CHAPTER

26

Spatial Processing of Environmental Representations

James R. Brockmole and Ranxiao Frances Wang

ABSTRACT

In everyday life, we accomplish tasks that require the storage and access of mental representations of many familiar locations. Humans store this environmental knowledge in a series of representations in memory. We discuss recent studies on how humans attend to and process multiple environmental representations to reason and act within the spatial world. These studies suggest that people access one environment at a time, and they automatically update their relationship to their immediate environment, but not to remote environments. Navigation across environments involves reorienting to upcoming environments at certain spatial regions and dropping the old environment from the processing. This selective processing sheds light on the nature and structure of human environmental representations.

I. INTRODUCTION

Throughout this book, attention and selective processing have been shown to play an important role in various visually based cognitive tasks, such as visual search, change detection, object identification, scene processing, and learning. In studies on the nature and structure of spatial representations in humans, however, the role of attention and selective processing has received little consideration. As a result, traditional models of environmental representations in humans largely ignore issues of information processing load, a critical oversight given the vastness of our world. In this chapter, we discuss the limitations of human spatial processing ability, show how traditional models of environmental knowledge encounter difficulty when attention and processing limitations are considered, and describe an alternative view of spatial processing that is consistent with these limitations.

II. THE HIERARCHICAL MODEL OF ENVIRONMENTAL REPRESENTATIONS

Humans live in amazingly complicated environments. On a daily basis, we encounter environments that span miles and include natural and artificial structures, as well as subjectively imposed divisions among areas such as neighborhoods, cities, and states. Given this complexity, humans cannot directly perceive the entire extent of their world from a single vantage point, yet exhibit an impressive ability to reason about, and navigate within, the spatial world. We can give directions to our house to a friend or describe the layout of our kitchen while sitting in our living room. We can turn from one direction to face another without losing track of our bearings with respect to regions of space beyond our field of view. We can navigate from one place to another without being able to see our desired destination.

To perform these tasks, humans must store, recall, and transform information about the world in memory. This information is referred to as an environmental representation. Environmental representations must encode environments at different scales. For example, people have spatial knowledge of small places such as the arrangement of objects on their desk or the location of furniture in a room as well as large places such as the arrangement of rooms in a building, or of buildings in a city. According to the hierarchical network model, the representation of an environment is composed of a number of distinct units, which encode information at different levels of detail.

Moreover, these units are organized into a systematic, hierarchical structure. For example, humans maintain separate representations of a room, the building the room is in, the block the building is on, and so forth. Within this network, the spatial relationships among locations in the world are explicitly encoded in memory only if the locations are encoded within the same representation (e.g., McNamara, 1986; Stevens and Coupe, 1978). When making spatial judgments about locations encoded in different representations, information from each representation and the corresponding superordinate representations must be accessed and combined to compute a solution (Hirtle and Jonides, 1985; McNamara 1986; McNamara et al., 1989; Stevens and Coupe, 1978; Taylor and Tversky, 1992; Wilton, 1979).

The hierarchical network model is intuitively appealing and explains several phenomena well. For example, when people judge the directional relationship between items encoded in separate units, they typically bias their judgments in accordance to the relationship shared by superordinate units. For example, Reno, Nevada, is overwhelmingly judged to be northeast of San Diego, California, although Reno is actually northwest. According to the hierarchical model, the relationship shared by Reno and San Diego is not encoded in memory but must be computed by combining the subordinate knowledge that Nevada is east of California and the subordinate knowledge that San Diego is in Southern California and Reno is in northern Nevada, which results in the biased judgment (Stevens and Coupe, 1978). In addition, within-unit distances are typically underestimated while across-unit distances are often overestimated (Hirtle and Jonides, 1985; McNamara, 1986), and spatial judgments about targets in a single representation are made faster than across-representation judgments regardless of the distance between the locations (McNamara, 1986; Wilton, 1979). These findings suggest that locations within a single unit are "mentally closer" than those in different units.

There is one aspect of the hierarchical network model that has traditionally been overlooked, however. To infer spatial relationships based on two pieces of information, they need to be available at the same time. That is, one needs to know that A is south of B and B is east of C simultaneously to infer that A is southeast of C. In other words, information from multiple units and levels in a hierarchical network may need to be accessed simultaneously to allow inferences about spatial relationships across units. Compared with the vastness of our world and the infinite capacity of long-term memory, however, human information processing abilities are comparatively limited; for example, the capacity of visual working memory is often estimated to be only four to six items (e.g., Irwin, 1992; Luck and Vogel, 1997). Thus, it is an empirical question whether multiple levels of a hierarchical structure can be accessed simultaneously.

