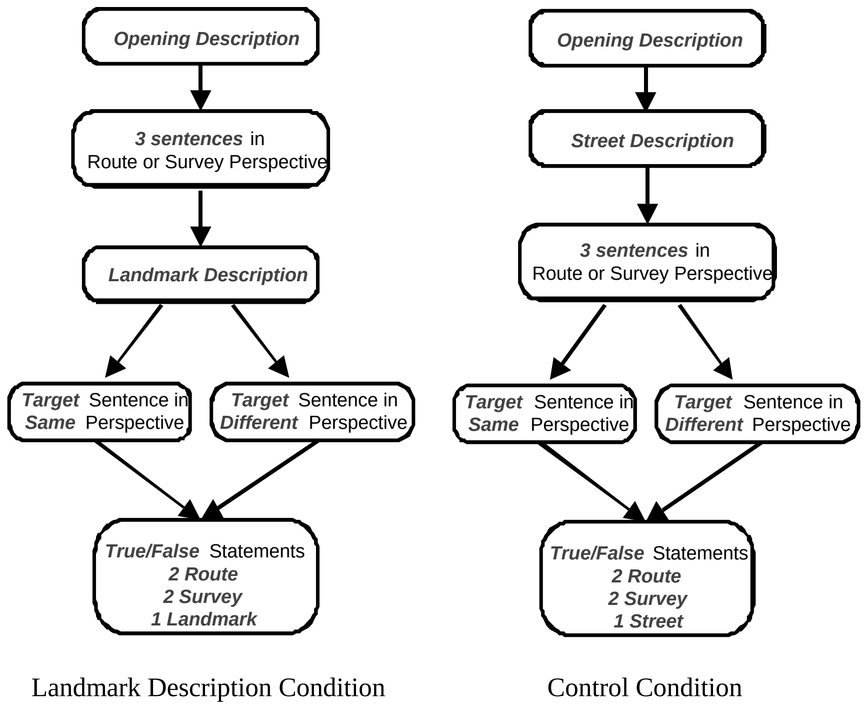

Figure 6. Overall Design of the Visual Landmark Description Experiment: Landmark Description Condition (left) and Control Condition (right)

Results

Repeated measures design was used to analyze reading times (RT) per syllable for spatial descriptions, target sentences, and true/false verification statements. The reading times replicated the previous findings (Lee & Tversky, 2001).

Prior to the target sentences, participants read spatial descriptions from route or survey perspective. For these sentences, they studied survey texts marginally longer (426 msec/syllable) than route texts (404 msec/syllable). $F(1,31) = 2.62$, $p < 0.12$. Following the study texts in route or survey perspective, target sentences were introduced. Half the target sentences switched perspective from the preceding study sentences and the other half kept the same perspective. Target sentences were analyzed for two factors: perspective of the target sentence and perspective consistency to the preceding study texts. An interaction between target perspective and perspective consistency was replicated (see Figure 7). $F(1,31) = 6.07$, $p < 0.02$.

Figure 7 shows that target sentences consistent with the previous perspective were read faster both for route targets (consistent, 317 msec/syllable; inconsistent, 575 msec/syllable) and for survey targets (consistent, 427 msec/syllable; inconsistent, 572 msec/syllable), and that the consistency advantage was greater for route than for survey targets. The three-way interaction among target perspective, perspective consistency, and presence of a landmark description prior to the target sentences was not significant. F (1,31) = 1.03, p > 0.3.

