



Topics in Cognitive Science 3 (2011) 741–759

Copyright © 2011 Cognitive Science Society, Inc. All rights reserved.

ISSN: 1756-8757 print / 1756-8765 online

DOI: 10.1111/j.1756-8765.2011.01153.x

Representations and Processes of Human Spatial Competence

Glenn Gunzelmann,^a Don R. Lyon^b

^a*Air Force Research Laboratory*

^b*L3 Communications at Air Force Research Laboratory*

Received 4 December 2009; received in revised form 9 January 2011; accepted 1 February 2011

Abstract

This article presents an approach to understanding human spatial competence that focuses on the representations and processes of spatial cognition and how they are integrated with cognition more generally. The foundational theoretical argument for this research is that spatial information processing is central to cognition more generally, in the sense that it is brought to bear ubiquitously to improve the adaptivity and effectiveness of perception, cognitive processing, and motor action. We describe research spanning multiple levels of complexity to understand both the detailed mechanisms of spatial cognition, and how they are utilized in complex, naturalistic tasks. In the process, we discuss the critical role of cognitive architectures in developing a consistent account that spans this breadth, and we note some areas in which the current version of a popular architecture, ACT-R, may need to be augmented. Finally, we suggest a framework for understanding the representations and processes of spatial competence and their role in human cognition generally.

Keywords: Spatial cognition; Spatial visualization; Orientation; Reference frames; Cognitive architecture; Computational model

1. Introduction

The importance of spatial information processing extends beyond cognitive maps and mental rotation. Its tendrils extend throughout human cognitive processing, from reaching and grasping (e.g., Mulliken, Musallam, & Andersen, 2008), to reasoning about numerical information (e.g., Gevers, Verguts, Reynvoet, Caessens, & Fias, 2006), to problem solving (e.g., Fincham, Carter, van Veen, Stenger, & Anderson, 2002). Further, language is replete with spatial metaphors related to non-spatial concepts (e.g., Boroditsky, 2000; Lakoff &

Correspondence should be sent to Glenn Gunzelmann, Cognitive Models and Agents Branch, 711 HPW/RHAC, 2620 Q Street, Building 852, Rm. 3-312, Wright Patterson AFB, OH 45433-7905. E-mail: glenn.gunzelmann@us.af.mil

Johnson, 1980), and it has been proposed that spatial structure in the environment may provide a foundation for early conceptual development in children (e.g., Mandler, 1992). Moreover, there is evidence that the origins of complex spatial information processing abilities across mammals, birds, reptiles, and teleost fish may derive from a common fish ancestor that inhabited the earth some 400 million years ago (Rodríguez et al., 2002).

The foundational assumption that we derive from the broad and extensive research on spatial cognition is that spatial information processing is central to cognitive processing more generally. More specifically, we view spatial processing as a flexible and powerful cognitive tool that is brought to bear by diverse components of cognition to improve the adaptivity and effectiveness of perception, cognitive processing, and motor action. A corollary of this position is that there are important interconnections between spatial processing and other components of cognitive functioning. Traditionally, research on spatial cognition has looked at spatial processing in relative isolation, leading to a proliferation of theoretical accounts that unfortunately fail to generalize beyond a small set of spatial tasks.

Likewise, models of spatial processing have been confined to specific accounts of particular phenomena or aspects of spatial competence. These models have taken diverse forms, from models of reaction time in a mental rotation task (e.g., Bejar, 1990; Shepard & Metzler, 1971), to robot models of rodent navigation (e.g., Burgess, Donnett, Jeffery, & O'Keefe, 1999; Touretzky & Redish, 1996), but they do not provide a framework for understanding diverse ways in which spatial processing is leveraged in human cognition.

Our goal is to pursue a more integrated theory, which ties spatial information processing to other components of cognitive functioning (e.g., Gunzelmann & Lyon, 2007). Such integration is necessary to explain how spatial cognition is brought to bear across the myriad tasks and domains where spatial knowledge is critical for efficient and effective performance.

A vital component of our research is the use of a cognitive architecture representing a general theory of cognitive processing. Cognitive architectures provide an important methodological advantage for this research, because they comprise validated mechanisms for many aspects of cognitive functioning that are necessary to account for behavior in particular task contexts (Anderson, 1983; Newell, 1990). This informs the process of identifying candidate mechanisms by constraining the set of alternatives based upon existing theoretical commitments of the architecture itself. In addition, it allows us to address issues related to the integration of spatial cognition with other components of cognitive functioning by leveraging existing theories, rather than becoming encumbered by the need to propose and validate numerous mechanisms that are ancillary to the central focus of the research.

Perhaps the biggest advantage of using a cognitive architecture is that one can work simultaneously at different levels of complexity. For example, cognitive architectures allow the modeling of behavior in complex, realistic tasks in which spatial cognition is crucial. There are many applied contexts, ranging from simulation-based training to geographic information systems, where understanding spatial cognition and its role in cognitive functioning could lead to innovations that improve quality of life, and even save lives. While our collaborators are building models of complex behavior in naturalistic task environments (e.g., Ball et al., 2009), we can use the same cognitive architecture to address in detail key

components of spatial cognition using simpler laboratory tasks to isolate phenomena of interest (e.g., Gunzelmann, 2008; Lyon, Gunzelmann, & Gluck, 2008).

In this paper, we present an example of this multi-level approach. We first describe a computational model of behavior in a complex, realistic task. We then show how laboratory tasks can be used to understand the spatial information processing that influences performance on the complex task, and how those processes can be modeled using the same cognitive architecture. In the course of these latter examples, we will show how the measurement and modeling of a particular building block of spatial competence cannot be understood easily without considering the influence of other aspects of the cognitive system.

2. Multilevel analysis of spatial cognition using a cognitive architecture

2.1. Top level: Performance on a complex task

Technological innovations, like remote sensing and simulated virtual environments, have created new application opportunities and new avenues for research on spatial cognition. Activities like training in simulated virtual environments and piloting unmanned aerial vehicles (UAVs) are increasingly prevalent. These technologies go well beyond more traditional spatial tasks like you-are-here maps, navigation, and perspective taking, which historically have received more research attention (e.g., Huttenlocher & Presson, 1979; Levine, Marchom, & Hanley, 1984; Malinowski & Gillespie, 2001). Moreover, they are having a substantial impact on modern society, yet the complex interplay of spatial information processing, perception, action, and higher-level reasoning involved in their use is still poorly understood (see Kheener, current issue). More sophisticated theories of spatial cognition are needed, which can explain performance in these new domains of human activity.

