Johnson, M.K. (1991). Reflection, reality monitoring, and the self. In R.G. Kunzendorf (Ed.), *Mental imagery* (pp. 3-16). New York: Plenum Press.

REFLECTION, REALITY MONITORING, AND THE SELF

Marcia K. Johnson, Ph.D.

Department of Psychology
Princeton University
Princeton, NJ 08544

For some time now, I have been investigating the problem of reality monitoring--how people discriminate, when remembering, between information that had a perceptual source and information that was self-generated from thought, imagination, fantasy, or dreams (Johnson, 1977; Johnson, Taylor, & Raye, 1977). This question has many intriguing facets. It makes contact with fundamental philosophical issues of epistemology; it is a stimulating theoretical puzzle in itself; it is critically important for our everyday functioning in the world. It also has a certain science fiction quality about it that is disconcerting but compelling. If, as the analysis of reality monitoring that I will describe suggests, our ideas of reality and fantasy originate from imperfect attributional processes, who knows what might be true?

These philosophical, theoretical, practical, and science fiction aspects of reality monitoring could easily sustain my interest but, as it turns out, there has been an added attraction: Exploring the problem of reality monitoring has profoundly shaped my thinking about memory and cognition in general. The ideas about memory that have evolved from these efforts with colleagues and students to understand reality monitoring are summarized in a cognitive architecture that we call MEM-- a Multiple-Entry, Modular memory system (Johnson, 1983, 1990; in press; Johnson & Hirst, in press; Johnson & Multhaup, in press). Drawing on these earlier papers, I will briefly describe MEM and then some of our work on reality monitoring that grows out of the idea that records of reflective (or "self-generated") operations provide cues to the origin of information in memory. In the last section I will consider this "self" that does the generating, and discuss some implications of the MEM framework for how the idea of a "self" might arise and be maintained.

MULTIPLE-ENTRY MODULAR (MEM) MEMORY SYSTEM

According to MEM, memory is the result of processes that are organized at the most global functional level into perceptual and reflective systems. The perceptual system records information that is the consequence of perceptual processes such as seeing and hearing. The reflective system records information that is the consequence of internally-generated processes such as planning, comparing, speculating and imagining.

The perceptual system consists of two subsystems, P-1 and P-2, and the reflective system consists of two subsystems, R-1 and R-2 (Figure 1a). Each of these, in turn, includes component subprocesses. Suggestions for what some of these component subprocesses might be are given in Figures 1b and 1c. Subprocesses of P-1 might include *locating* stimuli, *resolving* stimulus configurations, *tracking* stimuli, and *extracting* invariants from perceptual arrays (e.g., cues specifying the rapid expansion of features in the visual field). P-1 processes develop connections or associations involving perceptual information of which we are often unaware, such as the cues in a speech signal that specify a particular vowel, or the aspects of a moving stimulus that specify when it is likely to reach a given point in space. Learning in the P-1