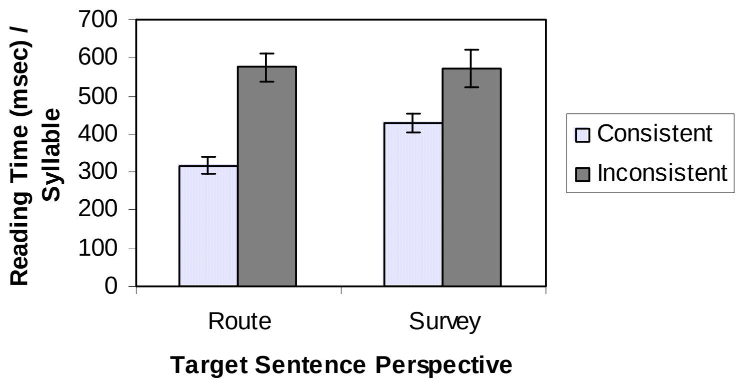

Figure 7: RT/syllable for Target Sentences: Switching Costs during Comprehension

The interaction between target and study perspective indicates that perspective consistency was more important for route descriptions. Since a route perspective is relative to a moving referent (i.e. navigator), knowledge of current location and orientation of the referent is necessary to process the incoming route information. Switching from survey to route descriptions forces the reader to establish the current orientation of the referent without the benefit of inferring this knowledge from the previous text. Since survey descriptions are based on a fixed orientation perspective, consistency has no advantage over switched perspective in establishing the orientation information.

Reading times for the true/false verification statements also replicated the previous findings (Lee & Tversky, 2001). There was a significant interaction between statement perspective and its consistency to study perspective. F(1,31) = 5.60, p < 0.024. Participants responded equally fast to route statements, irrespective of perspective (386 msec/syllable for route vs. 399 msec/syllable for survey) but responded faster to survey statements when they had studied from survey perspective (423 msec/syllable for survey; 526 msec/syllable for route). As in previous studies, there were no significant interactions between statement and target perspective.

Asymmetry in switching costs during retrieval may be due to the fact that true/false statements from a survey perspective maintained the same bird's eye viewpoint as the survey texts during learning. In contrast, each true/false statement in route perspective required the reader to establish a new location/orientation of the referent before verifying the statement, regardless of the study perspective. In essence, each route verification statement required the reader to establish a new viewpoint. These evidences suggest that switching costs during retrieval were driven mainly by viewpoint changes. The switching costs during retrieval were small, however, and disappeared altogether with extensive learning (Taylor and Tversky, 1992), extensive retrieval (Lee, 2002), or greater exposure to multiple viewpoints (Lee and Tversky, 2001).

The accuracy data showed the same significant interaction. $F(1,31) = 9.60$, $p < 0.004$. Participants answered perspective-consistent statements more accurately than perspective-inconsistent statements, but switching costs were greater for survey statements (71% accuracy for inconsistent perspective vs. 89% for consistent perspective) than for route statements (81% accuracy for inconsistent vs. 87% for consistent). This finding is consistent with the RT data so there was no speed-accuracy tradeoff.

Although replication of results is worthwhile, the main interest of this study was to determine how non-locative landmark descriptions affect comprehension of locative route/survey descriptions. Since landmark descriptions were added just prior to the target statements, we can ask if reading times of the target statements were affected by switching from spatial to visual information. Target reading times showed a significant interaction between the presence of a landmark description and perspective consistency to the study perspective (see Figure 8). $F(1,31) = 6.07$, $p < 0.02$.

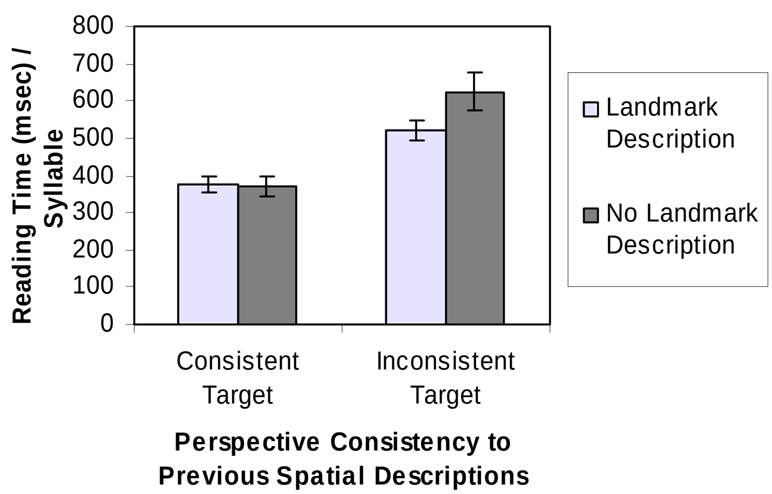

Figure 8. Effects of Landmark Descriptions on Perspective Switch

Figure 8 shows that introducing a landmark description during the route/survey description did not affect target reading times for consistent targets

(375 msec/syllable for landmark descriptions; 370 msec/syllable no landmark description; $t(31) = 0.15$, $p > 0.4$). However, introducing a landmark description just prior to a perspective switch *decreased* reading times for the switched target sentence (with preceding landmark description, 572 msec/syllable; without landmark descriptions, 625 msec/syllable; $t(31) = 3.33$, $p < 0.002$).