Dimperio, Gunzelmann, and Harris (2008) describe a computational process model that flies a UAV in a synthetic task environment to perform reconnaissance missions. It uses a virtual stick and throttle to maneuver a simulated Predator UAV so that high-resolution surveillance footage of a ground target is obtained through a hole in a layer of clouds. An illustration of the task environment is shown in Fig. 1. The goal is to maximize “time on target” while minimizing violations of various restrictions (e.g., “no-fly” zones, altitude limits). This is a complex task requiring spatial reasoning and planning but also involving significant non-spatial processing. For instance, substantial knowledge is required to simply control the aircraft appropriately to avoid situations that cost points in the simulation but that might put the aircraft at risk in an operational setting (e.g., violations and engine stalls).

As the task requires that various cognitive activities and functions be fully integrated, Dimperio et al. (2008) used a cognitive architecture—specifically the Adaptive Control of Thought-Rational, or ACT-R, architecture (Anderson, 2007)—as the theoretical foundation for the model. The model interacts directly with the task software, and it does a relatively good job of matching human performance on measures like time on target and penalty time (see Dimperio et al., 2008). It also produces routes that are qualitatively similar to the routes

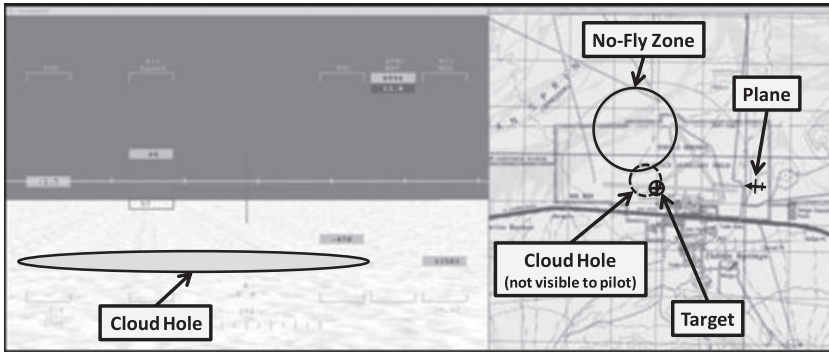


Fig. 1. Illustration of the unmanned aerial vehicle (UAV) synthetic task environment (STE) reconnaissance mission task. On the left is a nose camera view from the plane of the layer of clouds, with a cloud hole visible in the distance. Superimposed on this view is a heads-up display (HUD), providing critical information related to the UAV's performance. On the right is a map of the area, which also shows the location and orientation of the UAV and the target location. The location of the cloud hole is shown here for illustration, but it was not visible to participants in the experiment.

flown by some expert pilots (Fig. 2). Importantly, the spatial reasoning in the model is accomplished largely through mechanisms that are not grounded in psychological theories of spatial cognition. Consequently, as currently implemented there are important questions about the psychological validity of some aspects of the model. However, valid mechanisms for these processes are critical for developing a detailed scientific understanding of human performance in complex tasks, as well as in applied settings where psychological models may be utilized by decision makers on issues ranging from training, to workload, to interface design.

A full description of the model is beyond the scope of this paper. More important than the details, however, is a primary motivation for its development, which was to expose gaps in theoretical coverage within ACT-R, particularly relating to spatial information processing. Table 1 lists several spatial processes that are needed by our UAV model to accomplish the reconnaissance task but are not explained well using mechanisms that compose the current ACT-R architecture.

Consider one specific example from Table 1—route planning. In the task, the model must plan an initial route to fly over the cloud hole. Although the cloud hole may be visible through the nose camera on the plane, the field of view is limited (about 30°) and distance information is lacking (see Fig. 1). Sometimes, a no-fly zone blocks a direct path, necessitating some planning to maneuver around the no-fly zone and get back on the intended trajectory. Such planning is straightforward for humans, but in Dimperio et al. (2008), it was necessary to use trigonometric functions and waypoints implemented directly in code, due to a lack of validated computational mechanisms for performing such reasoning in a human-like manner. While effective in the task, these processes are unlikely to reflect cognitively valid mechanisms for this kind of planning.

Another example involves coordinating spatial information between the camera-based representation on the left display and the map-based perspective on the right in Fig. 1

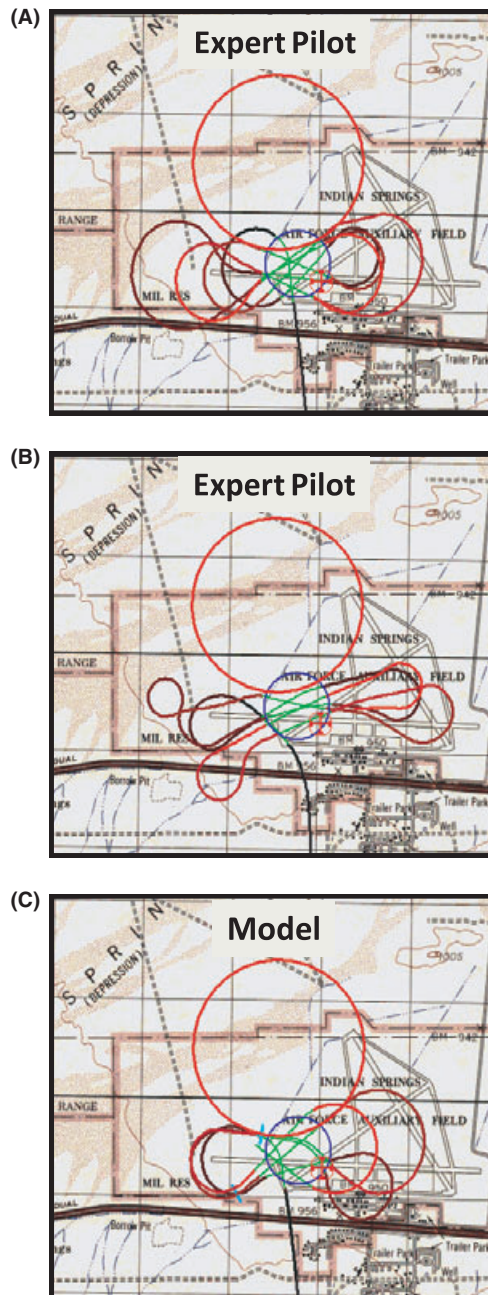


Fig. 2. Sample flight paths for two expert human pilots (A and B) and the model (C). The image presents a top-down view of the flight path during a 10-min trial. The lines transition gradually from black to red, with green portions illustrating periods where the camera was obtaining surveillance footage of the target. The cloud hole is illustrated for information purposes in blue. The model produces paths that are qualitatively similar to many human flight paths, although the human flight paths were quite variable.

