5

Thinking Outside the Brain: Spatial Indices to Visual and Linguistic Information

MICHAEL J. SPIVEY
DANIEL C. RICHARDSON
STANKA A. FITNEVA

In vision research, accumulating evidence suggests that the coherence of our visual experience involves not only internal representations in the brain but also the external visual environment itself. In this chapter, we discuss a collection of eye-movement experiments that lend further support for this important role of the external visual environment in visual imagery, in visual memory, as well as in linguistic memory and even in naturalistic conversation and insight problem solving. We argue that eye fixations serve as the cognitive liaisons (or "spatial indices" or "pointers") between internal and external objects and events. Essentially, the visual environment can be treated as an additional memory database, with eye movements being the most typical method for accessing such data. The spatial indices to which eye movements interface appear to be used not just for organizing perceptual-motor routines but also for organizing relatively high-level cognitive processes. These findings point to an externalist philosophy of mind, in which the concept of mental activity is not solely defined over neural states, but also includes peripheral bodily states, as well as objects and events in the surrounding environment.

INTRODUCTION

It just might be that your mind is bigger than your brain. Not because you have an ethereal soul that influences your brain via the pineal gland, as proposed by Descartes, but because your external physical environment contains information that you can *perceptually* access as quickly and directly as you can *cognitively* access information from internal memory. One might even say, what is in your immediate physical environment is "part of what you know." For example, do you know what time it is? If looking at your wristwatch is about as quick as (perhaps quicker than) recalling from memory what time it was 30 seconds ago when you last looked at your watch and involves functionally quite similar processes (i.e., content-addressable memory), then perhaps both processes can constitute "knowing the time."

In this chapter, we walk through a range of experimental demonstrations of ways in which people tend to rely on the external environment to store information for them rather than storing it all in their brains. On the surface, the phenomenon that we report—use of spatial indices, or deictic pointers—may appear intriguing, but not necessarily revolutionary. At a deeper level, however, this constellation of findings hints at the potential upheaval of some very old and fundamental assumptions in cognitive science: a mindset that philosophers have called *internalism* (Segal, 2001; see also Putnam, 1975).

INTERNALISM AND EXTERNALISM

Internalism holds that the contents of the mind at any point in time can be fully accounted for by a description of the state of the brain. While a full description of the state of the brain is, of course, impossible with current technology, it is noteworthy that an internalist account of mental content rules out any need for reference to the organism's environment in this description of mental content. Thus, although philosophy will certainly never be able to provide a full account of mental content (since it will not be the field that produces a full description of a brain-state), an internalist philosopher will at least tell us where not to look for one. The environment contains stimuli that influence the organism, and the environment undergoes changes due to that organism's actions, but the environment is not part of that organism's mind (cf. Newell, 1990). According to internalism, it is separate. This internalist conception of mental states seems intuitively obvious, but such intuitions can be severely troubled in the case of the "Twin Earth" thought experiments proposed by Putnam (1975) and others.[1]

Much of the debate between externalism and internalism employs such Twin Earth thought experiments to test for a relatively static inclusion of the environment in determining the truth value of belief states (e.g., Fodor, 1980;

Wilson, 1994; Segal, 2001). However, a recent version of externalism that focuses rigorously on the immediate participatory role of the environment (in addition to brain and body, of course) in constructing mind has been called *active externalism* (Clark & Chalmers, 1998). This perspective marshals demonstrations from self-organized artificial intelligence research (Beer, 1989; Brooks, 1991), demonstrations from dynamical systems theory (Kelso, 1995; Thelen & Smith, 1994), observations of situated action (Greeno, 1998; Suchman, 1987), of collective action (Hutchins, 1995), and collective intelligence (Lévy, 1997), as well as thought experiments (Wilson, 1994), to argue for the importance of "cognitive properties of systems that are larger than an individual" (Hutchins, 1995; for a review, see Clark, 2001). Haugeland (1995) has dubbed it the "embodied and embedded" account of mind. Not only does the central nervous system's *embodiment* in a particular vehicle with particular sensors and effectors pose as a crucial expansion of the old-fashioned concept of mind-as-just-brain, but that brain-body dyad's *embedding* in a particular environment makes the whole system a richly interactive brain-body-environment triad.

Although the case for an embodied and embedded mind is compelling for some (cf. McClamrock, 1995; Ross, 1997), with its robot implementations, computer simulations, natural observations, and thought experiments, the one thing this literature has been short on is controlled laboratory experimentation. Importantly, as some of the most devoted (and sometimes unwitting) customers of the mind-as-just-brain assumption, cognitive psychologists have found it easy to ignore this new embodied and embedded perspective precisely because it has lacked controlled experimental results. Perhaps it should not be surprising that so many people accept internalism, at first glance, as a foregone conclusion. Lakoff (1987, 1997) provides a number of naturally occurring linguistic examples of people taking for granted the conceptual metaphor "the mind is a container." If the mind is a container, then it must have discrete boundaries delineating what is "inside" and what is "outside," and in the case of the human mind, the human skull seems to be the best box for the job. Indeed, the intuition that a complete description of mental activity will come solely from properties of the organism's central nervous system is so powerful that it has successfully resisted quite a few attempts to dispel it. Not only has the internalist mindset survived the recent critiques of contemporary philosophers such as Putnam (1975), Haugeland (1995), Dreyfus (1996), and Clark (2001), but decades ago it survived Dewey (1896), Le Bon (1916), Ryle (1949), Merleau-Ponty (1962), and Gibson (1966), just to name a few.

INTERNALISM IN PSYCHOLOGY

As one example of manifestation of this internalist mindset in psychology, a popular framework for theories of visual perception, the "spatiotopic fusion hypothesis" critiqued by Irwin (1992), assumes that the successive retinal images that are acquired in between saccadic eye movements are metrically combined to construct and store an internal representation of the external visual world inside the brain (cf. Marr, 1980). This assumption of an "internal screen" (O'Regan, 1992) on which is projected an image of the external visual world for the inspection of some central executive has—despite its obvious homunculus problems—driven a great deal of research in visual psychophysics, visual neuroscience, visual cognition, as well as computer vision. A number of theories have been proposed to account for the problem of how such noisy, illusion-prone, ballistic optical devices as the eyes can avail the construction of a contiguous, metrically accurate, internally represented 3-D model of the visual environment (for a review, see O'Regan, 1992). O'Regan (1992) notes that over the years a number of researchers have proposed not to *solve* this problem, but instead to *dissolve* it (e.g., Gibson, 1950; Haber, 1983; Turvey, 1977; see also Bridgeman, van der Heijden, & Velichkovsky, 1994). If we do *not* actually have a contiguous, metrically accurate, internally represented 3-D model of the visual environment in our brains, then there is no need to figure out how our eyes and visual systems build one (and perhaps computer vision should stop trying things that way too, cf. Ballard, 1989). O'Regan (1992) suggests that, rather than visual perception being a passive process of accumulating retinal images from which to build an internal 3-D model, "seeing constitutes an active process of probing the external environment as though it were a continuously available external memory…if we so much as faintly ask ourselves some question about the environment, an answer is immediately provided by the sensory information on the retina, possibly rendered available by an eye movement." (p. 484). Not unlike the externalist philosophers, O'Regan and Noë (2001) claim that "activity in internal representations does not generate the experience of seeing. The outside world serves as its own, external, representation." (p. 939).

If it is the case that rather little of the external visual environment is actually internalized, then, logically, unexpected changes in the visual environment should go unnoticed. For example, one should be able to change the color, location, and other properties as well—even the very presence—of large objects in a complex scene and have it frequently go unnoticed. This, however, clashes sharply with our intuition that we are continuously aware of the complete contents of the visual scene laid out before our eyes. This logical, but counterintuitive, prediction of O'Regan's (1992) brand of visual externalism led directly to the recent cottage industry of change blindness research (for a review, see Simons, 2000).

Abrupt changes in a display will typically attract attention immediately if they take place during an uninterrupted fixation (e.g., Yantis & Jonides, 1990). However, it turns out that a range of minor ocular and attentional disturbances are sufficient to mask this ability. If the image flickers briefly during the scene change, participants rarely notice the change (Rensink, O'Regan, & Clark, 1997). If the scene is briefly overlaid by a few blobs, or "mudsplashes," flashed on the screen during the change—without occluding the region that changes—participants rarely detect the change (O'Regan, Rensink, & Clark, 1999). If the scene change takes place during a saccade, it is likely to go unnoticed (McConkie & Currie, 1996). And if the scene change takes place during a blink, it is rarely detected (O'Regan, Deubel, Clark, and Rensink, 2000). In fact, even if the eyes were fixating within a degree of the object to be changed, right before the blink, when the eyelids open back up, and the object has changed, participants notice the change only 40% of the time (O'Regan et al. 2000).