Thus, non-locative landmark descriptions had no negative impact on the comprehension of route/survey descriptions; in fact, their presence reduced the cognitive costs of perspective switches. One possible explanation for the reduced reading time of the target sentence may be that the participants simply ignored the landmark descriptions prior to the target sentence and therefore had extra time to process the spatial information. True/false verification statements of the landmark descriptions, however, showed that the participants attended to the descriptions. They accurately identified 88% and 95% of true and false description statements, respectively ($F(1,31) = 7.15$, $p < 0.012$; landmark and street description results were near identical and therefore were collapsed in the analyses). The accuracy data suggest that false statements were easier to verify than true statements, and the RT data corroborated that finding (301 msec/syllable for true statements; 228 msec/syllable for false statements). $F(1,31) = 44.19$, $p < 0.001$.

Landmark descriptions also increased recall of the described landmark from 42% without a description to 52% with a description ($\chi_1^2 = 4.90$, $p < 0.027$), suggesting that elaborations provided by landmark descriptions helped participants to better remember the described landmarks. The landmark descriptions did not affect the recall of the other landmarks. Finally, landmark descriptions were read equally fast when the descriptions followed either route (191 ms/syl) or survey descriptions (198 ms/syl). $F(1,31) = 1.00$, $p > 0.3$. Had the landmark descriptions been construed as either route or survey perspective, then they should have been understood faster after one of the two perspectives. The lack of difference in reading time indicates that there were no perspective biases for the landmark descriptions.

Landmark descriptions were read faster than street descriptions (landmark descriptions, 195 msec/syllable; street descriptions, 231 msec/syllable). $F(1,31) = 8.65$, $p < 0.006$. The recall of the described street was also significantly lower (23%) than the landmark recall. Although the lower recall may suggest that streets in general are harder to remember than buildings, the results were confounded by the fact that the street names in the experiment were often more vague (e.g. central corridor, main aisle) or context independent (e.g. High St., Green Ave) than the landmark names. Finally, the pattern of recall was similar for streets and landmarks, as streets that were described were recalled marginally better (36%) than those that were not described (24%; $\chi_1^2 = 3.57$, $p < 0.06$), again suggesting that the descriptions helped form a better memory trace for the described landmarks/streets.

Not surprisingly, landmark descriptions were comprehended faster than spatial descriptions (landmark descriptions, 195 msec/syllable; target descriptions, 415 msec/syllable; $F(1,31) = 237.54$, $p < 1e-7$). Since these descriptions were matched in their sentence lengths and the number of propositions, the results suggest that non-locative descriptions are much easier to process than locative descriptions. This is consistent with other studies showing reluctance to use projective spatial terms when simpler terms are appropriate (Mainwaring, Tversky, Ohg ishi, and Schiano, 2003).

Discussion

Participants read descriptions of simple environments from route or survey perspectives, knowing they would be asked questions about the spatial layouts of the environments. On half the trials, the last sentence of the description switched perspective. Participants were slower to read switched target sentences, indicating that their on-line construction of mental models of the environments was sensitive to perspective. Sensitivity to perspective diminished in subsequent testing of spatial relations of the environment from both perspectives, presumably because mental representations shifted from perspective-dependent to perspective-free during information retrieval from both perspectives. These findings replicate previous ones (Tversky and Lee, 2001).

New to the present study was the introduction of landmark descriptions just prior to the target on half the trials. Understanding the landmark descriptions required encoding visual as opposed to spatial information. Adding landmark descriptions did not increase the overall cognitive load during the acquisition of spatial mental representations. Intriguingly, reading a landmark description decreased the cost of switching perspective on subsequent target reading times. The landmark descriptions had no effect on target reading times when the perspective was not switched. These findings have a number of interesting implications.