Table 1

Spatial abilities required for the UAV model but not represented by validated mechanisms in ACT-R

Spatial Ability	Relevance in UAV Model
Encoding spatial location	Representing relative locations of visible objects, including the plane, the target, and the cloud hole.
Mental rotation	Reasoning about aircraft maneuvering based upon map-based representation of plane location.
Frame of reference transformations	Reasoning between the egocentric perspective from camera and the map-based representation.
Spatial reasoning for route planning	Finding and subsequently returning to the area where the target is visible through the cloud hole.
Mental simulation of spatial action	Anticipating consequences of maneuvers; reasoning about strategies for maximizing time on target.
Visualization and visuospatial memory	Constructing and maintaining mental images of key spatial information, such as: the map region from which the target is viewable; the predicted track of the aircraft; and the “cone” relating altitude to target-viewing area.

(*Frame of reference transformations* in Table 1). The model uses the first-person view for little more than initially locating the cloud hole and orienting the plane appropriately. Once that is accomplished, the spatial planning and reasoning is based almost entirely on map-based information and tends to be mostly heuristic in nature (e.g., fly higher and more slowly while over the cloud hole). Whereas this is sufficient in the context of this task, there are many other contexts where more explicit coordination of in situ and map-based information is required. In fact, the situation noted in the previous paragraph—planning a route to avoid a no-fly zone—represents an instance of this, as the cloud hole is visible in the first-person (egocentric) view, and the no-fly zone is indicated on the map (exocentric view). This kind of task, which exposes important gaps in the representational and processing capacities of ACT-R, has been the focus of the research discussed next, which utilized a simpler laboratory-based task to hone in on this process more closely.

2.2. Component level example 1: Reasoning about frames of reference

In the UAV reconnaissance task, and in a variety of real-world settings, information is distributed across multiple internal and/or external representations. The particular situation in the reconnaissance task is common—perceptual information about the world from a first-person perspective must be coordinated with map-based representations in a variety of navigation contexts (e.g., Levine, Jankovic, & Palij, 1982; Malinowski & Gillespie, 2001; Péruch & Lapin, 1993). We have conducted research on the coordination and integration required in this kind of task to address how spatial information processing abilities are leveraged to make various sorts of judgments (Gunzelmann, 2008; Gunzelmann & Anderson, 2006; Gunzelmann & Lyon, 2006). In one version of the task, participants are shown a visual scene, depicting a circular space containing a number of objects from a viewpoint somewhere along the edge. Along with this egocentric view, an exocentric map is shown,

which also shows the object locations. An example is illustrated in Fig. 3. Participants were asked to click on the edge of the map to indicate the viewpoint for the visual scene. The response was scored as being correct when it fell within $\pm 15^\circ$ of the actual viewpoint.

The task illustrated in Fig. 3 was challenging for participants to perform, and error rates were in the range of 25%–30%, even after significant practice (i.e., hundreds of trials). What is most interesting about the errors, however, is the distribution of responses relative to the correct answer. Fig. 4 shows the response proportion as a function of the angular deviation from the actual viewpoint (the first three points represent responses scored as correct). The results clearly show that errors in the task were not random. Rather, they tended to be quite close to the correct answer. The pattern suggests that participants were able to understand

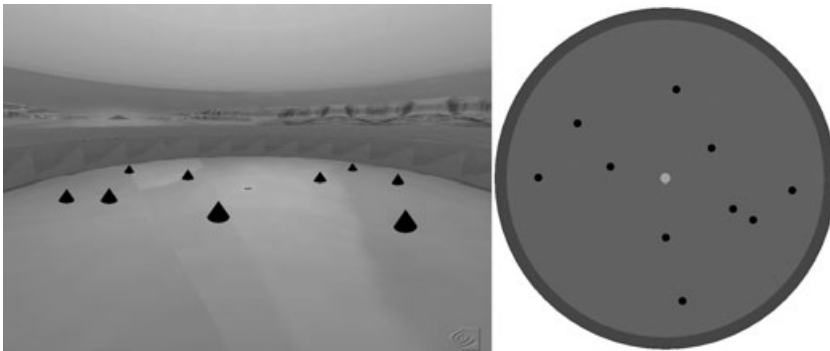


Fig. 3. Sample trial for a task where participants must identify their location on the map, given the first-person view of the environment on the right. Responses are made by clicking within the darker shaded ring around the edge of the map, and they were scored correct if they fell within $\pm 15^\circ$ of the actual location. In this trial, the viewer is positioned northeast of the center, looking southwest.

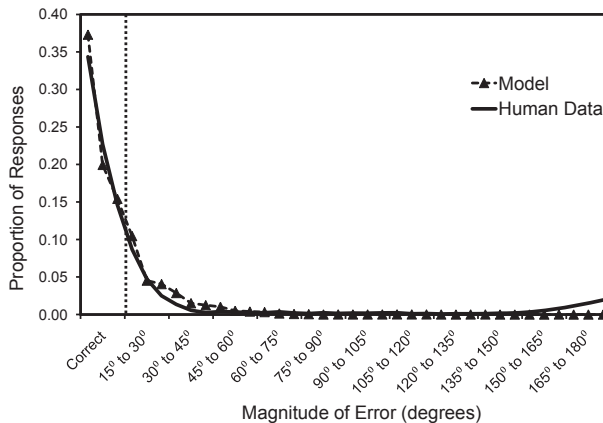


Fig. 4. Distribution of responses for the model and human participants as a function of angular deviation from the actual viewpoint. Responses were scored as correct when they fell within $\pm 15^\circ$ of the actual location.

and identify the correspondence between the map and the visual scene generally but often failed to estimate the location of the viewpoint precisely enough.