III. ACCESSING MULTIPLE ENVIRONMENTAL REPRESENTATIONS

To examine the issue of whether people are able to access spatial knowledge of nested environments (e.g., a building and a room in the building) at the same time, Brockmole and Wang (2002; see also Brockmole and Wang, 2003) employed a task-set switching paradigm where, across successive trials, participants made spatial judgments about objects drawn from the same environment or from different environments. If people can access representations of both environments at the same time, then participants should be able to switch between representations freely without a cost in performance. If, however, only one environment can be accessed at a time, then switching between environments would be associated with a temporal cost as subjects deactivate the old representation and activate the new representation. In one experiment, subjects evaluated target locations in a building and in an office within that building; in another experiment, locations on a college campus and a building within that campus were evaluated.

Consistent with the sequential access account, participants required additional time to judge spatial relationships immediately following a switch in the probed environment, an effect that could not be attributed to switching between two semantic categories (objects in a room environment versus rooms in a building environment). Strikingly, the effect also occurred despite the fact that each participant had at least 2 years of daily experience in the tested environments (see Fig. 26.1).

These results suggest that representations of nested environments are not accessed simultaneously, but rather, they are accessed one at a time. When a switch of representation is required, the inactive representation must be retrieved, a process that takes time. These findings put constraints on how spatial inference may occur in the hierarchical network model. That is, spatial inference across representations cannot be the simple combination of two or more concurrently accessed representations at different hierarchical levels. Either other mechanisms are needed to mediate such a combination process, or these environments, though familiar, are not represented in an integrated hierarchical network.

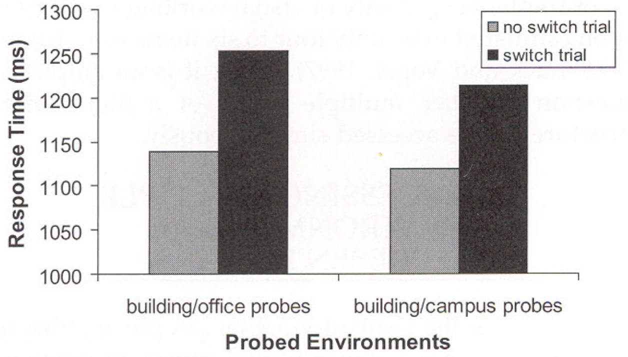

FIGURE 26.1 Results from Brockmole and Wang (2002). Response time to trials that previously probed a different environments (switch trial) was slower than that to trials that previously probed the same environment. A switch in probed environment entailed a switch in representation.

IV. SPATIAL UPDATING IN NESTED ENVIRONMENTS

Further evidence that representations of multiple, nested environments are not processed simultaneously comes from research on spatial updating in nested environments. Successful navigation requires that navigators keep track of their relationship to their surroundings as they locomote through the world (i.e., spatial updating). For example, wherever you are while reading this, look up from the page and notice what is around you. Close your eyes and turn 90° to your right. How has your relationship to those objects changed? To answer this question, you had to update the viewer-to-object relationships of the objects surrounding you. For example, what was in front of you now is on your left.

Spatial updating seems to require time (e.g., Rieser, 1989) and appears to occur automatically in some cases (e.g., Farrell and Robertson, 1998). For example, after an observer moves in an environment, he has great difficulty ignoring that movement and behaving as though his position relative to surrounding objects did not change. There is some indication that attention may play a role in the spatial updating process. For example, performance in recognizing an object after participants walked to a new viewpoint depends on whether they were instructed to focus on the object along the path or to focus on their own movements (Amorim and Stucchi, 1997).

Little research, however, has examined the processing limitations of spatial updating. It is not clear, for example, whether people update their relationship (consciously or not) with respect to multiple, nested environments simultaneously, or how many environments one can update at a time. Every object in the universe shares some relationship to you, and by moving, all of those relationships change. Thus, the amount of processing necessary to update one's position with respect to all known locations may be quite high. For example, if people represent their environments as an integrated network of spatial relations, as in a hierarchical network model, then knowing one's position in one environment would allow one to infer his or her position with reference to other environments in the network. As a result, one should be able to keep track of one's relation to all environments as a whole. On the other hand, if environmental representations are relatively independent, spatial updating may have limited capacity leading to the updating of some environments but not others.