First, it should be noted that there were no cognitive costs associated with intervening visual information during the construction of the spatial mental model, contrary to some previous findings (e.g. Knauff and Johnson-Laird, 2002). The difference in the findings may be due to the fact that for learning environments, the visual details were relevant, so they aided the on-line comprehension of spatial information by providing a better visual representation of the pertinent landmarks in the spatial environment. In contrast, the visual details in the studies of Knauff and Johnson-Laird (2002) were irrelevant to the reasoning process.

Secondly, comprehension times supported the claim that landmark descriptions were not construed as either survey or route since such construal would have resulted in increased cognitive costs when the target sentences were given in the opposite perspective. In fact, landmark descriptions were comprehended equally fast when preceded by either route or survey perspective, suggesting that these landmark descriptions were perspective-neutral.

An alternative explanation may be that landmark descriptions invoked perspective-dependent representations, but the perspective taken varied depending on the preceding spatial perspective. For example, a description "The Picasso gallery is a sparse room that has a high ceiling with skylights." may invoke an image of a three dimensional room viewed from the side for a route perspective while it may invoke a bird's eye view from the ceiling for a survey perspective. If landmark descriptions can malleably take either route or survey perspective depending on the perspective of the preceding descriptions, they can be fully integrated into the overall spatial representation without perspective switching costs. However, this explanation is not consistent with the reduced perspective switching costs after landmark descriptions. If landmark descriptions invoked the perspective of the preceding sentences, a perspective switch after a landmark description should have had the same cost as a switch without the description. Instead, the reduction of switching costs after landmark descriptions demonstrates that the descriptions facilitated perspective switches in both directions.

The facilitation effect of landmark descriptions is a counterintuitive and surprising finding. Taken together with the finding that landmark descriptions had no effect for consistent target sentences, it suggests that not only are visual landmark descriptions perspective-neutral, switching focus to visual details during on-line comprehension reduces perspective dependence of the spatial mental model.

Another possible explanation of the facilitation effect may be the additional time between the perspective descriptions and the target sentence. Specifically, reading landmark descriptions may have given participants more time to process the overall spatial information. This hypothesis was not supported by the data. First evidence is the lack of facilitation effects due to landmark descriptions when perspective was not switched, whereas the temporal interval explanation would have predicted faster comprehension time of the target sentence after reading the landmark descriptions. Next, in the second experiment, described in the following section, there was no facilitation after reading factual descriptions of the landmarks despite the increased time before the target sentence.

Finally, landmarks that were described were recalled better than the other landmarks, demonstrating the advantage of landmark descriptions for long-term memory. Since recall can be enhanced through descriptions without any immediate cognitive costs during comprehension, landmark descriptions provide an effective way to acquire and maintain route knowledge.

Although it is unclear how visual landmark descriptions encouraged the construction of perspective-free spatial mental models, we can speculate on the mechanisms based on the previous findings. One important element in perspective switching costs during comprehension is the cost associated with switching reference frame/terms (Lee and Tversky, 2001). In contrast, reference frame/terms seem to exact very little or no cost once the spatial mental model is constructed. Interestingly, a key difference between locative route and survey descriptions and non-locative landmark descriptions is that landmark descriptions do not invoke a particular reference frame. At the same time, because visual landmark information is central to mental representations of environments, it is integrated into the spatial mental model. It could be that integrating information that is visual but not linked to a particular reference frame encourages establishment of perspective-free mental models. If this is the case, then inserting information about landmarks that is not likely to be integrated into a spatial mental model should not facilitate perspective switching. This possibility is examined in the next experiment, where the landmark descriptions are factual rather than visual.

Experiment 2: Effects of Factual Landmark Descriptions on Construction of Mental Models from Spatial Descriptions

Landmark descriptions appeared to be seamlessly integrated into the general spatial representation and reduced costs for perspective switches during on-line comprehension. Furthermore, recall of described landmarks was better than the remaining landmarks, demonstrating more benefits of landmark descriptions.