In modeling human performance on this task, we have drawn upon other behavioral and neuropsychological research that has argued for a distinction between qualitative and quantitative spatial processes (Baciu et al., 1999; Kosslyn, Sokolov, & Chen, 1989). The model makes a response by executing a series of productions that implement a two-step strategy for the task. The first is a qualitative step, where the model identifies corresponding groups of objects in the visual scene and on the map. This information is used to narrow the potential response area to a region along the edge of the map (i.e., an arc where the qualitative spatial relations are preserved). Then, in a second step, the model uses quantitative estimates of the egocentric bearing to particular groups in the visual scene to refine its estimate of the viewpoint location within that potential response area on the map. Specifically, the model responds when it finds a location on the edge of the map where the bearing estimates are *close enough* to what was encoded from the visual scene, based upon prior experience with the task. Finding the correct response location is challenging because of the differing reference frames but also because encoding of quantitative information is imprecise. As a result, the model makes errors, and it produces a pattern of responses that closely approximates the performance of human participants (Fig. 4).

This model begins to illustrate how qualitative and quantitative spatial reasoning may be integrated in human performance to accomplish operations in Table 1 like spatial encoding and reference frame transformations. Variations of this task have also been used to show how a single general strategy may be adapted for use in different particular task contexts (Gunzelmann, 2008). However, even in this task, the *spatial mechanisms* are largely abstracted and captured in parameters like *rotation speed* or *spatial update time*. In other words, this task is more about using spatial knowledge than isolating the representations and processes that are involved at the architectural level. To focus even more precisely on those details, we have developed a task that taxes spatial processing specifically, and which allows us to address some of the foundational mechanisms involved in visualizing spatial information to support decision making.

2.3. Component level example 2: Spatial visualization

Visualizing spatial material is a key aspect of solving many kinds of spatial problems, including route planning, mental simulation of spatial action, and visuospatial memory processes mentioned in Table 1 in the context of our UAV reconnaissance task. To further sharpen our focus on understanding the elementary representations and processing mechanisms that underlie these abilities, we studied visualization using the *path visualization* task (Lyon et al., 2008). The goal was to tap the process of visualizing complex spatial material, while minimizing extraneous aspects of task complexity. The simplest versions of the path visualization task are similar to the experience of hearing or reading a verbal description of how to walk or drive from one point to another, except that paths can be in three dimensions. The participant hears or reads a sequence of segments for a path wandering randomly through an undifferentiated, $5 \times 5 \times 5$ grid space (for example “Right 1,” “Up 1,” “For-

ward 1,” with segment length always 1 step). Each time a segment is presented, a key must be pressed to indicate whether or not the new path segment has intersected with any previous part of the path. Both accuracy and response time are recorded; the latter primarily to assure that any effects on accuracy are not due to speed-accuracy tradeoffs.

Path visualization is similar in some ways to tasks used by Brooks (1968), Attneave and Curlee (1983), Kerr (1987, 1993), Diwadkar, Carpenter, and Just (2000), and others. However, we have argued (Lyon et al., 2008) that this task is particularly good at forcing participants to actually perform complex, very effortful visualization, rather than using alternative strategies such as verbal rehearsal or numeric recoding. We believe that this provides us with an objective measure of accuracy in spatial visualization, thereby revealing some of the details regarding the underlying mechanisms in this component of human cognitive functioning.

We have built and tested a detailed computational model (implemented in ACT-R) of visualization accuracy for this task (Lyon et al., 2008). The model proposes several processes that operate on chunks representing segments of the path, localized within a cognitive representation of space. According to this model, the following three processes are the major sources of error in this kind of spatial visualization: (a) decay of chunk activation; (b) associative interference; and (c) spatial interference. Decay and associative interference are part of virtually all ACT-R models of declarative memory, but spatial interference is a new process that helps to explain the strong effects of spatial proximity found in our data. This mechanism represents greater confusability between elements being visualized when they are closer together. In path visualization, this leads to incorrectly identifying an intersection when a nearby node has been visited. These proximity effects in visualization could prove to be related to crowding effects in vision itself (Lyon, 2009), although evidence for this idea is currently inconclusive.

The model was based on experiments in which verbal descriptions of path segments were exocentric, so for example, “Up” was always toward the top of the space and did not vary with the facing direction of a hypothetical traveler along the path. There is, however, another important class of situations in which spatial visualization plays a role. People frequently attempt to visualize paths described egocentrically, with reference to their own virtual facing direction (“Turn right, go 2 blocks, then turn left”). Although the description is egocentric, often they need to interpret the result in terms of an exocentric, map-like representation.

It is easy to study both exocentric and egocentric path descriptions using the path visualization task, because any path can be described either way. For instance, a square can be described in egocentric terms as a sequence of four “Right 1” segments. The corresponding description in an exocentric reference frame would be “Right 1,” “Back 1,” “Left 1,” “Forward 1.” When we compared performance using these two alternatives for presenting path information, we found that egocentrically described paths produced substantially more errors (and longer response latencies) than exocentrically described paths (Lyon, Gunzelmann, & Gluck, 2007). A possible explanation for this result is that the additional errors in the egocentric condition are due to an egocentric-to-exocentric translation process that is used to “construct” a representation of the path in an exocentric reference frame. There-

fore, we extended our computational model to incorporate this process. Our data show that the time required to perform egocentric-to-exocentric translation increases substantially as the orientation of one's (virtual) body differs from upright and facing forward (Fig. 5). Once this factor was incorporated into the model, we could account for the difference in accuracy between egocentric and exocentric conditions. Interestingly, the model explained this result based solely on the increased decay associated with the longer latencies produced by the translation process (Lyon & Gunzelmann, 2009).

Another important aspect of the path visualization task is that the same set of paths can be presented using a variety of different methods. The results discussed above came from experiments in which people visualize paths that are described verbally. However, UAV operators experience both verbal descriptions of paths and the virtual experience (from the nose camera feed) of moving along a path, turning, climbing, etc. To examine the influence of this perceptual information, we compared error rates for egocentrically described paths to paths experienced through virtual motion (Lyon, Gunzelmann, & Gluck, 2006). The result was clear—virtual motion did not improve visualization accuracy. Of course this result applies only to the process of visualizing the abstract path itself, including the ego-to-exo translation process. Visualizing landmarks is a very different matter, and it would be expected to benefit from visual depiction over verbal description.