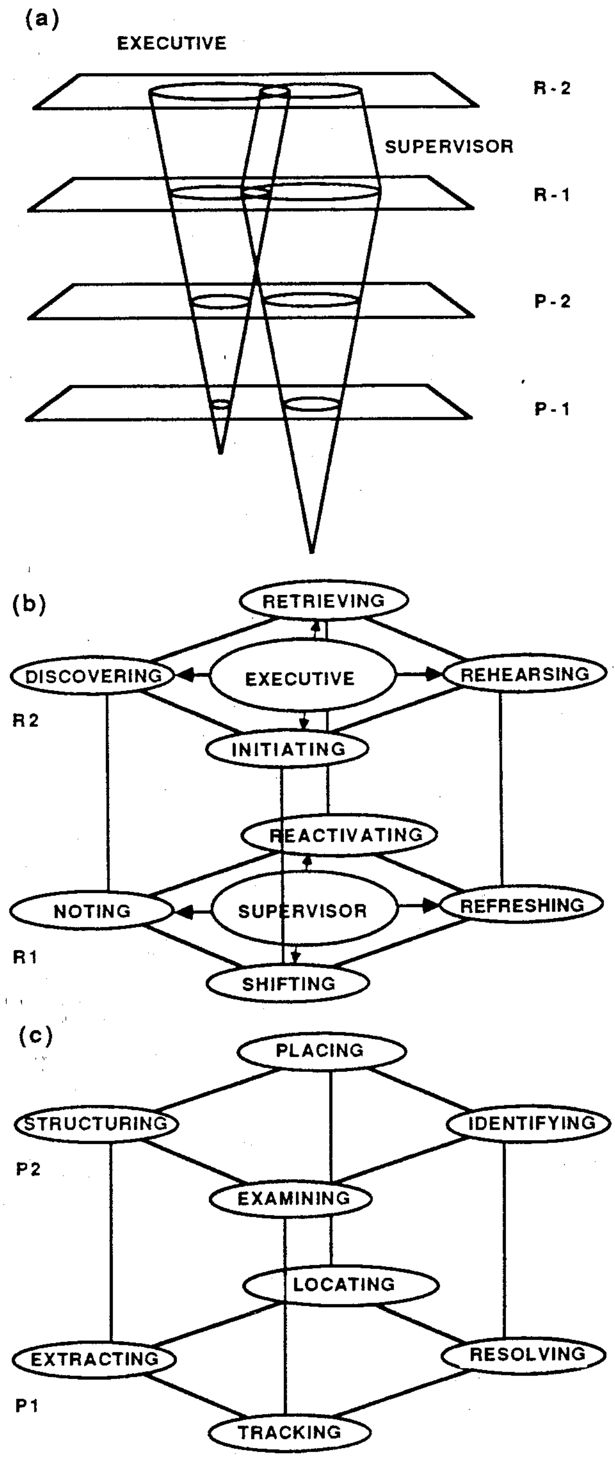

Figure 1. A multiple-entry, modular memory system:
(a) configuration of subsystems;
(b) component subprocesses of R-1 and R-2;
(c) component subprocesses of P-1 and P-2.

subsystem allows us to do things such as adjust to a person's foreign accent, or to anticipate the trajectory of a baseball. Subprocesses of P-2 might include *placing* objects in spatial relation to each other, *identifying* objects, *examining* or perceptually investigating stimuli, and *structuring* or abstracting a pattern of organization across temporally extended stimuli (e.g., syntactic structure). P-2 processes are involved in learning about the phenomenal perceptual world of objects such as chairs and balls, and events such as seeing a person sit down in a chair or catch a ball.

In contrast to perceptual processes, reflective processes occur independent of sensory stimulation. They may sometimes be initiated by perception, but reflective processes allow one to go beyond external cues. They are generative; they allow us to manipulate information and memories, anticipate events, imagine possible alternatives, compare these alternatives, and so forth. R-1 processes and R-2 processes differ in the complexity of the tasks they can handle. For example, R-1 processes would be sufficient for anticipating a picnic, but R-2 processes would be necessary for planning one.

Both R-1 and R-2 involve component processes that allow people to sustain, organize, and revive information. Some of these component processes in R-1 are *noting* relations, *shifting* attention to something potentially more useful, *refreshing* information so that it remains active and one can easily shift back to it, and *reactivating* information that has dropped out of consciousness. Component processes in R-2 include *discovering*, *initiating*, *rehearsing*, and *retrieving* (see Johnson, 1990, and Johnson & Hirst, in press, for a discussion of these component processes).