Change blindness also works in dynamic real-world scenarios. For example, inspired by a gag from the old Candid Camera television show, Simons and Levin (1998) had a confederate accost passersby on the Cornell University campus and ask for directions on a map. During the conversation, two young men carrying a door walked between the confederate and the passerby. The confederate and one of the door carriers exchanged places, and the door carrier took up the conversation as if nothing unusual had happened. Only about 50% of the time did the passerby notice that the person he was talking to had changed!

The dramatic effects observed in change blindness experiments provide compelling support for an externalist claim that the locus of perception is as much in the environment itself as it is in the organism interacting with that environment (e.g., Noë, Pessoa, & Thompson, 2000). This is not to say that nothing about the environment is stored internally. As the reports show, many of these scene changes are detected as much as 50% of the time (e.g., Hollingworth & Henderson, 2002), and implicit measures often reveal greater change detection than that seen with explicit verbal protocol (Hayhoe, 2000; Hollingworth, Williams, & Henderson, 2001). Thus, certain attended aspects of the scene are stored in internal memory, and when those aspects are altered in the scene, the mismatch between internal and external representations is detected at least somewhere in the visual system. This point will become especially important in our later discussion of exactly how visual properties that are *not* stored internally can be accurately indexed and accessed from the external environment, via an internally stored label for the index.

THINKING OUTSIDE THE BRAIN

If the external environment is even just occasionally relied upon as a source of visual memory, one can ask whether it is possible, in those circumstances, to purposefully take advantage of and optimize that external memory? In fact, Kirsh (1995; see also Kirsh & Maglio, 1994) cites numerous real-world examples of people doing exactly that. Kirsh (1995) makes the observation that we physically “jig” our environment with physical constraints that structure and optimize our interaction with it. For example, when moving into a new house, deciding what utensils, dishes, and pans to put in which kitchen drawers and cabinets is often done with imagined plans of when and where the various accoutrements will be needed during cooking and cleaning. When arranging one’s office desk, the computer, the telephone, the stapler, the tape dispenser, “in” and “out” boxes, etc. are all placed in locations that the user expects will maximize their coordinated and sequential use. Similarly, a colleague of ours, who worries that he paces too much while lecturing, deliberately places chairs, overhead projectors, etc, blocking the way of the most natural pacing routes. These are all examples of physically jigging one’s environment so that accessibility and restriction of various objects and actions is optimally timed. This means that information is being built into the environment and thus that information will not always need to be cognitively represented. In a way, a properly jigged work environment can be counted on to “do some of the thinking for you.”

Additionally, Kirsh (1995) notes that one way of *informationally* jigging an environment is to “seed” it with attention-getting cues. For example, to help oneself remember to bring a book to school, one might place the book next to the front door inside one’s house. Also, many people have specific wall-hooks or dishes near the front door inside their house where they keep their keys. Thus, the knowledge that one’s keys will be needed when leaving the house need not be an active component of the cognitive plan to go to the store because that knowledge is built into the environment to become perceptually salient at just the right time. In these kinds of circumstances, we’ve externalized (offloaded, if you will) information onto our environment, thereby freeing up internal processing capacity, and thus certain crucial bits of information that are necessary for complex behavior are provided not by neural-based memory representations but by the environment itself on a need-to-know basis.

A concrete example of this kind of phenomena comes from a recent study by Grant and Spivey (in press), in which participants’ eye movements were recorded while they attempted to solve a diagram-based version of Duncker’s (1935) classic tumor-and-lasers problem. The schematic diagram was simply a filled oval, representing the tumor, with a circumscribing oval representing the stomach lining (which must not be injured). Nothing else in Duncker’s problem description was depicted in the schematic diagram. As this problem is a very difficult insight problem, only a third of the participants solved it without needing hints. Although the eye-movement patterns were very similar for successful and unsuccessful solvers, one difference stood out. Successful solvers tended to look at the stomach lining more than unsuccessful solvers. We then used this observation to try to influence participants’ cognitive performance by manipulating the perceptual salience of components of the diagram.

In a second experiment, the schematic diagram was animated (with a single pixel increase in diameter pulsating at 3 Hz) to subtly increase the perceptual salience of either the stomach lining, or the tumor. A control condition had no animation. In the control and pulsating tumor conditions, one third of the participants solved the problem without hints, as expected. However, in the pulsating stomach lining condition, *two* thirds of the participants solved the problem without hints! Grant and Spivey (in press) hypothesized that the increased perceptual salience of the stomach lining helped elicit patterns of eye movements and attention that were conducive to developing a *perceptual simulation* (Barsalou, 1999) of the correct solution, which involved multiple weak lasers from different locations converging on the tumor. Essentially, Grant and Spivey (in press) "jigged" the environment, with a subtle manipulation in perceptual salience, such that a creative cognitive inference was facilitated. Thus, one might say, having an intelligent environment is just as important as having an intelligent brain.

POINTERS IN SPACE

In the next sections, we will outline several examples of the bidirectional interaction between the environment and cognition and discuss examples of salient external information triggering internal processes, as well as internally generated information being linked back to external objects and locations. In fact, we humans have quite a penchant for externalizing our internal information. Of course, we communicate to others by linguistic means (speaking and writing) as well as nonlinguistic means (hand gestures, facial expressions, prosody, etc.). But we also externalize internal information purely for our own benefit. We recite phone numbers out loud to ourselves so that the environment can deliver the information to our ears, mimicking the phonological loop. We make lists of things to do and of groceries to buy. Some of us talk to ourselves. Some of us even write on our hands. We write down appointments on calendars. We occasionally point a finger at an object when we're silently reminding ourselves to do something with it. And sometimes when we imagine things, our eyes virtually paint our imagery on the world.

In a recent headband-mounted eye-tracking experiment, Spivey and Geng (2001, Experiment 1; see also Spivey, Tyler, Richardson, & Young, 2000) recorded participants' eye movements while they listened to spoken descriptions of spatiotemporally dynamic scenes and faced a large white projection screen that took up most of their visual field. For example, "Imagine that you are standing across the street from a 40-story apartment building. At the bottom there is a doorman in blue. *On the 10th floor, a woman is hanging her laundry out the window. On the 29th floor, two kids are sitting on the fire escape smoking cigarettes. On the very top floor, two people are screaming.*"

While listening to the italicized portion of this passage, participants made reliably more upward saccades than in any other direction. Corresponding biases in spontaneous saccade directions were also observed for a downward story, as well as for leftward and rightward stories. Thus, while looking at ostensibly nothing, listeners' eyes were doing something similar to what they would have done if the scene being described were actually right there before them. Instead of relying solely on an internal "visuospatial sketchpad" (Baddeley, 1986) on which to illustrate their mental model of the scene being described, participants also recruited the external environment as an additional canvas on which to depict the spatial layout of the imagined scene.

Although eye movements may not be required for vivid imagery (Hale & Simpson, 1970; but cf. Ruggieri, 1999), it does appear that they often naturally accompany it (e.g., Antrobus, Antrobus, & Singer, 1964; Brandt & Stark, 1997; Demarais & Cohen, 1998; Neisser, 1967; see also Hebb, 1968). But what is it that the eyes are trying to do in these circumstances? Obviously, it is not the case that the eyes themselves can actually externally record this internal information. When the eyes move upward from the imagined 10th floor of the apartment building to the imagined 29th floor, no physical mark is left behind on the external location in the environment that was proxying for that 10th floor.

Rather than a physical mark, perhaps what they "leave behind" is a deictic pointer, or spatial index. According to Ballard, Hayhoe, Pook, and Rao (1997; see also Pylyshyn, 1989, 2001), deictic pointers can be used in visuomotor routines to conserve the use of working memory. Instead of storing all the detailed properties of an object internally, one can simply store an address (or pointer) for the object's location in the environment, along with some labeling information, and access those properties perceptually when they are needed.

In the case of Spivey & Geng's (2001) eye movements during imagery, a few pointers allocated on a blank projection screen will obviously not reference any external visual properties, but they can still provide perceptual-motor information about the relative spatial locations of the *internal* content associated with the pointers. If one is initially thinking about x (e.g., the 10th floor) and then transitions to thinking about y (e.g., the 29th floor), then storing in working memory the relation *above (y,x)* may not be necessary if the eye movements, and their allocation of spatial indices, have embodied that spatial relationship already (cf. Pylyshyn, 1989). In this way, a "low-level" motor process, such as eye movements, can actually do some of the work involved in the "high-level" cognitive act of visual imagery.

Although it is the address in the pointer that allows one to rely on the external environment to store information, the *label* for the pointer is also a very important ingredient in this recipe. The internally represented label could be something as simple as "target," or it could be rich information such as "the doorman in blue at the bottom of the 40-story apartment building." A pointer must have some internal content attached to it so that one can know when and why to use it. Otherwise, you wind up like Ernie, on Sesame Street, trying to explain to Bert why he has a string tied around his finger when he can't remember what it was that the string was supposed to remind him about. A pointer with no internal information attached to it is useless.