These results may seem at odds with other research suggesting that integration of locative and non-locative descriptions increases cognitive load and decreases comprehensibility. For example, Schneider and Taylor (1999) gave participants route and survey descriptions with non-spatial filler descriptions, such as “Mount Lebanon Park is the site of the annual Octoberfest Celebration.” and “Last year there was construction on Highway 19.” Participants took notes while listening to the descriptions and later rated them for understandability and clarity. When participants recorded a greater number of non-spatial descriptions, their ratings of description understandability and clarity were correspondingly lower, regardless of the number of spatial descriptions recorded.

Note, however, that Schneider and Taylor’s landmark descriptions did not provide visual information. It is possible that benefits of landmark descriptions are restricted to descriptions of visual information, information that is relevant to understanding of the visuospatial nature of the environment. This possibility is addressed in the next study – a near replication of the first with factual landmark descriptions replacing the visual ones used in the first experiment.

Method

Participants. Thirty-six undergraduates, 12 male and 24 female, from Stanford University participated individually in partial fulfillment of a course requirement. Two men and one woman were eliminated from the analyses using the accuracy criteria of the previous experiment. Data from the remaining thirty participants, 10 male and 20 female, were analyzed and are reported below.

Materials and Procedure. The stimuli were identical to those used in the previous experiment except for the landmark descriptions. Instead of describing physical details of the landmarks, the descriptions gave other pertinent information, such as historical relevance (see Figure 9). The descriptions were designed to evoke as little visual imagery as possible, but some descriptions were inevitably visual. The reported quality of images of the two types of descriptions differed, however, since visual descriptions in the first study seemed to evoke images of the landmarks themselves whereas "factual" descriptions in this study evoked images about events related to the landmarks.

Landmark Descriptions	*Street Descriptions*
The Picasso gallery is the most popular, drawing millions of visitors every year.	The Central Corridor is the main access to the site and draws the most traffic.
The mortgage bank has the highest interest rate in town but it still has many customers.	The cobblestones on High St were brought from a quarry hundreds of miles away.

Figure 9. Examples of Factual Landmark Descriptions

The overall procedure was identical to the previous experiment. Participants read descriptions of sixteen environments with and without factual landmark descriptions and then answered four true/false questions for each description in both route and survey perspectives, as well as a true/false question related to the landmark or street description (see Figure 10).

Landmark T/F Statements	*Street T/F Statements*
Millions of visitors come to the museum every year to see the Picasso gallery.	The Central Corridor is quite busy as it provides the main access to the museum.
The mortgage bank has the highest interest rate in town so there are not many customers.	The cobblestones on High St were brought from a nearby quarry.

Figure 10. Examples of Factual Landmark T/F Verification Statements

Results

Replicating previous findings, participants studied the initial survey texts longer (458 msec/syllable) than route texts (407 msec/syllable). $F(1, 29) = 12.55$, $p < 0.001$. There was also an interaction between target perspective and perspective consistency with the preceding study texts. Switching from a route to a survey perspective exacted a lower cost (538 msec/syllable for the switched perspective; 454 msec/syllable for the consistent perspective) than switching from survey to route (617 msec/syllable for the switched; 324 msec/syllable for the consistent, $F(1,29) = 20.15$, $p < 0.0002$).

For the true/false verification statements, the response times did not replicate the statistical significance of previous experiments, although the data trended in the same direction as those in the previous experiments. Route statements were verified marginally faster than the survey statements (379 msec/syllable for route, 409 msec/syllable for survey; $F(1,29) = 2.74$, $p < 0.11$), but there was no main effect of perspective consistency with the study perspective (387 msec/syllable for consistent perspective, 401 msec/syllable for inconsistent; $F(1,29) = 1.11$, $p > 0.3$). The interaction between statement perspective and study perspective was also not significant. $F(1,29) = 1.35$, $p > 0.25$.

Unlike previous experiments (e.g. Lee and Tversky, 2001), there was a significant interaction between statement perspective and study perspective in the accuracy data. $F(1,29) = 9.60$, $p < 0.003$. Survey true/false statements were answered more accurately after survey learning (86%) than route learning (75%), but route true/false statements were answered with similar accuracy after route learning (89%) and survey learning (86%). The shift in the significant interactions from RT to accuracy data may reflect a shift in participants' strategy from an emphasis for speed in the first study to an emphasis for accuracy in this study.