The R-1 and R-2 subsystems also include control and monitoring component processes. For R-1 these are collectively referred to as *supervisor* processes and for R-2 as *executive* processes. Supervisor and executive processes set up goals and agendas and monitor or evaluate outcomes with respect to these agendas (Miller, Galanter, & Pribram, 1960; Nelson & Narens, 1990; Stuss & Benson, 1986). Furthermore, they recruit other reflective component processes for these purposes. There is more effort, will, or control (e.g., Hasher & Zacks, 1979; Norman & Shallice, 1986; Shiffrin & Schneider, 1977) associated with executive than with supervisor processing. The difference between supervisor and executive processes is something like the difference between "tactical" and "strategic" control, or the difference between habitual and deliberate reflective processes. For example, under the guidance of an R-1 intention to *listen attentively to a story* told by your dinner companion, you might generate tacit implications of sentences, notice relations between one part of the story and an earlier part you are reminded of, and so forth. Under guidance of an R-2 agenda to *critically evaluate the story*, you might generate objections to the logic of events in the story, actively retrieve other stories for comparison, and so forth. Normal cognitive functioning draws on different component processes of reflection as needed. Disruption of various combinations of component processes results in various patterns of cognitive deficit (Johnson & Hirst, in press).

P-1, P-2, R-1, and R-2 activities go on simultaneously and they produce corresponding changes in memory. These changes are the representations of experience. At some future time, exactly which of these records are activated will depend on the kind of task probing memory (i.e., according to the encoding specificity principle, Tulving, 1983). A task in which you had to identify random syllables spoken by your dinner companion that a devious psychologist embedded in white noise would draw primarily on representations formed by P-1. A recognition task in which you had to discriminate pictures of people who were and were not at the dinner party should draw primarily on representations formed in P-2. Recall of your dinner companion's story would draw on R-1 and R-2 records.

Under ordinary circumstances, subsystems interact with each other, although exactly how they interact needs further investigation. During remembering, representations from one subsystem may directly activate related representations from another, or interactions between perceptual and reflective memory may take place through supervisor and executive components. For example, an agenda initiated by the R-2 executive, such as *look for a restaurant,* might activate relevant perceptual schemas from perceptual memory (e.g., look for building with ground level window, tables visible, menu in window). It might also activate reflective plans adapted to the current situation (e.g., check the restaurant guide for this part of town).

In Figure 1a, supervisor and executive processes are depicted as cones passing through planes representing different subsystems. The sizes of the ellipses at the intersects of cones and planes reflect the relative degree of involvement of supervisor and executive processes in each subsystem's activities. Typically, executive functions have greater access to reflective memory than to perceptual memory, and greater access to P-2 than to P-1 subsystems. An especially important aspect of reflection is that the supervisor and executive processes in R-1 and R-2 can recruit and monitor each other, as depicted by their overlap in Figure 1a. For example, an R-2 agenda to *check the restaurant guide* can initiate an R-1 goal to *note the number of stars by each entry.* Interaction between R-1 and R-2 provides a mechanism for sequencing subgoals. It also gives rise to the phenomenal experience of reflecting on reflection which, as I will discuss later, is intrinsic to our sense of self. In addition, access to information about one's own cognitive operations provides a salient cue for identifying oneself as the origin of information (Johnson, Raye, Foley, & Foley, 1981).

I have used the MEM framework to organize empirical facts obtained from cognitive-behavioral studies (Johnson, 1983). It also has prompted and been shaped by our research on anterograde amnesia (Hirst, Johnson, Kim, Phelps, Risse, & Volpe, 1986; Hirst, Johnson, Phelps, & Volpe, 1988; Johnson & Kim, 1985; Johnson, Kim, & Risse, 1985; Weinstein, 1987; see also Johnson, 1990), as well as our work on reality monitoring (Johnson & Raye, 1981; Johnson, 1988a,b, in press), and we have used it to discuss the relation between cognition and emotion (Johnson & Multhaup, in press). Several behavioral dissociations support the idea that the division between perceptual and reflective memories may capture functional organizations within the nervous system as well. For example, an argument can be made that memory for reflective processing develops later than memory for perceptual processing (e.g., Flavell, 1985; Moscovitch, 1985; Perlmutter, 1984; Schacter & Moscovitch, 1984); and that P-2 develops later than P-1 and R-2 later than R-1. Moreover, memories for reflective processing appear to be disrupted more easily by stress, depression, aging, and the use of alcohol and other drugs than are memories for perceptual processes (Craik, 1986; Eich, 1975; Hasher & Zacks, 1979; 1984; Hashtroudi & Parker, 1986). Furthermore, the breakdown in memory functioning found in patients with anterograde amnesia appears to fall disproportionately on reflective memory (Johnson, 1983, 1990; Johnson & Hirst, in press; see also chapters in Cermak, 1982).