POINTERS TO OBJECTS

To illustrate the use of such spatial indices in visual attention, Pylyshyn introduced a multiple object tracking task (e.g., Pylyshyn & Storm, 1988; Scholl & Pylyshyn, 1999). In this task, participants view an initial display of indistinguishable discs or squares, of which a subset flash several times to indicate that they are the targets. Then all the objects begin to move in pseudorandom directions across the screen, and the participant's task is to "keep track" of the handful of target discs while maintaining central fixation. Participants can successfully track up to about four or five such targets, but if there are more than that, they begin to make errors (attributing targethood to nontarget objects). As participants must maintain central fixation throughout this task, these spatial indices are clearly being allocated and updated extrafoveally.

In another experimental paradigm that demonstrates the use of spatial indices in natural visuomotor processing, Ballard, Hayhoe, and Pelz (1995) recorded participants' eye movements during a block pattern copying task, with a model pattern, a resource of blocks, and a workspace in which to copy the model. In this kind of framework, eye position serves the function of allocating spatial pointers for working memory, in which a pointer stores an *address* in spatial coordinates along with little more than a *label* for when and why to use the pointer. For example, a pointer's address might be something like "the block just to the right of the top-leftmost block in the model," and its label might be "the block I am working on now." Thus, if the participant has just finished placing the previous block in the incomplete block pattern in the workspace, then this pointer can guide the eyes to this new block in the model block pattern in order to access and store its color. With the color of this block now stored internally, the eyes can then move to the resource space, containing many blocks of various colors, and search for a block of the same color. Once that new block is picked up, in order to put it in the appropriate location in the workspace, one needs to know its position relative to the other blocks in the incomplete block-pattern. As the pointer's address itself may make reference to blocks that have not yet been placed in the workspace, the eyes must once again call up this pointer allocated to "the block just to the right of the top-leftmost block in the model" and perceptually access its spatial relationships with the adjacent blocks. With this new information stored in working memory, the eyes can move down to the workspace for placement of the new block. The pointer with the label "the block I am working on now" must then delete its current address and find a new one elsewhere on the model block pattern, and begin the process all over again. This sequence of fixating the model, then the resource, then back to the model, before finally looking at the workspace for block placement was indeed the modal pattern of eye movements observed in Ballard et al.'s (1995) experiments.

But what happens if the external information referred to by these spatial indices changes? According to the framework, one should expect the person copying the block pattern not to notice when a block changes color, except under those circumstances where the process is at a stage where the visual property that's been changed is the one currently being stored in working memory. This is, indeed, exactly what happens (Hayhoe, 2000; Hayhoe, Bensinger, & Ballard, 1998). If a few deictic pointers have been allocated to particular objects or regions of space, and the current task calls upon the label of one of those pointers, the system will automatically seek the address associated with that pointer—fixate the indexed object or location—and perceptually access the external information at that address. If neither the pointer's label nor working memory contain information that conflict with this externally accessed information, then, naturally, any change that took place in that external information will go undetected. The newly accessed visual properties will be trusted as if they had been that way all along.

POINTERS TO ABSENT OBJECTS

Interestingly, this accessing of a pointer when its label is called upon is so automatic that it can even happen when the object to which it was originally allocated is no longer present at all. In Spivey and Geng's (2001) second experiment, they presented four different shapes of varying colors, tilted 15 degs leftward or rightward, in the four quadrants of the screen. Participants were instructed to look at the object in each quadrant, and then back to a central fixation cross. One of the four shapes then disappeared, and participants were asked to recall either its color or its direction of tilt. On as many as 50% of the trials, as they formulated their answer, participants spontaneously fixated the empty quadrant that used to contain the shape being queried—despite the fact that they could easily determine in peripheral vision that the object was no longer there. Participants rarely looked at the other remaining shapes. This is exactly what one should expect if observers are employing pointers to rely on the external world to store object properties in addition to what is stored in the pointers' labels themselves and in working memory. The task calls upon the shape's name (e.g., "diamond"), which activates the pointer with that label, and queries a property of that shape (e.g., "color"). If the pointer's label does not include the attribute (e.g., "green"), then the pointer's address to the external environment is the next obvious resource. A relatively automatic eye movement to that address verifies that the queried information is absent from the external environment. At this point, internal working memory is the only resort. On the trials where participants fixated the empty quadrant, as well as on the trials where they did not fixate it, the same information resource, internal working memory, is used to answer the question. Thus, one should actually *not* expect a difference in memory accuracy between trials in which the empty quadrant was fixated and those in which it was not. And that is, indeed, what Spivey and Geng (2001, Experiment 2) found.

Spivey and Geng (2001) concluded that, since there is no improvement of memory, the eye movement to the empty quadrant does not appear to be an attempt to recruit visual surroundings in order to encourage a context-dependent improvement of memory. Nor is it a deliberate, strategic, attempt to answer the question by looking at the queried object because participants can easily tell from peripheral vision, as well as from previous trials, that the object is not there. Rather, the eye movement to the empty quadrant is an automatic attempt by an embodied working memory system to access the contents of a pointer's address in the external environment. Just as in the change blindness studies, this embodied working memory system does not know that the content in that external location has been removed until it accesses the pointer with that address. Although it is possible to attend to and access these pointers without eye movements when the task instructions require it (Pylyshyn & Storm, 1988), a wide range of research indicates that eye movements naturally follow such allocations of attention (e.g., Ballard et al., 1997; Corbetta & Shulman, 1999; Henderson, 1993; Hoffman, 1998; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995).

HOLLYWOOD SQUARES

It might not be too surprising that the embodied working memory system, relying on pointers that reference visual objects, elicits eye movements to the addresses of those pointers when the system is trying to access memory of *visual* properties. But what about when the queried content associated with that pointer is not visual, but auditory? In a series of experiments, affectionately referred to as Hollywood Squares because the task somewhat resembles the television game show, Richardson and Spivey (2000) presented four talking heads in sequence, in the four quadrants of the screen, each reciting an arbitrary fact and then disappearing (e.g., "Shakespeare's first plays were historical dramas. His last play was The Tempest."). With the display completely blank except for the lines delineating the four empty quadrants, a voice from the computer delivered a statement concerning one of the four recited facts, and participants were instructed to verify the statement as true or false (e.g., "Shakespeare's first play was The Tempest."). See Figure 5.1.

While formulating their answer, participants were twice as likely to fixate the quadrant that previously contained the talking head that had recited the relevant fact than any other quadrant. Despite the fact that the queried information was delivered auditorily, and therefore cannot possibly be visually accessed via a fixation, something about that location drew eye movements during recall. Richardson and Spivey (2000) suggested that spatial indices had been allocated to the four quadrants to aid in sorting and separating the events

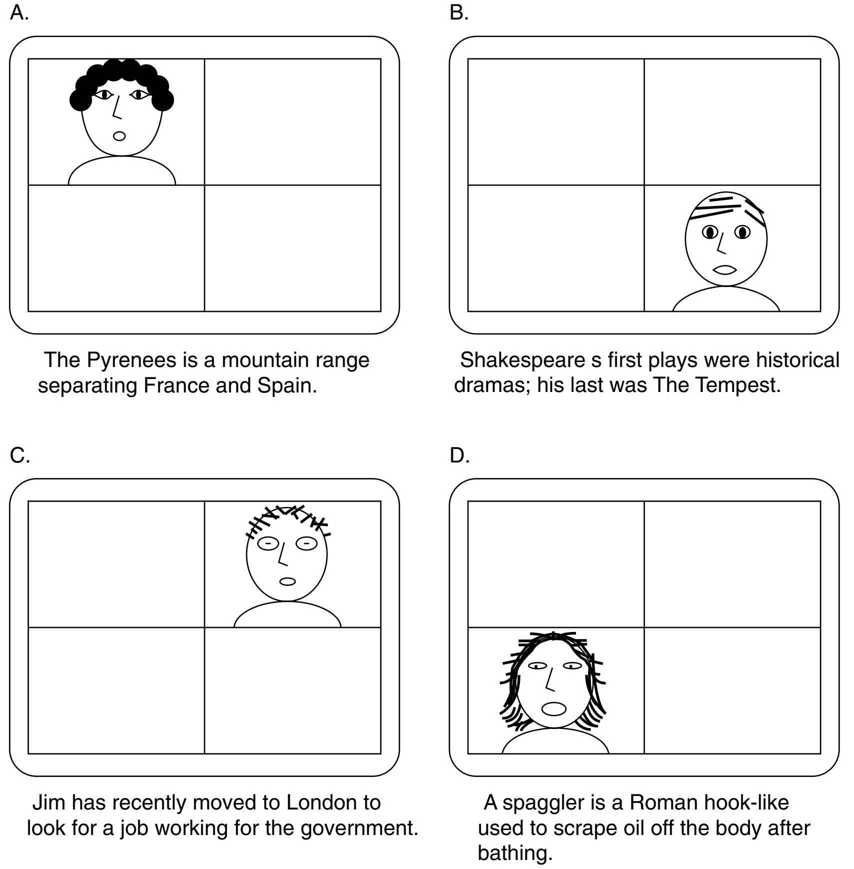

FIGURE 5.1. In the 'Hollywood Squares' experiment, participants looked at talking heads that delivered arbitrary facts (A-D). At the end of the trial, with the four quadrants empty, the computer delivered a spoken statement that the participant verified as true or false (e.g., "Shakespeare's first play was The Tempest.").

that took place in them. Thus, when the label of one of those pointers was called upon (e.g., "Shakespeare"), attempts to access the relevant information were made both from the pointer's address in the external environment and from internal working memory. As before with Spivey and Geng's (2001) findings, since the external environment no longer contained the queried information, internal working memory was the sole determinant of memory accuracy. Therefore, verification accuracy was the same on trials that did have fixations of the queried quadrant as on trials that did not.