The purpose of this study was to examine whether factual landmark descriptions yielded similar switching benefits as the visual landmark descriptions. The reading times for the target sentences resulted in a significant interaction between the presence of landmark descriptions and perspective consistency but the interaction was the reverse of that of the previous experiment (see Figure 11). $F(1,29) = 4.25$, $p < 0.05$. The presence of a factual landmark description *increased* reading times for an inconsistent target (551 msec/syllable without a description; 605 msec/syllable after a description) but decreased reading times to a consistent target (407 msec/syllable without a description; 371 msec/syllable after a description). However, planned contrasts on target sentences failed to show significant difference for either the consistent ($t(29) = 0.83$, $p > 0.2$) or the inconsistent targets ($t(29) = 1.24$, $p > 0.1$).

Therefore, factual landmark descriptions did not reduce cognitive costs during perspective switches, unlike visual landmark descriptions. Furthermore, they failed to improve their recall of the described landmarks. Participants recalled 57% of the landmarks with descriptions vs. 55% of the landmarks without descriptions. $\chi_1^2 = 0.21$, $p > 0.6$. Similarly, streets were recalled equally well when they were described (30%) as those that were not described (30%). $\chi_1^2 = 0.04$, $p > 0.8$.

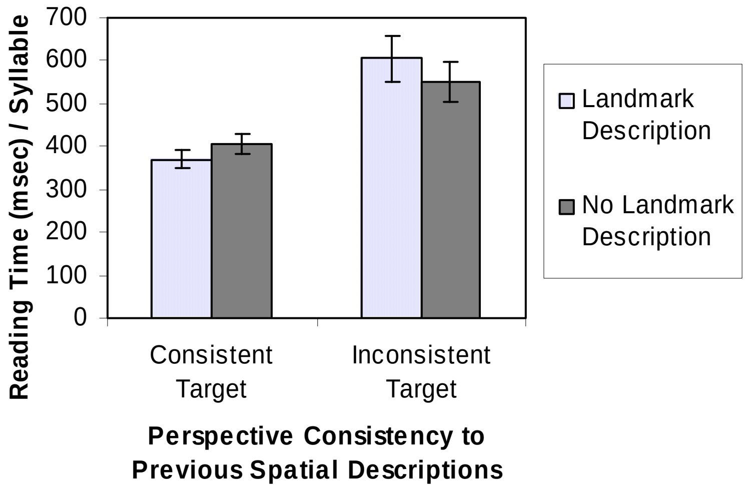

Figure 11: RT/syllable for Target Statements:
Effects of Factual Landmark Descriptions on Perspective Switch

As in the first study, true/false verification statements of the landmark descriptions showed that the participants attended to the descriptions. Collapsing similar results for landmark and street descriptions, participants accurately identified 94% and 92% of true and false description statements, respectively. $F(1,29) = 0.75$, $p > 0.3$. Similar to the first study, the RT data suggested that false statements were easier to verify than true statements (318 msec/syllable for true statements; 225 msec/syllable for false statements). $F(1,29) = 27.51$, $p < 0.001$.

Otherwise, the findings on landmark and street descriptions replicated the first study. First, the preceding spatial perspective had no significant effect on the comprehension time of landmark descriptions (175 msec/syllable for route; 199 msec/syllable for survey). $F(1,29) = 2.22$, $p > 0.14$. Secondly, landmark descriptions were read faster than street descriptions (landmark descriptions, 187 msec/syllable; street descriptions, 209 msec/syllable). $F(1,29) = 9.69$, $p < 0.004$. Finally, the recall of the described street was also significantly lower (30%) than the landmark recall. Participants comprehended factual landmark descriptions much faster (187 msec/syllable) than target spatial locative descriptions (432 msec/syllable; $F(1,29) = 240.40$, $p < 1e\text{-}7$), suggesting again that locative descriptions are harder to process than non-locative ones. Comprehending factual landmark descriptions took about the same amount of time as comprehending visual landmark descriptions in the first study.

Discussion

In the first study, visual landmark descriptions facilitated switching perspectives. They were also integrated with route/survey descriptions without adding cognitive costs and had long-term benefits of better recall. Such benefits seem to be limited to visible features of the landmarks since factual descriptions in this study failed to show any facilitation for perspective switching or better subsequent recall. Since participants accurately remembered the factual information, the absence of facilitation suggests that factual information about landmarks, in contrast to spatial information, does not get integrated into spatial mental models of environments.