A system like MEM is an extremely powerful cognitive architecture. For example, guided by learning in the perceptual subsystems, we can hit a tennis ball and at the same time reflectively think about a strategy for playing the game. As I mentioned before, the interaction between R-1 and R-2 permits sequencing of subgoals to perform complex tasks such as going on picnics and writing research papers. The availability of both perceptual and reflective subsystems allows us to build up a veridical representation of the world through P-1 and P-2 processes and yet to imagine, or reflectively generate, worlds as they might be.

But the same reflective processes that underlie our planfulness and creativity produce a crucial dilemma for us. Reflection generates events that take place only in thought and imagination. So, how do we know the life we remember is the product of actual experiences rather than only experiences we have thought, fantasized, or dreamed?

REALITY MONITORING

Reality monitoring is the term Carol Raye and I suggested for the processes involved in discriminating memories that originated from perception from those that arose from thought, imagination, fantasy, dreams and other self-generated processes (Johnson & Raye, 1981; Johnson, 1985; in press). Reality monitoring failures occur when people confuse the origin of information, misattributing something that was reflectively generated to perception or vice versa. According to this view, reality is not directly given in remembering, but is an attribution that is the outcome of judgment processes. To understand reality monitoring confusions we have to consider both the phenomenal characteristics of memories and the decision processes people apply to them.

We proposed that memories for perceived and imagined events differ in average value along a number of dimensions. Memories originating in perception typically have more perceptual information, (e.g., color, sound), more contextual information such as time and place, and more meaningful detail, whereas memories originating in thought typically have more accessible information about cognitive operations--that is, about those perceptual and reflective processes that took place when the memory was established. Judgment processes monitored by R-1 capitalize on these differences. That is, differences between externally and internally derived memories in average value along these dimensions or attributes form one basis for deciding the origin of a memory. For example, one would be likely to decide that a memory with very little cognitive operations information and a great deal of perceptual information was externally derived (Johnson, Raye, Foley, & Foley, 1981; Johnson, Raye, Wang, & Taylor, 1979).

In contrast, R-2 processes control a second type of decision process based on reasoning:

This may include, for example, retrieving additional information from memory and considering whether the target memory could have been perceived (or self-generated) given these other specific memories or general knowledge (e.g., Johnson, Foley, Suengas, & Raye, 1988). For example, I might have a memory of a challenging question asked by a colleague during a talk I gave (and my perfect answer), but correctly attribute this to imagination on the basis of knowledge I have that he was out of town when I gave the talk. In addition, judgments will be affected by people's opinions or by "metamemory" assumptions about how memory works.

For normally functioning adults, most reality monitoring is guided by R-1 supervisor processes; that is, reality monitoring typically takes place rapidly, in a nondeliberative fashion, based on the qualitative characteristics of memories that are activated (e.g., amount or type of perceptual detail). The generally slower, more deliberate retrieval of supporting memories and initiation of reasoning processes (e.g., Does this seem plausible given other things I know?) are R-2 functions and are engaged less often. Among other things, R-2 processes allow us to look back on ideas we initially accepted and question them.