To explore spatial updating across environmental representations, Wang and Brockmole (2003a) examined whether people can simultaneously update multiple targets both in their immediate surroundings and in a more remote environment. Participants sat in a swivel chair and learned the relationship between themselves and objects immediately surrounding them as well as familiar campus landmarks. Subjects were then blindfolded. One group was instructed to turn to face the objects in the room in sequence. Another group was told to turn to face the campus landmarks in sequence. This manipulation either foregrounded or backgrounded each environment in terms of explicit spatial updating. From their final position, both groups were asked to point to the room and campus targets. Accurate pointing reveals a maintained knowledge of viewer-to-object relationships.

The group that explicitly turned relative to campus locations was equally accurate at pointing to room and campus targets (see Fig. 26.2). This indicates that they updated their heading relative to both environments. The group that explicitly turned relative to objects in the room, however, could only identify the locations of room objects with accuracy. The accurate relationship between themselves and campus locations was lost. This indicates that participants in this group updated only their heading relative to the room. Together, these results indicate that people automatically update only object locations in their immediate surroundings. More remote environments are not implicitly updated. Keeping track of one's heading in remote environments requires that the observer explicitly attend to her or his movements with respect to that environment. Thus, spatial updating does not always operate with respect to all known environments simultaneously; what is automatically updated is determined by the nature of the environment.

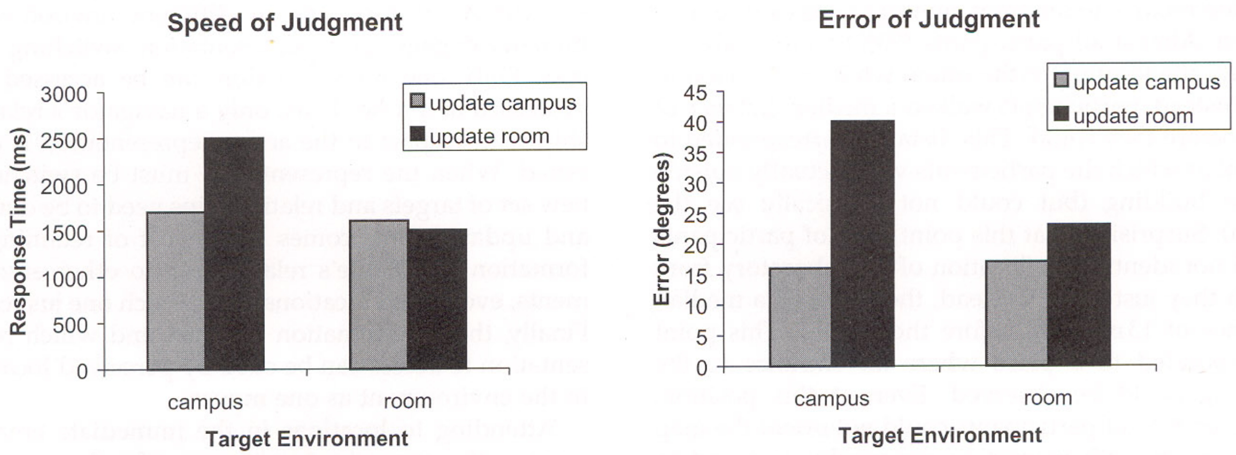

FIGURE 26.2 Results from Wang and Brockmole (2003a). Participants who updated according to room locations were slower and more error-prone when probed about campus locations, but participants who updated according to campus locations were not hindered when probed about room locations.

V. NAVIGATION IN NESTED ENVIRONMENTS

The sequential access of environmental representations and the failure to always update one's heading with respect to multiple nested environments has important implications for navigation. How is the separation of, and the misalignment among, multiple environmental representations resolved during navigation? That is, how do we maintain an accurate representation of our relationship to the world as we navigate across nested environments? Wang and Brockmole (2003b) considered this question by examining what knowledge participants have of their position relative to environments in which they are not currently located and when changes in that information occur.