Richardson and Spivey (2000, Experiment 2) replicated these results using four identical spinning crosses in the quadrants during delivery of the facts, instead of the talking heads. Participants seemed perfectly happy to allocate pointers to the four facts in those four locations, even when spatial location was the only visual property that distinguished the pointers. Moreover, in the "tracking" condition (Richardson & Spivey, 2000, Experiment 5), participants viewed the grid through a virtual window in the center of the screen. Behind this mask, the grid moved, bringing a quadrant to the center of the screen for fact presentation. Then, during the question phase, the mask was removed. Even in this case, when the spinning crosses had all been viewed in the center of the computer screen, and the relative locations of the quadrants implied by translation, participants continued to treat the quadrant associated with the queried fact as conspicuously worthy of overt attention. In fact, even if the crosses appear in empty squares which move around the screen following fact delivery, participants spontaneously fixate the square associated with the fact being verified (Kirkham & Richardson, submitted, Experiment 1). Thus, once applied, a deictic pointer— even one that attempts to index auditorily delivered semantic information—can dynamically follow the moving object to which it was allocated (e.g., Scholl & Pylyshyn, 1999; see also Tipper & Behrmann, 1996).

It actually should not be surprising that an embodied working memory system using deictic pointers would attempt to index information from events that are over and done with. The pointer doesn't "know" that the sought-after information at its address is long gone precisely because it has offloaded that knowledge onto the environment —it wouldn't be a pointer otherwise. These findings demonstrate the robustness and automaticity with which spatial indices are relied upon in order to employ the body's environment as sort of notice board of "virtual post-it notes" that complement our internal memory.

ON CONTEXT-DEPENDENT MEMORY

Some researchers acquainted with context-dependent memory results have expressed bemusement—in some cases, even disappointment—that memory accuracy in these studies was not improved on the trials where the participant fixated the quadrant that had been associated with the fact being verified. There might, in principle, be the possibility for visual context helping a pattern completion process for internally accessing a memory in this task (e.g., Eich, 1980). However, the nondistinctive visual contexts at the participants' disposal in this display would clearly not have been conducive to such an effect.[2]

Figure 5.2 demonstrates why it should not be surprising at all that memory was not improved on trials where the participant fixated the critical quadrant, compared to trials in which he/she did not. As the visual input is extremely similar regardless of which quadrant the participant looks at, it seems unlikely that the view of any one quadrant would provide sufficiently unique contextual memory cues compared to any other quadrant.

Importantly, this makes it all the more striking that we continue to observe participants producing this effect in study after study. Participants get no memory benefit from looking at the empty quadrant; indeed they shouldn't ex-

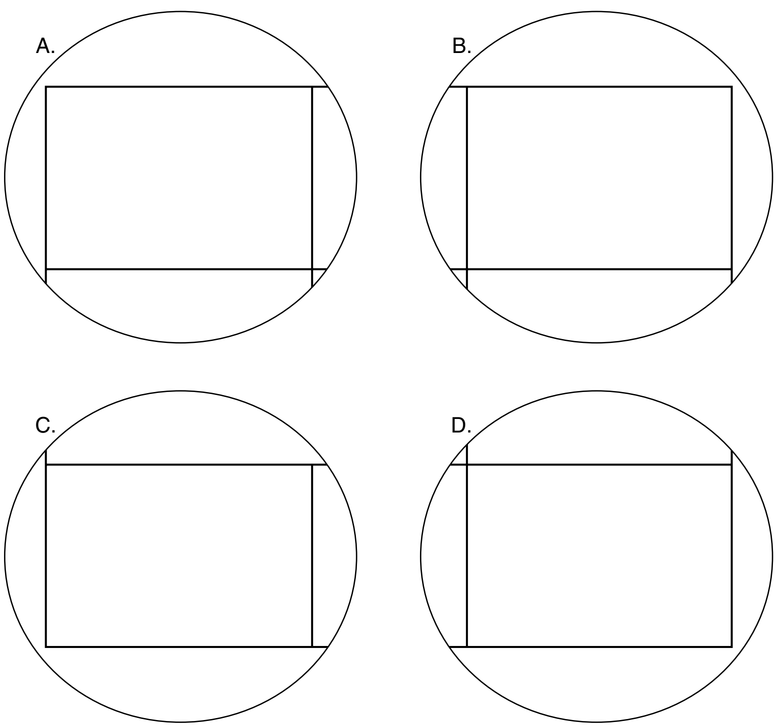

FIGURE 5.2. Panels A-D are schematic examples of the images on the central retina when viewing the empty quadrants during the verification phase of the Hollywood Squares task. Rather little distinguishing information is available for assisting a context-dependent memory effect.

pect to since there is little if any distinguishing contextual input provided from doing so. And yet the spatial indexing system is so accustomed to information in the physical environment being relatively stable from moment to moment that it erroneously treats even ephemeral information sources, such as spoken utterances, as if they are “still out there,” waiting for their content to addressed. This is externalism.

POINTERS IN INFANCY

Is externalism something we come to only late in life? In other words, is the type of spatial indexing we have discussed only a feature of the mature, literate (perhaps even computer-literate) adult brain? Or is the tendency to treat objects and spatial locations as though they were “keepers of information” evinced at early stages of development? Leslie, Scholl and colleagues have argued that theories of adult spatial indexing can be related to infants’ emerging concept of “‘objecthood” (e.g., Leslie, Xu, Tremoulet, & Scholl, 1998; Scholl & Leslie, 1999). Spatial indexing is believed to play an important role in infants’ developing ability to individuate and enumerate items in the world. Thus, theories of spatial indexing have typically been employed in developmental research primarily as an explanation for infants’ developing object knowledge. Surprisingly, though, this hypothetical mechanism has not itself been directly studied until recently.

In their Experiment 2, Kirkham and Richardson (submitted) familiarized 6-month-olds with animated pictures of toys that danced inside square frames in time to their distinctive sounds. As in the Hollywood Squares experiments, the sounds always came from the same pair of stereo speakers located on the left and right sides, equidistant from the center of the display. In the test phase, one of the sounds was played while the two squares remained empty, and the infants’ eye movements were recorded. The infants spent more time gazing at the empty square frame that had previously contained the toy associated with that sound. This first result with the infants is reminiscent of findings by Marshall Haith (e.g., Haith, Hazan, & Goodman, 1988; Haith, Wentworth, & Canfield, 1993), and may be essentially an expectation-based gaze.

However, Kirkham and Richardson’s (submitted) Experiment 3 shows something distinctly more sophisticated. Here, the square frames traveled to new locations on the screen right before the test sound was presented. Once again, the 6-month-olds spent more time gazing at the empty square frame that had previously contained the toy associated with that sound.. At this age, it has been found that infants are only just beginning to move from retinocentric to egocentric reference frames for representing locations in space (Gilmore & Johnson, 1998). However, even at this early point in development, infants already have the ability to index a sound to a location in space, and even to allow that location to be spatially updated—just like the updating of pointers' addresses in experiments with adults (Kirkham & Richardson, submitted, Experiment 1; Scholl & Pylyshyn, 1999; see also Wang, 1999).

POINTERS IN READING AND CONVERSATION

This ability to use space to organize components of sensory input is relevant not just for memory, attention, and action, but also for language processing as well. Spatial organization is used in many aspects of language processing (cf., Bryant, 1997; Chaterjee, 2001; Glenberg & Robertson, 1999; Richardson, Spivey, Barsalou, & McRae, in press), and particularly explicitly in sign language (e.g., Emmorey, 2002). For example, signers of American Sign Language will use a *signing space*, of perhaps 2 feet in diameter in front of them, and after discourse entities have been introduced and assigned to specific locations, they can be deictically referred to by pointing at the appropriate location in the signing space. In fact, transitive events with a subject and a direct object can often be communicated by simply signing the transitive verb in a fashion such that the hand's trajectory begins at the location assigned to the entity acting as subject and ends at the location assigned to the entity acting as direct object (Emmorey, 2001). No explicit reference or pointing to the entities is necessary. Thus, during sign-language production and comprehension, locations in space must be kept track of as place holders for the various objects and entities in the discourse.