Using this framework, consider the various ways in which reality monitoring could break down (Johnson, in press): Disrupted reality monitoring would result from any circumstance that decreases differences between phenomenal qualities of perceived and imagined events such as unusually vivid imagery or reduced cognitive operations associated with imagined information. Difficulty in retrieving relevant supporting information would also cause problems in reality monitoring, as would any disruption in R-1 or R-2 judgment processes, including more lax criteria (e.g., requiring less perceptual information to decide something had been perceived). In addition, reduced motivation to engage in reality monitoring would result in more confusions between fact and fantasy. Thus any one of these factors, or any combination, would reduce the accuracy of reality monitoring. Furthermore, we should be able to characterize clinically significant disruptions of reality monitoring such as occur in hallucinations (Bentall, 1990; Horowitz, 1978), delusions (Johnson, 1988a), confabulation (Johnson, in press), and hypnosis (Hilgard, 1977; Kihlstrom, 1987; Kunzendorf, 1986), in terms of the various factors suggested by this framework.

Most of our laboratory studies have been directed at exploring the viability of this framework for reality monitoring in normally functioning individuals. For example, one straightforward prediction is that the more imaginations are like perceptions in perceptual detail, the more subjects should confuse imaginations with perceptions. In one experiment testing this hypothesis (Johnson, Raye, Wang, and Taylor, 1979), we varied the number of times subjects saw pictures and the number of times they imagined each picture. Later, we asked subjects how many times they saw each picture. The more often subjects saw pictures, the higher their frequency judgments. More important, the more often subjects imagined the pictures, the more often they thought they had seen them. Furthermore, compared to poor imagers, good imagers were more affected by the number of times they had imagined a picture.

In another experiment on this point (Johnson, Foley & Leach, 1988), subjects imagined themselves saying some words and heard a confederate say other words. Later, subjects were asked to discriminate the words that they had thought from the words the confederate had actually said and they were quite good at this. In another condition, the procedure was the same except that subjects were asked to think in the confederate's voice; in this case subjects later had much more difficulty discriminating what they had heard from what they had thought. This study, as well as the good/poor imager study, is consistent with the idea that the more perceptual overlap there is between memories derived from perception and memories generated via imagination, the greater will be the confusion between them. (This is one reason to be careful not to make the imaginary conversations and arguments you have with people too perceptually detailed.)

REFLECTIVE COGNITIVE OPERATIONS

The perceptual quality of images is, perhaps, their most obvious and intriguing phenomenal characteristic, and certainly an aspect of them that has prompted a great deal of empirical research (e.g., Finke & Shepard, 1986; Kosslyn, 1980). The fact that perceptual detail in a memory can cause one to mistake an imagination for a perception while remembering is fundamental, but not too surprising. What has been more surprising to us is the critical importance that records of cognitive operations appear to play in reality monitoring (e.g., Johnson, Raye, Foley, & Foley, 1981; Rabinowitz, 1989). Thus, here, I want to stress an aspect of images other than their perceptual quality, namely, how they get generated. Images are under varying degrees of reflective control, and the degree of reflective control has implications for how easily they will be distinguished from percepts.

According to the reality monitoring framework, remembered cognitive operations can be a cue to the origin of a memory. If an image is generated with considerable reflection or effort, then the internally generated memory should be easier to distinguish from a memory for an external event than if the image is evoked without volition. We have done a number of studies exploring this hypothesis.

In one study (Johnson, Kahan, & Raye, 1984) based on this idea, we had people who lived together report their dreams to each other or report to each other "dreams" that they made up. Then we brought the people into the lab and had them attempt to distinguish their own reports from the reports they had heard from their partners. People experienced more confusion about the source of actual dreams than made up dreams. Our interpretation of this finding is that, compared to dreams made up under voluntary control, memories for real dreams contain less information about the cognitive operations used in the generation process, making them harder to distinguish from memories for perceived events.

The outcome of the dream study was encouraging, but studying dreams does not allow much in the way of experimental control, so we have attempted to manipulate cognitive operations in the lab. That is, our hypothesis is that memories based on normal, waking imagery will be confused with memories for perceived events, depending on how easily the images are generated. We would predict that images that are easier than others to generate would lead to greater confusion in reality monitoring.