The results using the path visualization task expose some details regarding the mechanisms of spatial visualization identified in Table 1. They suggest that the mechanisms that cause errors in visualizing a mental map operate in a similar way for very different percep-

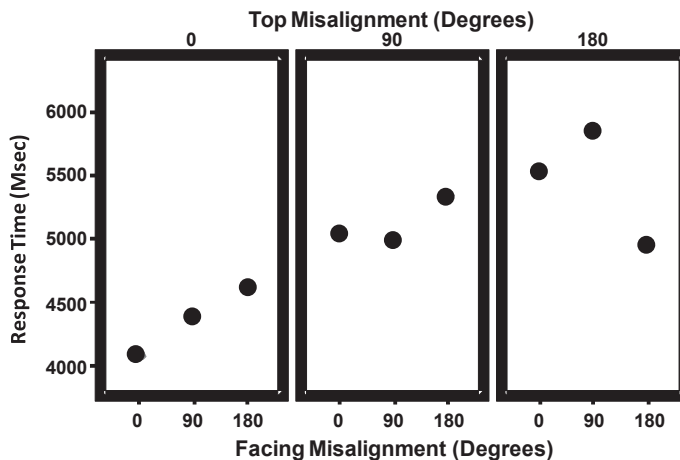


Fig. 5. Speed of mentally constructing each segment of a 3D allocentric “map” of (virtual) movements through a grid, shown as response times (ms) for correct responses in the task as a function of the misalignment of the imagined orientation. The time to visualize a map segment increases with increasing misalignment of one's virtual body from an upright, forward-facing orientation. For example, a *top misalignment* of 180° would be standing on one's head; a *facing misalignment* of 90° would be facing either right or left. Note that the faster response time in the 180°–180° misalignment condition appears to reflect a somewhat easier situation, where the rotations produce a mirror-image rotation that can be computed more quickly (e.g., Gunzelmann, 2008).

tual inputs (Klatzky, Wu, & Stetten, 2008). In addition, the model illustrates that maintaining visualized information may be subject to many of the same dynamic influences as other forms of declarative knowledge. These findings speak to the nature of the representations and processes that are needed in developing a quantitative account of human spatial competence.

2.4. Summary

The tasks and models described in this section represent an example of an architecture-guided, multilevel research strategy that addresses both complex, naturalistic tasks (UAV reconnaissance) and controlled, laboratory examinations of spatial processing in two different tasks. By using a cognitive architecture that is capable of modeling both complex tasks and key component processes, we can hope to eventually show exactly how spatial competence is integrated with non-spatial cognitive processes, and how elementary spatial processes support complex cognitive functions like those mentioned in Table 1. In the next section, we look at the other side of the coin, namely, the gaps and shortcomings that we have encountered in trying to apply the architecture-guided strategy using a specific cognitive architecture—ACT-R.

3. Critique of the architecture-guided approach to spatial competence

The focus of this section is on evaluating what additional theoretical components may be necessary to create a robust and general ability to represent and manipulate spatial knowledge within a general theory of human cognition. We discuss this in the context of ACT-R. However, we concentrate on components tied directly to spatial information processing capacities and avoid, to the extent possible, ACT-R jargon and details. While the context is ACT-R, we believe that our claims are generally consistent with the current state of most other cognitive architectures as well.

3.1. Spatial representations

Spatial cognition requires some way of representing spatial information perceived from the environment. Klatzky et al. (2008) argued that spatial representations may be amodal and derive from multiple sensory modalities, or indeed, emerge from their integrated contributions. Currently, ACT-R (and other cognitive architectures) lacks a capacity to derive spatial knowledge from perceptual input (*Encoding Spatial Location* in Table 1). Rather, the default ACT-R representation of spatial location is as screen coordinates, with depth represented as a static value based upon a standard viewing distance of an individual from a computer display.

To appropriately capture the dynamics of human interaction with the environment, more psychologically based representations are required, specifically representations grounded in both egocentric and exocentric frames of reference. This reflects evidence that both kinds

of representations play important roles in human spatial cognition (e.g., Burgess, 2006; Franklin & Tversky, 1990; Mou & McNamara, 2002; Sholl & Nolin, 1997), though we suggest that higher-level processing may be necessary to derive exocentrically based representations (see Harrison & Schunn, 2003; Klatzky, 1998). This conjecture is supported by our research using the path visualization task (Lyon & Gunzelmann, 2009).

Egocentric representations play a key role in immediate action, like reaching or grasping (e.g., Daprati & Gentilucci, 1997), and object avoidance during navigation (Harrison & Schunn, 2003). At the same time, a long history of research in neuroscience has established the role of exocentric representations in spatial processing, particularly the role of *place cells* (e.g., O'Keefe & Dostrovsky, 1971; O'Keefe & Nadel, 1978), and more recently *boundary vector cells* (e.g., Hartley, Trinkler, & Burgess, 2004) in coding the spatial location of oneself and other objects in the external environment. It is these representations, and the corresponding processing mechanisms, which we hypothesize form the foundation for human spatial competence across the diverse range of tasks and domains mentioned in the introduction. Our research using path visualization has helped to uncover basic characteristics of the underlying representational structures, while our experiments investigating reasoning about reference frames speak to the precision and the flexibility with which they can be used.

3.2. Spatial transformation processes

There is considerable evidence that the human brain implements processes specialized for performing various transformations on the spatial aspects of perceptual input. These transformations are believed to be critical for coordinating perception and action as part of the dorsal stream in visual processing (e.g., Milner & Goodale, 1993). Moreover, neural mechanisms that elegantly support reference frame transformations have been identified in parietal cortex (Andersen, Essick, & Siegel, 1985; Andersen & Mountcastle, 1983; Zipser & Andersen, 1988). Specifically, collections of cells in this area create *gain fields*, which can efficiently transform spatial representations from one coordinate system to another (Xing & Andersen, 2000). For instance, gain fields can transform retinotopic information derived from vision to 'head-centered' coordinates to guide visual attention (e.g., Smith & Crawford, 2005), or to a 'hand-centered' reference frame that is suitable for reaching and grasping an object (e.g., Mulliken et al., 2008).

Note that head-centered and hand-centered reference frames are, in spatial terms, different types of egocentric reference frames, as they represent spatial location with respect to the coordinate system defined by a part of the individual. A remaining issue relates to whether and how gain fields may be used to convert egocentric visual input to an exocentric representation, perhaps leveraging boundary vector cells and place cells to support the encoding and transformation processes. This is a critical process in spatial visualization and is one focus of our current research (Gunzelmann & Mohebbi, 2010). This research extends our research using path visualization by investigating the details of the egocentric-to-allo-centric translation process that is an important mechanism in the model, but where the details of the processing mechanism are lacking (see Table 1, *Frame of reference transformations*).