Maintenance of multiple spatial indices appears to be employed in naturalistic spoken conversation as well. In fact, in an experimental design that extends the Hollywood Squares paradigm to somewhat more ecologically valid circumstances, Fitneva (in preparation) recently collected data suggesting that this kind of use of spatial place holders occurs while people talk about different documents in the room with them. Participants were given two separate passages to read at different tables in the room. One was on corporate law, the other on ancient Egyptian burial practices. After the two passages had been read, they were enclosed in folders on their individual tables, and the participant sat facing the experimenter, with the two tables to his/her left and right. The experimenter read aloud random questions about the two passages from index cards (keeping her eyes on the cards), and the participant's eye movements were recorded during his/her attempts to answer the questions. The left/right position of the two passages was counterbalanced between subjects, and the experimenter was always blind to which passage was at which table.

Somewhat similar to this experimental design, Glenberg, Schroeder, and Robertson (1998) reported that people tend to look away from a speaker when they answer difficult questions (see also Kendon, 1967). However, in the present study, participants did not simply look away anywhere, they conspicuously looked *toward* an object that was relevant to, though unmentioned in, the question. Participants tended to gaze left of center when answering a question that concerned the content of the document on the participant's left-hand side, and right of center when answering a question that concerned the content of the document on the participant's right-hand side. As with the previous studies, accuracy was no better on the trials where participants did look at the correct side versus those where they did not.

To be sure, the participants in this study were fully aware that, during the question period, they could not possibly read anything on the pages enclosed in their folders on the tables 6 feet away. These fixations of the tables were certainly not explicit attempts to acquire external information from the environment. Rather, we suggest that they were relatively automatic—absent-minded, if you will—perusals of the spatial indices that had been allocated to the bundles of information inside the respective folders. A portion of the queried information was still available in short-term memory, and among that information was a pointer to the source of the entire set of information to be tested on. Participants' fixation of that source while recalling its content might even be thought of as a concrete manifestation of the process of source monitoring in memory (e.g., Johnson, Hashtroudi, & Lindsay, 1993).

POINTERS IN CONVERSATION AND GESTURE

As noted above, the idea that spatial properties of hand gestures can be informative in language processing seems natural enough for *signed* languages. However, gestures also play an important role in *spoken* languages (e.g., Goldin-Meadow, 1997; Rauscher, Krauss, & Chen, 1996). There are many categories of hand gesture that have been catalogued in natural speech, including *deictic* gestures (e.g., Levelt, Richardson, & la Heij, 1985), *interactive* gestures (Bavelas Chovil, Lawrie, & Wade, 1992), *iconic representational* gestures (McNeill, 1992), but the category that is most relevant to this next experiment is that of *metaphoric representational* gestures (McNeill, 1992). A metaphoric representational gesture is one in which the speaker indicates spatial locations in the present environment to metaphorically refer to nonpresent objects, entities, or temporal phases in a story. For example, one might extend a hand outward in front of one's body while saying, "So, he's getting C's in college and flirting with a cocaine habit," then move the hand rightward 1 foot, and say, "Then he enters big business and gets rich off of investors' money," then move the hand rightward again one foot, and say, "And now he's responsible for decisions that directly affect the entire world." In this example, the indicated spatial locations stand as place holders for periods of time, which can even be gestured to again later when revisiting discussion of the appropriate time period.[3] A listener's sensitivity to these kinds of gestures was tested in a recent experiment, again extending the Hollywood Squares paradigm to somewhat more ecologically valid circumstances. Richardson (in preparation) played digital video clips for participants, in which an actor talked about two alternative types of objects, animals, people, holiday destinations, etc. Each time the monologue discussed one alternative, a particular hand would always make a subtle movement (e.g., imagine the sort of gestures someone might make when saying 'On the one hand…but on the other hand…'). Although, these small movements resemble simple prosodic *beat* gestures (McNeill, 1992), after several correlated occurrences (e.g., left-hand moves for cats, right-hand moves for dogs), it becomes clear that they are acting as metaphoric representational gestures that metaphorically "place" one referent in one location in space and the other referent in another location in space. Left and right sides were counterbalanced between subjects by using mirror-image reversals of the digital video clips.

After each monologue, a blank display was shown, and a recorded voice asked a question about one of the referents (e.g., "What animal does my aunt own?"). While giving their answers, they tended to look at the half of the blank screen that the actor's subtle gestures had associated with the question's referent. For example, if the actor's right hand had been making metaphoric representational gestures during comments about dogs, then the participant tended to look at that side of the blank screen when giving the answer "a Labrador." Thus, much like in American Sign Language, a region of space had been recruited to serve as a place holder for the referent "dogs," and a different region of space had been recruited to serve as a place holder for the referent "cats." As in a signer's *signing space*, this layout of spatial indices not only allows for a topographic arrangement of the objects, entities, and topics of discussion, but also supports later deictic reference.

THE LOOK OF A MIND

As we have shown, spatial indices appear to be employed not just in low- and mid-level perception, such as perceptual-motor routines (Hayhoe, 2000) visual working memory (Ballard et al., 1995) and visual imagery (Spivey & Geng, 2001), but also in higher-level cognition, such as memory for semantic information (Richardson & Spivey, 2000) and even naturalistic conversation (Fitneva, in preparation; Richardson, in preparation). Based on these findings, we suggest that the objects of thought, the very things upon which mental processes directly operate, are not always inside the brain (e.g., O'Regan & Noë, 2001). The cognitive processing that gives rise to mental experience may be something whose functioning cuts across the superficial physical boundaries between brain, body, and environment (cf. Jarvilehto, 1998).

But what does such a non-brain-based mind look like? The mind, to an externalist, must be a rather graded entity, like a fuzzy set (Zadeh, 1973). In fuzzy set theory, the inclusion of members in a set is graded rather than all or none. For example, in traditional set theory, the set of *apples* would include, with full membership, all objects that are genuinely apples and nothing else. In contrast, fuzzy set theory would assign degrees of membership on a scale from 0 to 1. A fuzzy set is often depicted as something like a probability distribution, with a mode and tails that gradually approach zero. Fuzzy set theory is useful to an externalist because determining the discrete boundary in the external environment where things suddenly go from being "part of the mind" to being "not part of the mind" would presumably be impossible. Instead, one can hypothesize graded membership of external objects and events to the set of *mental contents*, gradually falling off with greater distance and with more mediated causes (e.g., Clark & Chalmers, 1998).

According to this version of externalism, the fuzzy set for *your mental contents* would include your brain, your body, as well as objects in the environment, and partially overlapping—at multiple spatial scales—with other mind-like fuzzy sets. Figure 5.3 presents an idealized sketch of this fuzzy set. The small oval in the middle of the diagram represents the classical set of *your brain contents*. Things inside that Venn diagram are part of your brain. Things outside it are not. The circumscribing oval represents the classical set of *your body contents*. Things inside that Venn diagram are part of your body. Things outside it are not. The fuzzy set of *mental contents* subsumes these two sets, and extends somewhat beyond them in x- and y-space. The third dimension of height in the diagram indicates degree of membership.

Importantly, the fuzzy set of *mental contents* includes to varying degrees, not just physical material in the present (such as a brain, a body, and other objects in the immediate environment), but also causal forces in that fuzzy set's history. As one traces back the causal forces of the environment's role in determining the set of *mental contents*, one must include—with some nonzero degree of membership—social influences accrued over days, parental influences accrued over decades, cultural influences accrued over centuries, and evolutionary influences accrued over many millennia.