We (Finke, Johnson, & Shyi, 1988) explored this idea using forms such as those in Figure 2. Previous studies have shown that subjects can judge symmetry more rapidly for vertically symmetrical patterns than for horizontally symmetrical patterns (e.g., Corballis & Roldan, 1975). This finding suggested to us that if we showed subjects only half of a pattern and had them imagine it as completed, they would find it easier to do this when the form is symmetrical about the vertical axis than when it is symmetrical about the horizontal axis. We confirmed this expectation in a preliminary study in which subjects were shown half forms and were instructed to imagine each half form as being completed about the axis of symmetry to make a symmetrical whole form. Once they had done so, they were to rate how easy or difficult it was to imagine completing the form. As predicted, the mean rated difficulty of completing the forms was less for vertical stimuli than for horizontal stimuli.

In the main experiment, we used the fact that vertical forms are easier to imagine as complete than are horizontal forms to test our prediction that easily generated images are more likely to be confused with perceptions than are images that are more difficult to generate. Some subjects were randomly assigned to an imagery condition, and some to a control condition. In both conditions, subjects saw stimuli in random order for five seconds each. Some of the items were forms presented in whole versions and some were presented in the incomplete versions. Subjects in both conditions were asked to rate the complexity of the forms. Control subjects

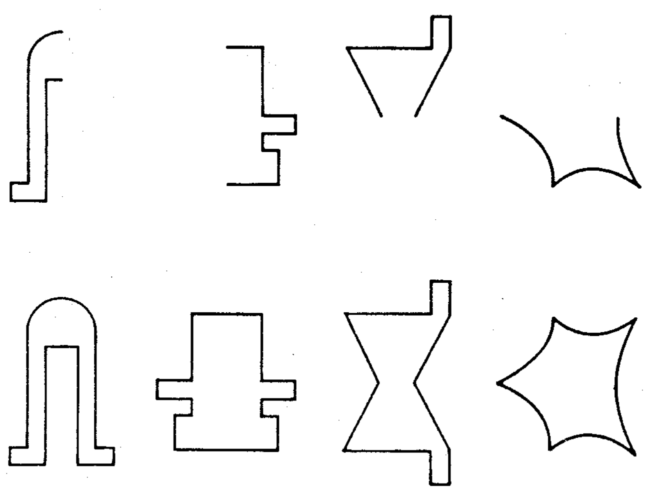

Figure 2. Example of forms used by Finke, Johnson, & Shyi, 1988.

rated each form as it was presented. Imagery subjects also rated the whole forms as presented but were instructed that if the slide contained a half form, they were to imagine it as a completed, symmetrical form, and then rate the complexity of the imagined form.

After completing their ratings, subjects received a surprise memory test. They were shown test patterns in random order consisting of whole forms they had just seen, whole forms corresponding to completions of the half forms they had just seen, and new whole forms. Subjects were asked to indicate whether each pattern had been seen as a whole, contained half of a form they had seen, or had not been shown as a whole or as a half. We explained to the imagery subjects that a form that was presented in half form but completed by imagination should be identified as a half.

First, the data were scored without regard to whether the subject remembered if the form had been presented as a half or a whole. These recognition scores are the number of correct "old" responses plus the number of correct rejections of new items, divided by the total number of test items. The overall proportion correct for this old/new discrimination score was about .72 and there were no differences among conditions. These old-new recognition results make two points. The forms were equally memorable in terms of recognition when presented vertically and when presented horizontally. Second, the performance of the imagery and control groups on old-new recognition was equivalent; hence, any differences between instructional groups in their ability to remember whether forms were presented as halves or wholes can be attributed to differences between groups in reality monitoring rather than in recognition.

To assess reality monitoring, half-whole discrimination scores were also computed. This score reflects the proportion of items identified as old that were also correctly identified as presented in half or whole form. Imagery (.75) and control (.75) subjects did not differ in their ability to discriminate half from whole forms in the horizontal condition, whereas the performance of the imagery (.67) subjects was significantly worse than that of the control (.77) subjects in the vertical condition. The most straightforward interpretation of these results is that reality monitoring is more difficult in the vertical than in the horizontal condition because images of vertically symmetrical forms are more easily constructed; thus the memory for the construction includes less information about cognitive operations. Later, in evaluating the origin of a memory, low values for cognitive operations for imagined as well as perceived events would make it difficult to distinguish them.