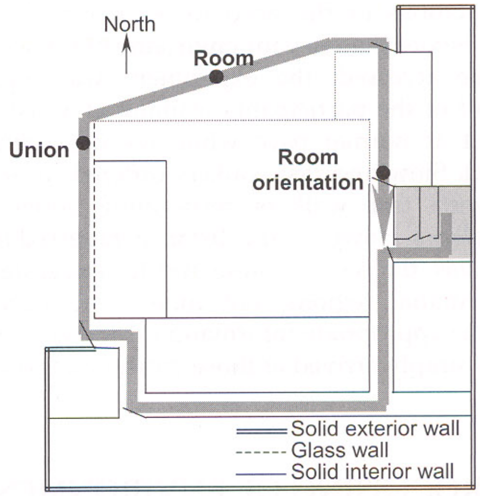

FIGURE 26.3 Method and results from Wang and Brockmole (2003b). The thick solid line shows the path of navigation. The filled dots represent the median response positions for the student union, the room, and the room's orientation. Participants alternated awareness of their position with respect to immediate and remote environments during navigation.

Participants walked along a prespecified path that originated in our laboratory, traversed several hallways, exited the building and continued along two sidewalks, and finally reentered the building through a different entrance and, ultimately, the laboratory once again (see Fig. 26.3). This path spanned three immediate environments: our laboratory, the building housing our laboratory, and the campus the building is on. While walking along this path, participants were queried about the location of various targets both internal and external to the building. While still in our laboratory (before embarking on the path) participants were asked to point to the student union, a familiar building on campus. If they could not immediately identify the union's direction, they followed the experimenter along the path and stopped and pointed to the union as soon as they were certain of its direction. At this point, they were asked to point to the room from which they came. Again, if they could not identify its location, they continued on the path until they could. Finally, they were asked to orient a map of our laboratory according to geography. If they could not, they continued on the path until they could.

Response accuracy was quite good, but all participants displayed clear alternation of awareness of their position relative to different environments during navigation. Almost all participants (93%) were unable to identify the location of the union while in our laboratory. Instead, participants walked a median distance of 55 m before they could. This distance corresponded to a point at which the participants were actually outside of the building (but could not physically see the union). Surprisingly, at this point, 80% of participants could not identify the location of the laboratory from which they just came! Instead, they walked a median distance of 13 m more before they could. This point corresponded to a place where an entrance to the building could be observed. Even at this position, however, 87% of participants could not orient the map of the room with respect to geography and had to continue a median distance of 22 m more before they could. This point corresponded to a position at which an entrance to the laboratory could be observed.

To examine whether the location changes that were needed to respond to each query resulted from time to access information from the relevant environmental representations or the need to reach certain spatial regions before the relevant environmental knowledge could be accessed, the experiment was replicated with half of the participants walking at a fast speed, and half at normal pace while response time was recorded. Slow and fast walkers differed in response time, with fast walkers responding sooner than slow walkers. However, the distance traversed by each group was the same. These results suggested that certain spatial regions, not time, were needed to access the appropriate information from memory: fast walkers simply arrived at those critical locations more quickly.

VI. PROCESSING OF ENVIRONMENTAL REPRESENTATIONS

Together, these studies suggest that people have separate representations for different environments, which are not organized into an integrated network such as a hierarchy, as the dominant view of spatial representations argues. The traditional hierarchical network model may require simultaneous access to information from multiple representations to make spatial inferences. Moreover, an integrated network of spatial relations, such as a hierarchical network, allows spatial updating that occurs within one unit to be generalized to other units. Thus, updating does not need to operate on each unit individually and should not be "capacity limited."

Human performance fails to support these predictions, however. When completing tasks that require knowledge of locations in different environments, humans engage in a representation switching process. Only one representation can be accessed and processed at a time. Thus, only a navigator's relationship with respect to the active representation is computed. When the representation must be switched, a new set of targets and relationships need to be derived and updated. This comes at the cost of retaining information about one's relationship to other environments, even those locations from which one just came. Finally, the determination of when and which representation to access can be cued by perceived locations in the environment as one moves.

Attending to locations in the immediate environment or the upcoming environment for these computations is most appropriate given nonequivalence in the utility of all locations for this process. For example, upcoming targets serve a more functional role in guiding action. Once one passes through an area of space, continued orientation to that area is not necessary for progress toward the goal location. Furthermore, distant targets cannot be updated as accurately as near targets (the further away a location is, the more difficult it is to ascertain an accurate vector between oneself and that target). Thus, assuming spatial updating is a limited-capacity process, then it is no surprise that targets have to be constantly dropped and reintroduced at given spatial locations where cues (certain landmarks, etc.) can be found to activate those representations.