An individual's personal (seemingly internally generated) sense of "intention" actually self-organizes across multiple coupled time scales from a combination of evolutionary, biological, cultural, parental, and social constraints (e.g., Gibbs, 1999; Juarerro, 1999; Van Orden & Holden, 2002; Van Orden, Holden, & Turvey, in press), not the least of which is—for evidence admissible to cognitive psychology, anyway—experimenter instructions and specific task constraints. In this view, mind becomes something not completely dependent on the body, although certainly not completely independent of it either. Rather, mind appears to be an emergent property that arises among

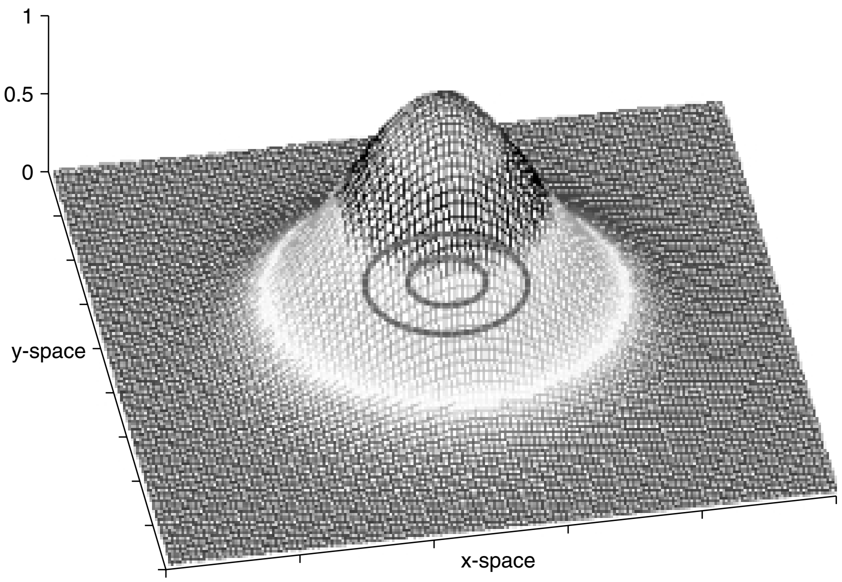

FIGURE 5.3. Along two spatial dimensions (x and y) the classical set of *body contents* (larger oval) circumscribes the classical set of *brain contents* (smaller oval). However, according to externalism, the fuzzy set of *mental contents* subsumes them both, as well as some of the properties of the surrounding environment, with a distribution function indicating degree of set membership (z-axis). Non-spatial dimensions that are likely to be relevant, such as semantic features and causal forces, are not depicted.

the interactions of a brain, its body, and the surrounding environment—which, interestingly, often includes other brains and bodies. Multiple brains, bodies, and environmental properties will often interact and function in a coherent manner that most decidedly resembles what we mean when we say "mind," as seen in collaborative task performance, mimicry, and other examples of social embodiment and embeddedness (Barsalou, Niedenthal, Barbey, & Ruppert, in press; Hutchins, 1995; Knoblich & Jordan, 2000; Stary & Stumptner, 1992).

Thus, the temporal dynamics of these fuzzy minds/sets become crucial for their accurate description— especially when one considers what happens as one fuzzy mind/set interacts with other fuzzy minds/sets over time. Figure 5.4a presents a schematic depiction of three bodies (and brains), like the one in Figure 5.3, moving in space as a function of time. Only one spatial dimension is shown so that the second dimension, of time, can be easily graphed. In Figure 5.4a, two bodies travel near one another for a period of time, then they diverge, and one of them begins traveling near a different body. As time is probably fractal, or self-similar, in this framework, the scale of the temporal dimension for these interactions could be just about anything. The bodies

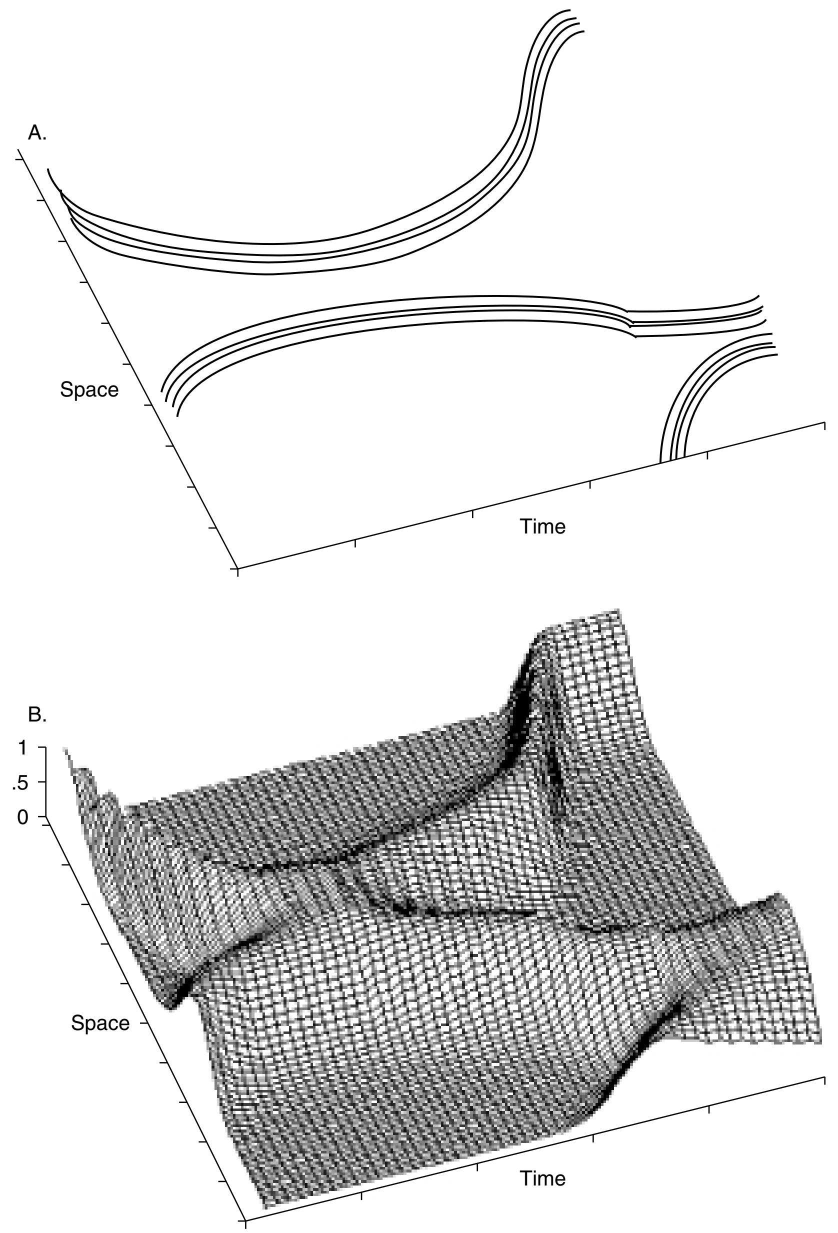

FIGURE 5.4. Using only one of the spatial dimensions from Figure 5.3, and adding a temporal dimension, panel A presents spatial trajectories of three bodies interacting over time. In panel B, the fuzzy set distributions intended to characterize the minds of those bodies do more than interact, they merge into one another at times.

could be interacting over the course of minutes (going from one hallway conversation to another), over the course of hours (going from one meeting to another), over the course of weeks or years or decades (going from one friendship/relationship to another).

For a fuzzy externalist, the depiction of these trajectories looks importantly different when they are defined over how the *minds* interact instead of how the *bodies* interact. Note, in Figure 5.4b, how the fuzzy set distributions merge as they approach one another. When two bodies are particularly close in space (and presumably close in other nondepicted semantic dimensions), the envelope of their distributions approaches having *one mode* instead of two. This demonstration offers a portrayal of how multiple different brains can cohere to such a degree that they function, at least to some extent, as though they were one mind. We suggest, in this chapter, that many of the interfacing links that maintain such phenomena are the spatial indices that connect bundles of information in one brain to bundles of information in other brains and in the rest of the environment.

A BROADER MINDSET

By looking at the use of spatial indices in a wide range of tasks and environments, we believe we have demonstrated some of the ways in which perception and cognition rely heavily on external objects and locations as the very stuff of mental activity. If we picture "mind" as the conglomeration of cooperating processes from both internal and external sources, we are logically forced—a little bit like the metaphysical functionalists (cf. Block, 1980; Fodor, 1981)[5] — to hypothesize group minds, nation minds, and even technologically expanded minds (e.g., Hewitt & Scardamalia, 1998).

Imagine sitting at your computer with the Internet browser open, and you've got an annoying mental block on a particular actor's name. You've spoken his name a hundred times before, you know his work well, but for some reason you have a habit of blanking on his name. (And by now you've probably *trained* the neural networks in your brain to do that.) You're in a "tip-of-the-tongue" state. You can recall that he was in Spike Lee's *Jungle Fever*, Quentin Tarantino's *Pulp Fiction*, as well as Francois Girard's *The Red Violin*, and many, many other movies as well. You can see his face in your mind clearly. His last name is a common two-syllable one that starts with J, and he uses his middle initial in his name. Aside from that, you can't quite dredge up his full name, but you *are* staring at your web browser. So you go to the Internet movie database, look up one of those films, skim through the cast credits, and his name pops right out.

What went on there? You knew some bits and pieces of a memory internally, but were unable to access one of the most important properties of that memory: its label. You were, however, able to internally access some potential addresses for that missing piece of memory. By going to the uniform resource locator (URL) http://www.imdb.com and finding its Web page for *Pulp Fiction*, you accessed the external content of one of the addresses that was available internally. Just like when you look at your watch to tell someone what time it is. And just like when our experimental participants, attempting to recall certain information, looked at particular regions in space that used to be home to a perceivable version of that information. When interfaced with the Internet, we have "minds" more expansive than ever before dreamt of in human history.