The results suggest that spatial and visual information, both of which are in visuospatial domain, have tighter interrelations than spatial and factual information. They also suggest that only a switch between spatial and visual information encourages perspective-free representations, perhaps because visual but not factual information are relevant to the formation of mental models.

Although factual descriptions did not provide any benefits, they also did not significantly increase the overall cognitive costs as suggested by previous research (Ruddle, Payne, & Jones., 1997; Schneider & Taylor, 1999). However, that research incorporated a much larger proportion of factual description. It is possible that a greater number of factual descriptions may be needed before they show added costs.

Conclusions

Environments are typically described from one of two perspectives, route or survey, or a combination of both. In a route perspective, the viewpoint is that of a traveler within the environment; in a survey perspective, the viewpoint is that from overhead. Spatial descriptions can induce mental models of the environments they describe, models that preserve the elements and the spatial relations among them. The present experiments replicated earlier studies (Lee and Tversky, 2001) in demonstrating comprehension time costs in constructing on-line mental models when spatial descriptions change perspective. With repeated perspective switches during information retrieval, those costs diminish and disappear. Repeatedly retrieving information from different perspectives apparently leads to perspective-free mental representations. Descriptions of environments, especially route directions, are often enriched with visual information in addition to the spatial, in particular, visual descriptions of landmarks. This visual information aids travelers to identify landmarks in the situation. Visual and spatial information are carried in different tracks in the brain; might that affect integration of these in comprehending spatial descriptions?

The present experiments examined the relations between visual and spatial information in spatial mental models by adding descriptions of landmarks, common in spontaneous spatial descriptions, to step-by-step locative descriptions. When landmark descriptions preceded perspective switches and contained visual information, the costs of switching perspective decreased. This reduction in perspective-switching costs may occur because incorporating visual information about landmarks into a mental model of the environment entails switching from encoding spatial locative information to encoding visual information. This switch, because it is in the visuospatial domain, may also encourage formation of perspective-free representations. Only visual descriptions of landmarks facilitate forming perspective-free representations. The reduction of switching costs was not obtained when the landmark descriptions were not visual, when they contained historical or social facts rather than information that could be imagined. Because that information is not visuospatial, it is not integrated into the visuospatial mental model, and it does not serve to reduce perspective bias in mental models. It appears, then, that spatial and visual information is integrated in spatial mental models.

The reduction in costs of switching perspectives obtained from visually described landmarks is comparable to the reduction in switching costs that comes from retrieving information from different perspectives. However, unlike switching spatial perspective, switching to visual information does not in and of itself increase comprehension time. This may be because comprehending visual information is easier than comprehending spatial information requiring understanding spatial relations in a reference frame. Surprisingly, visual descriptions of landmarks have benefits without costs. Their benefits extend from formation of mental models to memory for the landmarks.

The findings suggest a higher level of interconnectivity between visual landmark and spatial information than between factual and visual information. Recall of landmarks was better when the physical attributes of the landmarks were described but not when the descriptions were factual. In addition, the presence of a landmark description reduced the cognitive costs of switching perspectives when landmark descriptions were visual. In contrast, factual descriptions seem to increase the cognitive costs of switching perspectives. Since landmarks are critical to route knowledge, it is clearly desirable if landmark memory can be improved without increasing cognitive costs during comprehension. The data suggest that only the visual landmark descriptions can do both.

Descriptions of environments typically contain both spatial and visual information. When viewed rather than described, these kinds of information travel on different pathways in the brain. Nevertheless, in descriptions of space, as in actual navigation of space, they appear to be seamlessly integrated. In order to successfully navigate an environment, people need to know the structure of the paths on which they can navigate; they also need to know how to recognize landmarks when they see them. Spatial mental representations are used for judgment as well as navigation, such as judgments of distance or direction. Many of these judgments depend on survey rather than route perspectives. It follows from this analysis and these findings that to be useful for navigation and judgment, mental representations of environments should inherently incorporate more than one perspective and visual as well as spatial information. Given that spatial mental models are used for retrieving visual information and spatial information from several perspectives, it makes sense that retrieval tasks requiring different perspectives and visual information encourage perspective-free representations.