In a new series of experiments (Johnson, Finke, Danzer, & Shyi, in preparation), we have manipulated the ease of generation of images in a different way. Stimuli consisted of black and white patterns created by filling in squares on a five by five grid (Figure 3a). Some of the stimuli were familiar alphanumeric characters, and some were unfamiliar patterns. From these whole patterns, two incomplete versions were created. One incomplete version had three missing contiguous squares, each indicated by a dot. The other incomplete version had three missing squares randomly located throughout the figure, also indicated by dots. It should be easier to imagine the completions of familiar than unfamiliar forms, and easier to imagine forms with contiguous rather than random squares missing.

Subjects were shown some whole and some incomplete forms (equal numbers of familiar, unfamiliar, random and contiguous items) for five seconds each. Subjects in the control condition rated the complexity of the patterns as they were presented, ignoring the dots on the incomplete patterns. Subjects in the imagery condition rated the complexity of the whole patterns as presented; they imagined the incomplete patterns as whole by mentally filling in the dots and then rated the complexity of the imagined, whole pattern. This inspection phase was followed by an unexpected memory test in which only whole patterns were shown--some that were presented during the inspection phase, some that were whole versions of the incomplete patterns shown during the inspection phase, and some that were new whole patterns; subjects were asked to indicate which was which.

Again, first consider the old-new recognition. Subjects in the imagery group (.74) were actually better able to discriminate old from completely new forms than were subjects in the control group (.69). In contrast, the control group (.82) was better at discriminating whole from incomplete forms compared to the imagery group (.73). Thus, as in the last experiment, the difficulty imagery subjects had in reality monitoring cannot be attributed to an overall poorer memory for the items as indicated by old/new recognition.

We then looked at a confusion measure that reflects the proportion of items for which subjects claimed to have seen a complete pattern when an incomplete one had been presented (misattributions in Figures 3b and 3c). As you can see in Figure 3b, overall, imagery subjects were more likely to say a whole form had been presented when only half had been. Furthermore, the difference between the imagery and control groups was greater for familiar items than for unfamiliar items. This finding is consistent with the idea that familiar items were easier for the imagery group to imagine and then subsequently harder to discriminate from perceived items. In Figure 3c, imagery and control groups have been combined because for both imagery and control groups the pattern was the same -- no difference between contiguous and random conditions for familiar items, and more misattributions on contiguous than random items for unfamiliar patterns. Both imagery and control subjects were least likely to claim to have seen whole figures when the presented figure was unfamiliar and contained randomly missing squares.

In summary, several factors appear to influence the likelihood that subjects will claim to have seen an entire pattern when only a partial pattern was presented. Familiar patterns, or those with contiguous parts missing, are sometimes spontaneously filled in even without an explicit intention to make images, and then later may be misidentified as patterns actually presented whole. Explicit instructions to engage in imagery increased misattributions but less so for unfamiliar patterns. Our interpretation of this pattern is that imagery subjects were better at reality monitoring for unfamiliar than familiar patterns because unfamiliar patterns require more cognitive operations to fill in; these cognitive operations can later be used to identify oneself as the origin of the completion.

The fact that even control subjects made some misattributions in this last experiment suggests that we not only create images under voluntary control, that is, on purpose, but we also create images spontaneously, elicited by ongoing experiences as a way of filling in, or as a natural byproduct of perception and comprehension. The results of another experiment (Durso & Johnson, 1980) illustrate that such spontaneous imaginations may be the most difficult of all to later distinguish from perceptions. Subjects saw a list consisting of some words and some pictures (line drawings of common objects). Then we gave subjects a surprise test that asked them to indicate whether each item had appeared as a word or a picture (there were new items as well). We varied how the subjects processed the items initially in ways that should have affected later availability of information about cognitive operations. At acquisition, some subjects indicated the function of the referent of each item. For example, if they saw a picture of a knife (or the word knife) they might say "you *cut with* it". Still other subjects identified a particularly relevant feature of each object, for example, *blade* for knife. Other subjects had an artist time judgment task: If a picture was presented, they rated how long it took the artist to draw it. If a