RAMIFICATIONS

As with the Internet, where knowledge is about navigation, not about storage, the mind, too, is best measured by its capabilities, not by its capacities—by its processing, not by its representations (cf. Jones & Smith, 1993; see also Pirolli & Card, 1999). Crucially, the mind's capabilities and processing are inextricably linked to the organism's interaction with the environment. Indeed, according to active externalism, it is that interaction between organism and environment from which "mind" emerges.

It should be noted that a wide adoption of this externalist concept of mind would have profound and far-reaching consequences for society. More than just reshaping the theories and experimental methods of psychology and cognitive science, by legitimizing the concepts of distributed cognition, transactive memory systems, and the collective mind (Jonasse, 1995; Nowak, Vallacher, & Burnstein, 1998; Yoo & Kanawattanachai, 2001), externalism promises new and different applied understandings of social behavior, group decision making, and relationships (e.g., Hutchins, 1995; Larson, Christensen, Franz, & Abbott, 1998; Pedersen & Larsen, 2001). For example, when you spend time with different groups from different demographic backgrounds, you don't just *act* like someone else, you *are* someone else. And for a couple to "be one" becomes more than a pleasing metaphor; it becomes scientific fact (cf. Hollingshead, 1998). Externalism has implications for culture as well, explaining how a tradition or fashion or sociological pattern might literally "have a mind of its own" (Cole & Engestroem, 1997). Indeed, the re-examination of the concept of individual responsibility instigated by externalism would shake the very foundations of (at least Western) legal theory, shifting much of the focus of reform from individual criminals to the criminogenic conditions that foster them (Haney, 2002). Finally, and possibly most important of all, a sincere espousing of externalism radically alters one's phenomenological sense of self. When the self is no longer conceived of as an ivory tower in the skull, it can be understood as an amalgam of interweaving influences from both internal and external sources. And so, perhaps, despite the intuitive appeal of one of Lakoff's (1987) favorite conceptual metaphors, the mind is not a container after all.

NOTES

1. Take, for example, Twin Earth, where Twin Gerry interacts with a fluid he calls "water" in just the same way that Gerry on Earth interacts with a fluid he calls "water," but the two fluids actually have very different chemical structures and thus are fundamentally different things. So what happens when our Gerry visits Twin Earth to go swimming with Twin Gerry, and they exclaim in unison, "Gosh, I like swimming in this water.?" If you think their respective mental states are not quite identical, then, like it or not, you're an externalist.
2. Although future research might benefit from providing distinctive markings in each stimulus port in order to provide unique visual contexts for memory primes, that would be more a study of context-dependent memory than of spatial indices.
3. Another example of a metaphoric representational gesture, that has a great deal in common with literal deictic gestures, is when someone points directly at an empty chair that was recently vacated by someone who has left the room (let's say, John) and says something like, "Wasn't John just making that same point?"
4. A similar analysis is easily yielded for light traveling through the eyes and producing the mental experience of vision.
5. Except that externalism itself is agnostic as to whether an organic nervous system is required as part of the conglomeration in order for the system to be considered "a mind."

ACKNOWLEDGMENTS

We are grateful to Larry Barsalou, Fernanda Ferreira, Monica Gonzalez-Marquez, Antje Meyer, Melinda Tyler, Becca Webb, and the group-mind of the entire workshop for helpful discussions and comments. We are also grateful to Carmela Alcántara, James May, Janice Ng, Adam November, and Anna Waisman for assistance with data collection. The new work described herein was supported by a Neuroscience Research Fellowship from the Alfred P. Sloan Foundation and by a Research Grant from the Consciousness Studies Program at the University of Arizona.

REFERENCES

Alibali, M. W., Heath D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language, 44,* 169-188.

Antrobus, J. S., Antrobus, J. S., Singer, J. L. (1964). Eye movements accompanying daydreaming, visual imagery, and thought suppression. *Journal of Abnormal and Social Psychology, 69,* 244-252.

Baddeley, A. D. (1986). *Working memory.* Oxford: Oxford University Press.

Ballard, D. H. (1989). Behavioral constraints on animate vision. *Image and Vision Computing, 7,* 3-9.

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience, 7,* 66-80.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., and Rao R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences, 20,* 723-767.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22,* 577-660.

Barsalou, L. W., Niedenthal, P. M., Barbey, A. K., & Ruppert, J. A. (in press). Social embodiment. In B. H. Ross (Ed.), T*he psychology of learning and motivation* (Vol. 43). San Diego, CA: Academic Press.

Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse Processes, 15,* 469-489.

Beer, R. (1989). *Intelligence as adaptive behavior.* New York: Academic Press.

Block, N. (1980). Introduction: what is functionalism? In N. Block (Ed.), *Readings in philosophy of psychology* (pp.171-184). Cambridge, MA: Harvard University Press.

Brandt, S. A., and Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience, 9,* 27-38.

Bridgeman, B., Van der Hejiden, A. H. C., Velichkovsky, B. M. (1994). A theory of visual stability across saccadic eye movements. *Behavioral and Brain Sciences, 17,* 247-292.

Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence, 47,* 139-159.

Bryant, D. J. (1997). Representing space in language and perception. *Mind and Language, 12,* 239-264.

Chatterjee, A. (2001). Language and space: Some interactions. *Trends in Cognitive Sciences, 5,* 55-61.

Clark A. (2001). Reasons, robots and the extended mind. *Mind and Language, 16,* 121-145.

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

Cole, M., & Engestroem, Y. (1997). A cultural historical approach to distributed cognition. In S. Gavriel (Ed), *Distributed cognitions: Psychological and educational considerations* (pp. 46).

Corbetta, M., & Shulman, G. L (1999). Human cortical mechanisms of visual attention during orienting and search. In G. Humphreys & J. Duncan, John (Eds.), *Attention, space, and action: Studies in cognitive neuroscience* (pp. 183-198). London: Oxford University Press.

Demarais, A. M., & Cohen, B. H. (1998). Evidence for image scanning eye movements during transitive inference. *Biological Psychology, 49,* 229-247.

Dewey, J. (1896). The reflex arc concept in psychology. *Psychological Review, 3,* 357-370.

Dreyfus, H. L. (1996). The current relevance of Merleau-Ponty's Phenomenology of Embodiment. Electronic *Journal of Analytical Philosophy, 4* (Spring).

Duncker, K. (1945). On problem solving. *Psychological Monographs, 58,5.doc”> No. 270.*

Eich, J. E. (1980). The cue dependent nature of state dependent retrieval. Memory and Cognition, 8, 157-173.

Emmorey, K. (2001). Space on hand: The exploitation of signing space to illustrate abstract thought. In G. Merideth (Ed), *Spatial schemas and abstract thought* (pp. 147-174).

Emmorey, K. (2002). *Language, cognition, and the brain: Insights from sign language research.* Mahwah, NJ: Erlbaum.

Fodor, J A. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences, 3,* 63-109.

Fodor, J. A. (1981). The mind body problem. *Scientific American, 244,* 114-123.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14,* 178-210.

Gibbs, R. W. Jr. (1999). *Intentions in the experience of meaning.* NY: Cambridge University Press.

Gibson, J. J. (1950). *The perception of the visual world.* Boston: Houghton Mifflin.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston: Houghton Mifflin.

Gilmore, R. O., & Johnson, M. H. (1998). Learning what is where: Oculomotor contributions to the development of spatial cognition. In S. Francesca & G. Butterworth (Eds.), *The development of sensory, motor and cognitive capacities in early infancy: From perception to cognition* (pp. 25-47). Hove, UK: Psychology Press.

Glenberg, A. M., Schroeder, J. L., & Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory and Cognition, 26,* 651-658.

Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes, 28,* 1-26.

Goldin-Meadow, S. (1997). When gestures and words speak differently. *Current Directions in Psychological Science, 6,* 138-143.

Grant, E. R., & Spivey, M. J. (in press). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science.*

Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American Psychologist, 53,* 5-26.

Haber, R. N. (1983). The impending demise of the icon: A critique of the concept of iconic storage in visual information processing. *Behavioral and Brain Sciences, 6,* 1-54.

Haith, M. M., Hazan, C., & Goodman, G. (1988). Expectation and anticipation of dynamic visual events by 3.5 month old babies. *Child Development, 59,* 467-479.

Haith, M. M., Wentworth, N., & Canfield, R. L (1993). The formation of expectations in early infancy. *Advances in Infancy Research, 8,* 251-297.

Hale, S M., & Simpson, H. M. (1970). Effects of eye movements on the rate of discovery and the vividness of visual images. *Perception and Psychophysics, 9,* 242-246.

Haney, C. (2002). Making law modern: Toward a contextual model of justice. *Psychology, Public Policy, and Law, 8,* 3-63.

Haugeland (1995). Mind embodied and embedded. In Y. Houng and J. Ho (Eds.), *Mind and cognition.* Taipei: Academia Sinica.

Hayhoe, M. M. (2000). Vision using routines: A functional account of vision. *Visual Cognition, 7,* 43-64.