References

Aguirre, G. K. and D'Esposito, M. (1997). Environmental knowledge Is subserved by separable dorsal/ventral neural areas. *Journal of Neuroscience*., 17, 2512-2518.

Baddeley, A. D. & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.). *The Psychology of Learning and Motivation, 8,* 47-89.

Denis, M. (1997). The descriptions of routes: A cognitive approach to the production of spatial discourse. *Current Psychology of Cognition, 16,* 409-458.

Denis, M., Pazzaglia, F., Cornoldi, C., & Bertolo, L. (1999). Spatial discourse and navigation: An analysis of route directions in the city of Venice. *Applied Cognitive Psychology, 13,* 145-174.

de Renzi, E., Faglioni, P., & Villa, P. (1977). Topographical amnesia. *Journal of Neurology, Neurosurgery & Psychiatry, 40(5),* 498-505.

Diwadkar, V. A., & McNamara, T. P. (1997). Viewpoint dependence in scene recognition. *Psychological Science, 8(4),* 302-307.

Epstein, R., and Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, 392, 598-601.

Farah, M. J. and Hammond, K. M. (1988). Visual and spatial mental imagery: Dissociable systems of representation. *Cognitive Psychology*, 20, 439-462.

Knauff, M., Jola, C., & Strube, G. (2001). Spatial Reasoning: No Need for Visual Information. In D. R. Montello, (Ed.). *Spatial Information Theory. Foundations of Geographic Information Science.* pp. 447-457. Heidelberg: Springer-Verlag.

Knauff, M. and Johnson-Laird, P. N. (2002). Visual imagery can impede reasoning. *Memory & Cognition, 30(3),* 363-371.

Lee, P. U. (2002). Costs of Switching Perspectives in Route and Survey Description, *Dissertation Abstracts International, 63(4-B),* 2084B. (UMI No. 3048564).

Lee, P. U. and Tversky, B. (2001). Costs of Switching Perspectives in Route and Survey Description. *Proceedings of the twenty-third Annual Conference of the Cognitive Science Society, Edinburgh, Scotland*

Levinson, S. (1996). Frames of reference and Molyneux's question: Cross-linguistic evidence. In P. Bloom, M. A. Peterson, L. Nadel, & M. Garrett (Eds.), *Space and language* (pp. 109-169). Cambridge: MIT Press.

McNamara, T. P., Halpin, J. A., & Hardy, J. K. (1992). The representation and integration in memory of spatial and nonspatial information. *Memory & Cognition, 20*, 519-532.

Maguire, E. A., Frackowiak, R. S. J., & Frith, C. D. (1997). Recalling routes around London: Activation of the right hippocampus in taxi drivers. Journal of Neuroscience, 17, 7103-7110.

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 887-897.

Ruddle, R. A., Payne, S. J., & Jones, D. M. (1997). Navigating buildings in `desk top' virtual environments: Experimental investigations using extended navigational experience. *Journal of Experimental Psychology: Applied, 3,* 143-159.

Schneider, L. F. and Taylor, H. A. (1999). How do you get there from here? Mental representations of route descriptions. *Applied Cognitive Psychology, 13,* 415-441.

Shelton, A. L. and McNamara, T. P. (2004). Orientation and perspective dependence in route and survey learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 158-170.

Taylor, H. A., and Tversky, B. (1992). Spatial mental models derived from survey and route descriptions. *Journal of Memory & Language, 31(2),* 261-292.

Taylor, H. A., and Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory and Language, 35,* 371-391.

Tom, A. and Denis, M. (In press). Language and spatial cognition: Comparing the roles of landmarks and street names in route instructions. *Applied Cognitive Psychology.*

Mainwaring, S. D. Tversky, B., Ohgishi, M. & Schiano, D. J. (2003). Descriptions of simple spatial scenes in English and Japanese. *Spatial Cognition and Computation*, 3, 3-42.

Ungerleider, L. G. and Haxby, J. V. (1994). "What" and "where" in the human brain. *Current Opinions in Neurobiology, 4,* 157-165.