Once exocentric representations have been computed, they may be used in spatial computations like constructing complex representations of spatial layouts, distance and bearing estimates, and magnitude comparisons. Note that these operations are not free or automatic in this perspective but require effortful processing and attention to combine and integrate elemental spatial knowledge. We believe that a relatively small set of basic transformations exists, which can be combined in various ways to support a variety of higher-level spatial reasoning operations, from mental rotation (e.g., Shepard & Metzler, 1971), to map-based orientation and reasoning (e.g., Gunzelmann, 2008; Gunzelmann & Anderson, 2006), to complex mental simulation (e.g., Hegarty, 1992; Trickett & Traflet, 2007). These are all identified in Table 1 as existing limitations that play important roles in understanding performance on a complex task like UAV reconnaissance. Drawing upon the orientation-with-maps research described above, we suspect that both qualitative and quantitative processes are available, which likely reflect the functions of distinct cortical areas (e.g., Baciú et al., 1999; Kosslyn et al., 1989).

Unfortunately, these transformation processes are not easy to implement naturally in ACT-R. For one thing, the basic representations that they depend upon are not currently part of the architecture. This makes the goal of integrating research findings on spatial transformations with what is already known about other cognitive processes (such as declarative memory retrieval or goal-directed problem solving) all but impossible to accomplish without first attending to the representational issues described in Section 3.1.

3.3. Spatial visualization

When the transformation processes discussed above are applied to foundational spatial knowledge in the context of a spatial task, the result is, at some stage, a higher-level representation that integrates spatial and perceptual information to instantiate a representation of the emerging situation (c.f., Zwaan & Radvansky, 1998). Therefore, performance on complex spatial problems such as those embedded within the UAV reconnaissance task will be a function of both the efficacy of spatial operations and the accuracy of maintaining the spatial visualizations of the intermediate and final results.

As noted earlier, we have succeeded in adapting ACT-R's declarative memory component to model various phenomena in the visualization of complex paths (Lyon et al., 2008). We believe that this has led to some new and interesting insights about visualization (for example, the existence of lateral-masking-like interference in spatial images). However, research currently in progress suggests that this may not be the best long-term solution to modeling visualization in a general way and that a different kind of representation is needed.

One possibility is that spatial visualizations share representational features with perceptual encoding, as suggested by research on mental imagery (e.g., Kosslyn, 1994; Kosslyn, Thompson, & Ganis, 2006) and by the evidence for spatial interference in our path visualization results. However, even if this is the case, our results also clearly indicate that these representations are also affected by processes similar to those that influence the availability of semantic knowledge in declarative memory (e.g., decay and associative interference). Thus, one issue to be addressed is the relationship between

foundational spatial knowledge (discussed in Section 3.1) and spatial visualizations, which include more perceptually grounded information to contextualize the spatial information.

3.4. Episodic knowledge

Another requirement in our conceptualization of an architecture-based approach to spatial cognition is an episodic representation that integrates multisensory information to store episodes in declarative knowledge. In a sense, episodic representations are available in ACT-R, but there is a need for integration across sensory modalities to more specifically capture episodic *experiences* (or *situations*; Zwaan & Radvansky, 1998) that are not well-explained currently in the architecture. In terms of spatial processing, something like an “episodic chunk” would not only provide a robust encoding of particular events and experiences but also provide a means for creating and manipulating mental images. If episodes were stored as a form of declarative knowledge, this would imply that they may be retrieved, essentially re-instantiating a perceptual experience as a mental image (or visualization) for re-inspection. At that point, spatial transformations to that mental image, such as translation, rotation, and rescaling, could be performed.

In our view, episodic representations provide for another primary function of mental imagery besides the re-instantiation of a particular perceptual experience. The representations in this buffer must be rich enough to incorporate the *generation* of specific novel characteristics of visualized objects and spatial configurations. Researchers studying the nature of computation with visualized representations (e.g., Barkowsky & Freksa, 1997; Glasgow & Papadias, 1992) argue that at least some of the power of such representations comes from the requirement that spatial visualization must make commitments about many aspects of the situation that it is representing. A visualized line must have some extent and thickness; an object must have a particular position relative to others. ACT-R provides a possible mechanism for achieving this, as declarative knowledge can specify “default” values for attributes of different types of knowledge, which can augment spatial visualizations when details were not encoded or have been forgotten.

3.5. Summary

Fig. 6 is our attempt to construct a visual representation of the gaps we have encountered in trying to apply an architecture-based approach to spatial cognition. It depicts the current components of ACT-R (illustrated in general terms), which reflect a degree of consensus about the general processes involved in cognition, at least with regard to verbal and procedural content. The foundation is a perceptual-cognitive-action (motor) loop, which is influenced by stored knowledge. Spatial processing is represented as an additional influence on the flow of cognitive processing, which has deep connections to the perceptual, cognitive, and action processes. This reflects our claim that spatial information processing is central to effective and adaptive interaction with the environment. *Spatial competence* consists of this set of hypothetical new elements, along with interconnections among them and other com-

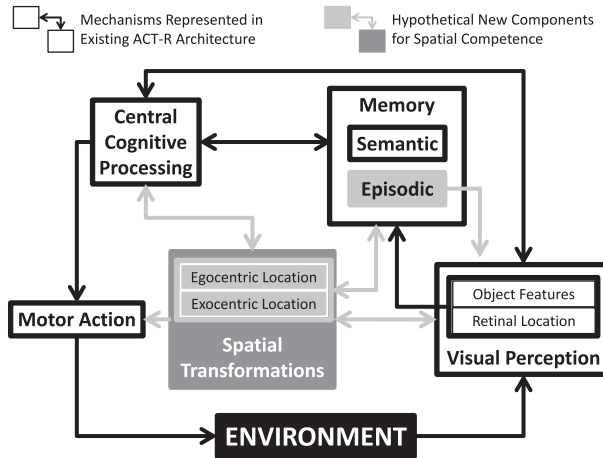


Fig. 6. Conceptual illustration of the components of our general theory of human spatial competence and their integration with other aspects of cognitive functioning.

ponents of cognitive functioning. If implemented in an empirically justified way, we believe this could fill many of the gaps in the current architecture.

4. Conclusion

In this paper, we have discussed research targeted at developing a detailed and general account of human spatial competence. By spatial competence, we mean to suggest something more than spatial information processing in isolation. Rather, we include the integration of those mechanisms with other components of cognition to understand the representations and processes supporting spatial ability, as well as their role in overall cognitive functioning. We have suggested that using a cognitive architecture to guide research and modeling can contribute to an integrated picture of cognition that includes spatial processes, but we have also noted gaps in one particular architecture's current representation of spatial thinking.