Hayhoe, M. M., Bensinger, D. G, & Ballard, D. H. (1998). Task constraints in visual working memory. *Vision Research, 38,* 125-137.

Hebb, D. O. (1968). Concerning imagery. *Psychological Review, 75,* 466-477.

Henderson, J. M. Visual attention and saccadic eye movements. In G. d'Ydewalle & J. Van Rensbergen (Eds.), *Perception and cognition: Advances in eye movement research. Studies in visual information processing* (Vol. 4, pp. 37-50). Amsterdam, Netherlands: North Holland/Elsevier Science Publishers.

Hewitt, J., & Scardamalia, M. (1998). Design principles for distributed knowledge building processes. *Educational Psychology Review,* 10, 75-96.

Hoffman, J. E. (1998). Visual attention and eye movements. In H. Pashler (Ed.), *Attention* (pp. 119-153). Hove, UK: Psychology Press.

Hollingshead, A. B. (1998). Retrieval processes in transactive memory systems. *Journal of Personality and Social Psychology, 74,* 659-671.

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 113-136.

Hollingworth, A., Williams, C. C. & Henderson, J. M. (2001). To see and remember: Visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin and Review, 8,* 761-768.

Hutchins, E. (1995). *Cognition in the wild.* Cambridge, MA: MIT Press.

Irwin, D. E. (1993). Perceiving an integrated visual world. In D. Meyer & S. Kornblum (Eds.), *Attention and performance 14: Synergies in experimental psychology, artificial* *intelligence, cognitive neuroscience* (pp. 121-142). Cambridge, MA: The MIT Press.

Jarvilehto, T. (1998). The theory of the organism environment system: I. Description of the theory. *Integrative Physiological and Behavioral Science, 33,* 321-334.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*, 3-28.

Jonasse, R. (1995). Collectively seeing the wind: Distributed cognition in smokejumping. *Mind, Culture, and Activity, 2,* 81-101.

Jones, S. S., & Smith, L. B. (1993). The place of perception in children's concepts. *Cognitive Development, 8,* 113-139.

Juarrero, A. (1999). *Dynamics in action.* Cambridge, MA: MIT Press.

Kelso, J. A. S. (1995). *Dynamic patterns*. Cambridge, MA: MIT Press.

Kendon, A. (1967). Some functions of gaze direction in social interaction. *Acta Psychologica, 26,* 22-63.

Kirkham, N. Z., & Richardson, D. C. (submitted). Multi modal events and moving locations: Eye movements of adults and 6 month olds reveal dynamic spatial indexing.

Kirsh, D. (1995). The intelligent use of space. *Artificial Intelligence, 73,* 31-68.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science, 18,* 513-549.

Knoblich, G., & Jordan, S. (2000). Constraints of embodiedness on action coordination. In L. R. Gleitman and A. K. Joshi, (Eds.), *Proceedings of the 22nd annual conference of the Cognitive Science Society* (pp. 764-769). Mahwah, NJ: Lawrence Erlbaum Associates.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Lakoff, G. (1997). The internal structure of the self. In U. Neisser & D. A. Jopling, (Eds.), *The conceptual self in context: Culture, experience, self understanding. The Emory symposia in cognition* (pp. 92-113). New York: Cambridge University Press.

Larson, J. R., Jr., & Christensen, C., Franz, T. M., & Abbott, A. S. (1998). Diagnosing groups: The pooling, management, and impact of shared and unshared case information in team based medical decision making. *Journal of Personality and Social Psychology, 75,* 93-108.

Le Bon, G. (1916). *The crowd, a study of the popular mind.* London: T.F. Unwin.

Leslie, A. M., Xu, F., Tremoulet, P, & Scholl, B. J. (1998). Indexing and the object concept: Developing "what" and "where" systems. *Trends in Cognitive Sciences, 2,* 10-18.

Levelt, W. J., Richardson, G., & la Heij, W. (1985). Pointing and voicing in deictic expressions. *Journal of Memory and Language, 24,* 133-164.

Lévy, P. (1997). Collective intelligence: Mankind's emerging world in cyberspace. New York: Plenum Press.

Marr, D. (1980). *Vision: A computational investigation into the human representation and processing of visual information.* New York: W.H. Freeman.

McClamrock, R. A. (1995). *Existential cognition: computational minds in the world.* Chicago: University of Chicago Press.

McConkie, G. W., & Currie, C. B. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 563-581.

McNeill, D. (1992). *Hand and mind: what gestures reveal about thought.* Chicago: University of Chicago Press.

Merleau-Ponty, M. (1962). *Phenomenology of perception* (C. Smith, Trans.). New York: Humanities Press.

Neisser, U. (1967). *Cognitive psychology.* Englewood Cliffs, NJ: Prentice Hall.

Newell, A. (1990). *Unified theories of cognition.* Cambridge, MA: Harvard University Press.

Nowak, A., Vallacher, R. R., & Burnstein, E. (1998). Computational social psychology: A neural network approach to interpersonal dynamics. In W. Liebrand & A. Nowak, (Eds.), *Computer modeling of social processes* (pp. 97-).

O'Regan, J K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology, 46,* 461-488.

O'Regan, J. K., Deubel, H., Clark, J. J., & Rensink, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition, 7,* 191-211.

O'Regan, J K., & Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24,* 939-1031.

O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change-blindness as a result of 'mudsplashes'. *Nature, 398,* 34.

Pedersen, M. K., & Larsen, M. H. (2001). Distributed knowledge management based on product state models the case of decision support in health care administration. *Decision Support Systems, 31,* 139-158.

Pirolli, P., and Card, S. (1999). Information foraging. *Psychological Review, 106,* 643-675.

Putnam, H. (1975). The meaning of 'meaning'. In K. Gunderson (Ed.), *Language, mind, and knowledge.* Minneapolis: University of Minnesota Press.

Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial index model. *Cognition, 32,* 65-97.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition, 80*, 127-158.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3,* 179-197.

Rauscher, F. H., Krauss, R. M. & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science, 7,* 226-231.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8,* 368-373.

Richardson, D. C. & Spivey, M. J. (2000). Representation, space and Hollywood Squares: looking at things that aren't there anymore. *Cognition, 76,* 269-295.

Richardson, D. C., Spivey, M. J., Barsalou, L., & McRae, K. (in press). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science.*

Ross, D. (1997). Critical notice of Ron McClamrock's "Existential Cognition." *Canadian Journal of Philosophy, 27,* 271-284.

Ruggieri, V. (1999). The running horse stops: The hypothetical role of the eyes in imagery of movement. *Perceptual and Motor Skills, 89,* 1088-1092.

Ryle, G. (1949). *The concept of mind.* New York: Barnes & Noble.

Scholl, B. J., & Leslie, A. M. (1999). Explaining the infant's object concept: Beyond the perception/cognition dichotomy. In E. Lepore & Z. Pylyshyn (Eds.), *What is cognitive science?* (pp. 26-73). Oxford: Blackwell.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology, 38,* 259-290.

Segal, G. (2001). *A slim book about narrow content.* Cambridge, MA: MIT Press.

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

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real world interaction. *Psychonomic Bulletin and Review, 5,* 644-649.

Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects. *Psychological Research, 65,* 235-241.

Spivey, M. J., Tyler, M. J., Richardson, D. C., & Young, E. E. (2000). Eye movements during comprehension of spoken scene descriptions. *Proceedings of the 22nd annual conference of the Cognitive Science Society* (pp. 487-492). Mahwah, NJ: Erlbaum.

Stary, C., & Stumptner, M. (1992). Representing organizational changes in distributed problem solving environments. *IEEE Transactions on Systems, Man, and Cybernetics, 22,* 1168-1177.

Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine communication.* Cambridge: Cambridge University Press.

Tanenhaus, M., Spivey Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information during spoken language comprehension. *Science, 268,* 1632-1634.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press.

Tipper, S. P. & Behrmann, M. (1996). Object centered not scene based visual neglect. *Journal of Experimental Psychology: Human Perception and Performance,* 22, 1261-1278.

Turvey. M. T. (1977). Contrasting orientations to the theory of visual information processing. *Psychological Review, 84,* 67-88.

Van Orden, G. C., & Holden, J. C. (2002). Intentional contents and self control. *Ecological Psychology, 14,* 87-109.

Van Orden, G. C., Holden, J. C., & Turvey, M. T. (in press). Self-organization of cognitive performance. *Journal of Experimental Psychology: General.*

Wang, R. F. (1999). Representing a stable environment by egocentric updating and invariant representations. *Spatial Cognition and Computation, 1,* 431-445.

Wilson, R. (1994). Wide computationalism. *Mind, 103,* 351-372.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 121-134.

Yoo, Y., & Kanawattanachai, P. (2001). Developments of transactive memory systems and collective mind in virtual teams. *International Journal of Organizational Analysis, 9,* 187-208.

Zadeh, L. A. (1973). Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man, and Cybernetics, 3,* 28-44.