If people only keep track of their relationship to a limited set of targets in the overall environment, for example, their current environment and the environment they are approaching, they may have very poor knowledge about their spatial relationship to many other environments that they are not actively processing. Wang and Brockmole (2003a) showed that people indeed are constantly disoriented relative to remote environments even when they are perfectly oriented to the immediate surroundings. Why do we rarely have the subjective experience of disorientation or "being lost" despite the fact that we are almost always "disoriented" relative to many environments? One possibility is that the sense of being oriented is like the sense of visual richness we perceive. Despite limitations on how much visual detail can be attended and retained in online visual memory, as shown in the research of change blindness (for reviews, see Rensink, 2002, and Simons, 2000), we experience a rich, detailed visual world as long as we have perceptual access to the visual details of those things that fall within the scope of attention. Similarly, we may experience perfect sense of orientation as long as we know our relationship to the immediate surroundings, despite our poor knowledge of our relationship to many, or even most, parts of the world.

VII. CONCLUSIONS

In summary, our environment is parsed into a series of independent representations that can only be accessed one at a time. Thus, when people keep track of their relationship to one environment, they do not necessarily update their relationship to other environments. Navigation across nested environments involves updating one's relationship to the upcoming environments, and changes in the active representation can be cued by locations in the environment that are perceived during motion.

These conclusions are inconsistent with integrated network models such as a hierarchical network model, which predicts that one may need to access information from multiple representations to make spatial inferences and, by the same token, that they may maintain knowledge of their relationship to multiple environments concurrently. Environmental representations appear to be fragmented in nature, and the mechanisms by which they are processed have a limited computational capacity and therefore do not apply to all environments simultaneously. This emerging incoherence underscores the importance of understanding the limitations in how environmental representations are processed to develop a more complete view of the nature and structure of spatial representations. Whether the hierarchical network model can be revised to account for the current findings or if new models of environmental representations need to be devised constitutes an important avenue for future research.

References

Amorim, M. A., and Stucchi, N. (1997). Viewer- and object-centered mental explorations of an imagined environment are not equivalent. *Cogn. Brain Res.* **5**, 229–239.

Brockmole, J. R., and Wang, R. F. (2002). Switching between environmental representations in memory. *Cognition* **83**, 295–316.

Brockmole, J. R., and Wang, R. F (2003). Changing perspective within and across environments. *Cognition* **87**, B59–B67.

Farrell, M. J., and Robertson, I. H. (1998). Mental rotation and the automatic updating of body centered spatial relationships. *J. Exp. Psychol. Learn. Mem. Cogn.* **24**, 227–233.

Hirtle, S. C., and Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Mem. Cogn.* **13**, 208–217.

Irwin, D. E. (1992). Memory for position and identity across eye movements. *J. Exp. Psychol. Learn. Mem. Cogn.* **18**, 307–317.

Luck, S. J., and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature* **390**, 279–281.

McNamara, T. P. (1986). Mental representations of spatial judgments. *Cogn. Psychol.* **18**, 87–121.

McNamara, T. P., Hardy, J. K., and Hirtle, S. C. (1989). Subjective hierarchies in spatial memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **15**, 211–227.

Rensink, R. A. (2002). Change detection. *Annu. Rev. Psychol.* **53**, 245–277.

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *J. Exp. Psychol. Learn. Mem. Cogn.* **15**, 1157–1165.

Simons, D. J. (2000). Current approaches to change blindness. *Visual Cogn.* **7**, 1–15.

Stevens, A., and Coupe, P. (1978). Distortions in judged spatial relations. *Cogn. Psychol.* **10**, 422–437.

Taylor, H. A., and Tversky, B. (1992). Descriptions and depictions of environments. *Mem. Cogn.* **20**, 483–496.

Wang, R. F., and Brockmole, J. R. (2003a). Simultaneous spatial updating in nested environments. *Psychon. Bull. Rev.* **10**, 981–986.

Wang, R. F., and Brockmole, J. R. (2003b). Human navigation in nested environments. *J. Exp. Psychol. Learn. Mem. Cogn.* **29**, 398–404.

Wilton, R. N. (1979). Knowledge of spatial relations: the specification of information used in inferences. *Q. J. Exp. Psychol.* **31**, 133–146.