Our research addresses some of the existing gaps. Table 1 shows several limitations of ACT-R in the context of a model that pilots a UAV to perform reconnaissance. In Section 2, we described ongoing research to isolate phenomena of interest and develop more detailed theoretical mechanisms. In Section 3, we discussed how the capacities and mechanisms we have identified could help to address many of the existing limitations. The ultimate goal is to create a more comprehensive theory of the human cognitive system, which is important both for our scientific understanding of the human mind and for real-world applications aimed at understanding and improving human performance in a variety of domains.

Note that the representations and processes associated with spatial competence occupy a central position in Fig. 6. This reflects our position, outlined in the introduction, that spatial information processing represents a powerful and flexible information processing tool that

is drawn upon pervasively by perceptual, cognitive, and action processes. This includes coordinating perceptual inputs with motor actions, as well as higher level cognitive functions ranging from reasoning and problem solving to language and mathematics. This enables the cognitive system as a whole to achieve more adaptive performance in a spatially rich and complex environment, making planning, problem solving, and decision making more effective in a variety of circumstances.

Finally, it is worth noting that our goal is to develop more than a superficial integration of these mechanisms with a particular unified theory of cognition—ACT-R. Instead, the focus is on using the architecture explicitly to aid in refining and specifying critical details, leading to a more comprehensive theory of the human cognitive architecture and enabling substantially greater functionality than exists today. Moreover, we do not mean to imply that the best research strategy would be to simply begin implementing all of these new potential components of the architecture. We wish to find the most parsimonious model that is powerful enough to account for the empirical phenomena in spatial cognition and general enough to be applied to important complex tasks. This is likely to depend on sustaining a close interaction between evaluations of models in complex task environments and the rigorous validation of mechanisms in carefully controlled laboratory paradigms to ensure that the accounts are psychologically valid and that they scale up to real-world tasks and domains.

Acknowledgments

The views expressed in this paper are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government. This research was sponsored by the Air Force Research Laboratory's (AFRL) Warfighter Readiness Research Division and by grants 05HE06COR and 10RH06COR from the Air Force Office of Scientific Research (AFOSR). Portions of this research have been presented at the Annual Meeting of the Cognitive Science Society (2006, 2007, 2009), the 17th Conference on Behavior Representation in Modeling and Simulation (2008), the International Conference on Cognitive Modeling (2009), and Spatial Cognition 2006. We appreciate comments from two anonymous reviewers and Dr. Wayne Gray during the editorial process.

References

- Andersen, R. A., Essick, G. K., & Siegel, R. M. (1985). The encoding of spatial location by posterior parietal neurons. *Science*, *230*, 456–458.
- Andersen, R. A., & Mountcastle, V. B. (1983). The influence of the angle of gaze upon the excitability of the light-sensitive neurons of the posterior parietal cortex. *Journal of Neuroscience*, *3*, 532–548.
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, J. R. (2007). *How can the mind exist in a physical universe?* New York: Oxford University Press.
- Attneave, F., & Curlee, T. E. (1983). Locational representation in imagery: A moving spot task. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 20–30.

- Baciu, M., Koenig, O., Vernier, M. P., Bedoin, N., Rubin, C., & Segebarth, C. (1999). Categorical and coordinate spatial relations: fMRI evidence for hemispheric specialization. *Neuroreport*, *10*, 1373–1378.
- Ball, J., Myers, C. W., Heiberg, A., Cooke, N. J., Matessa, M., & Freiman, M. (2009). The synthetic teammate project. *Proceedings of the 18th annual conference on behavior representation in modeling and simulation*. Sundance, UT: BRIMS.
- Barkowsky, T., & Freksa, C. (1997). Cognitive requirements on making and interpreting maps. In S. Hirtle & A. Frank (Eds.), *Spatial information theory: A theoretical basis for GIS* (pp. 347–361). Berlin: Springer.
- Bejar, I. I. (1990). A generative analysis of a three-dimensional spatial task. *Applied Psychological Measurement*, *14*, 237–245.
- Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition*, *75*(1), 1–28.
- Brooks, L. R. (1968). Spatial and verbal components in the act of recall. *Canadian Journal of Psychology*, *22*, 349–368.
- Burgess, N. (2006). Spatial memory: How egocentric and allocentric combine. *Trends in Cognitive Sciences*, *10*(12), 551–557.
- Burgess, N., Donnett, J. G., Jeffery, K. J., & O’Keefe, J. (1999). Robotic and neuronal simulation of the hippocampus and rat navigation. In N. Burgess & K. J. Jeffery (Eds.), *The hippocampal and parietal foundations of spatial cognition* (pp. 149–166). New York: Oxford University Press.
- Daprati, E., & Gentilucci, M. (1997). Grasping an illusion. *Neuropsychologia*, *35*, 1577–1582.
- Dimperio, E., Gunzelmann, G., & Harris, J. (2008). An initial evaluation of a cognitive model of UAV reconnaissance. In J. Hansberger (Ed.), *Proceedings of the 17th conference on behavior representation in modeling and simulation* (pp. 165–173). Orlando, FL: Simulation Interoperability Standards Organization.
- Diwadkar, V. A., Carpenter, P. A., & Just, M. A. (2000). Collaborative activity between parietal and dorsolateral prefrontal cortex in dynamic spatial working memory revealed by fMRI. *NeuroImage*, *12*, 85–99.
- Fincham, J. M., Carter, C. S., van Veen, V., Stenger, V. A., & Anderson, J. R. (2002). Neural mechanisms of planning: A computational analysis using event-related fMRI. *Proceedings of the National Academy of Sciences*, *99*(5), 3346–3351.
- Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General*, *119*(1), 63–76.
- Gevers, W., Verguts, T., Reynvoet, B., Caessens, B., & Fias, W. (2006). Numbers and space: A computational model of the SNARC effect. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(1), 32–44.
- Glasgow, J. I., & Papadias, D. (1992). Computational imagery. *Cognitive Science*, *17*(3), 355–394.
- Gunzelmann, G. (2008). Strategy generalization across orientation tasks: Testing a computational cognitive model. *Cognitive Science*, *32*(5), 835–861.
- Gunzelmann, G., & Anderson, J. R. (2006). Location matters: Why target location impacts performance in orientation tasks. *Memory & Cognition*, *34*(1), 41–59.
- Gunzelmann, G., & Lyon, D. R. (2006). Qualitative and quantitative reasoning and instance-based learning in spatial orientation. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th annual meeting of the cognitive science society* (pp. 303–308). Mahwah, NJ: Lawrence Erlbaum Associates.
- Gunzelmann, G., & Lyon, D. R. (2007). Mechanisms of human spatial competence. In T. Barkowsky, M. Knauff, G. Ligozat, & D. Montello (Eds.), *Spatial cognition V: Reasoning, action, interaction. Lecture notes in Artificial Intelligence #4387* (pp. 288–307). Berlin: Springer-Verlag.
- Gunzelmann, G., & Mohebbi, R. (2010). Spatial encoding in briefly presented schematic displays [Abstract]. In *Poster book: Association for Psychological Science 22nd annual convention*. (pp. 179). Washington, DC: Association for Psychological Science.
- Harrison, A. M., & Schunn, C. D. (2003). ACT-R/S: Look Ma, no ‘‘cognitive map’’! In F. Detje, D. Doerner, & H. Schaub (Eds.), *Proceedings of the fifth international conference on cognitive modeling* (pp. 129–134). Bamberg, Germany: Universitats-Verlag Bamberg.

- Hartley, T., Trinkler, I., & Burgess, N. (2004). Geometric determinants of human spatial memory. *Cognition*, 94(1), 39–75.
- Hegarty, M. (1992). Mental animation: Inferring motion from static diagrams of mechanical systems. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18(5), 1084–1102.
- Huttenlocher, J., & Presson, C. C. (1979). The coding and transformation of spatial information. *Cognitive Psychology*, 11, 375–394.
- Kerr, N. H. (1987). Locational representation in imagery: The third dimension. *Memory and Cognition*, 15, 521–530.
- Kerr, N. H. (1993). Rate of imagery processing in two versus three dimensions. *Memory and Cognition*, 21, 467–476.
- Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In C. Freksa, C. Habel & K. F. Wender (Eds.), *Spatial cognition – An interdisciplinary approach to representation and processing of spatial knowledge (Lecture Notes in Artificial Intelligence 1404)* (pp. 1–17). Berlin: Springer-Verlag.
- Klatzky, R. L., Wu, B., & Stetten, G. (2008). Spatial representation from perception and cognitive mediation: The case of ultrasound. *Current Directions in Psychological Science*, 17, 359–364.
- Kosslyn, S. M. (1994). *Image and brain*. Cambridge, MA: MIT Press.
- Kosslyn, S. M., Sokolov, M. A., & Chen, J. C. (1989). The lateralization of BRIAN: A computational theory and model of visual hemispheric specialization. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing comes of age: The impact of Herbert Simon* (pp. 3–29). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. New York: Oxford University Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.
- Levine, M., Jankovic, I. N., & Palij, M. (1982). Principles of spatial problem solving. *Journal of Experimental Psychology: General*, 111, 157–175.
- Levine, M., Marchom, I., & Hanley, G. L. (1984). The placement and misplacement of you-are-here maps. *Environment & Behavior*, 16, 139–157.
- Lyon, D. (2009). Crowding in the mental image: Evidence for a vision-like representation? [Abstract]. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st annual conference of the cognitive science society* (p. 951). Austin, TX: Cognitive Science Society.
- Lyon, D. R., & Gunzelmann, G. (2009). Visualizing egocentric paths: A computational model. In A. Howes, D. Peebles, & R. Cooper (Eds.), *Proceedings of the ninth international conference on cognitive modeling*. Manchester, UK: University of Manchester.
- Lyon, D. R., Gunzelmann, G., & Gluck, K. A. (2006). Virtual travel does not enhance spatial working memory for landmark-free paths. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th annual meeting of the cognitive science society* (pp. 2550). Mahwah, NJ: Lawrence Erlbaum Associates.
- Lyon, D. R., Gunzelmann, G., & Gluck, K. A. (2007). Visualizing egocentric vs. exocentric path descriptions. In D. S. McNamara & G. Trafton (Eds.), *Proceedings of the 29th annual meeting of the cognitive science society* (p. 1811). Mahwah, NJ: Lawrence Erlbaum Associates.
- Lyon, D. R., Gunzelmann, G., & Gluck, K. A. (2008). A computational model of spatial visualization capacity. *Cognitive Psychology*, 57, 122–152.
- Malinowski, J. C., & Gillespie, W. T. (2001). Individual differences in performance on a large-scale, real-world wayfinding task. *Journal of Environmental Psychology*, 21, 73–82.
- Mandler, J. M. (1992). How to build a baby: II. *Conceptual Primitives*. *Psychological Review*, 99(4), 587–604.
- Milner, A. D., & Goodale, M. A. (1993). Visual pathways to perception and action. *Progress in Brain Research*, 95, 317–337.
- Mou, W., & McNamara, T. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 162–170.

- Mulliken, G. H., Musallam, S., & Andersen, R. A. (2008). Decoding trajectories from posterior parietal cortex ensembles. *Journal of Neuroscience*, 28(48), 12913–12926.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
- O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely moving rat. *Brain Research*, 34, 171–174.
- O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford, England: Oxford University Press.
- Péruch, P., & Lapin, E. A. (1993). Route knowledge in different spatial frames of reference. *Acta Psychologica*, 84, 253–269.
- Rodríguez, F., López, J. C., Vargas, J. P., Broglio, C., Gómez, Y., & Salas, C. (2002). Spatial memory and hippocampal pallium through vertebrae evolution: Insights from reptiles and teleost fish. *Brain Research Bulletin*, 57(3/4), 499–503.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three dimensional objects. *Science*, 171, 701–703.
- Sholl, M. J., & Nolin, T. L. (1997). Orientation specificity in representations of place. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(6), 1494–1507.
- Smith, M., & Crawford, J. (2005). Distributed population mechanism for the 3-D oculomotor reference frame transformation. *Journal of Neurophysiology*, 93, 1742–1761.
- Touretzky, D. S., & Redish, A. D. (1996). Theory of rodent navigation based on interacting representations of space. *Hippocampus*, 6, 247–270.
- Trickett, S. B., & Trafton, J. G. (2007). “What if...”: The use of conceptual simulations in scientific reasoning. *Cognitive Science*, 31(5), 843–875.
- Xing, J., & Andersen, R. A. (2000). Models of the posterior parietal cortex which perform multimodal integration and represent space in several different coordinate frames. *Journal of Cognitive Neuroscience*, 12(4), 601–614.
- Zipser, D., & Andersen, R. A. (1988). A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature*, 331, 679–684.
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2), 162–185.