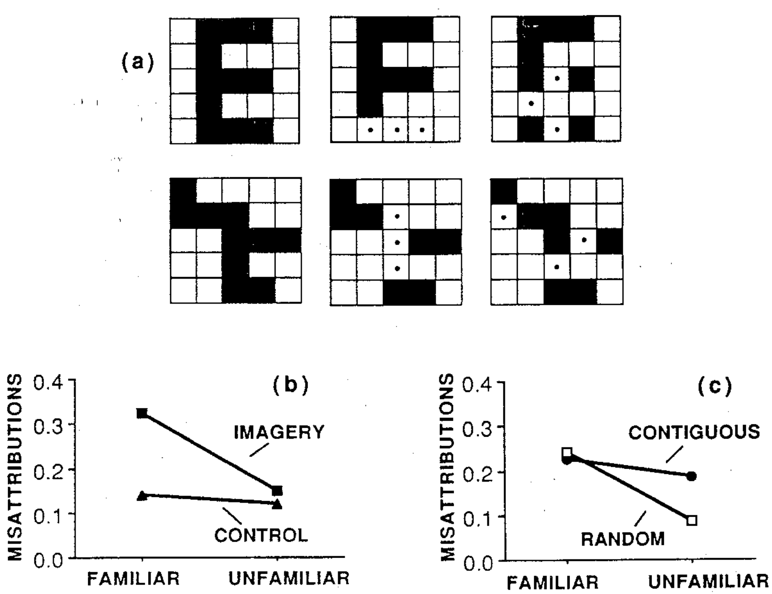

Figure 3. Johnson, Finke, Danzer, & Shyi (in preparation):
(a) Example stimuli are shown;
(b & c) Misattributions are incomplete forms that subjects claimed to have seen as complete forms.

Table 1. Mean Number of Times Subjects Said *Word* Given the Item Was a Picture (W/P) and *Picture* Given the Item Was a Word (P/W), Durso & Johnson, 1980.

	Number of Confusions W/P	P/W
Function	3.42	13.25
Relevant feature	2.08	11.08
Artist time judgment	5.42	3.50

word was presented, they constructed an image of a line drawing of the referent and then rated how long it would take the artist to draw the imagined picture.

One important difference between the first two tasks and the last one is that the artist-time judgment task involves explicit imagery, whereas the function and relevant feature tasks are likely to involve spontaneous, or incidental imagery. That is, in order to answer a question about the object's relevant features, you might think of a visual representation of the object and "pick out" a relevant feature. Explicit images are under voluntary, reflective control and thus the memories for them should contain more information about cognitive operations than the memories for spontaneous images. If so, people should later be better able to discriminate memories of voluntarily constructed images from memories of pictures than they can discriminate memories of spontaneous images from memories of pictures. The results (Table 1) were consistent with this prediction. Consider the number of times subjects claimed to have seen a picture when only a word was shown. In the artist time judgment task, subjects rarely said a word had been presented as a picture. In the function and relevant feature tasks, they much more often claimed to have seen pictures of objects that had only been named. Thus spontaneous or incidental images were more likely to be confused with perceptions than were consciously constructed ones.

Foley, Durso, Wilder, & Friedman (in press) recently used a similar paradigm with children to investigate age-related reality monitoring processes. It has often been proposed that children have a more difficult time than adults distinguishing fact from fantasy. In previous research, however, we had not found any greater tendency for children, compared with adults, to confuse pictures they imagined with pictures they perceived (Johnson, Raye, Hasher, & Chromiak, 1979). Foley et al. (in press) considered the possibility that we may have underestimated reality monitoring failures in children because children might not be as successful as adults at creating images on demand. If reality monitoring for children's *spontaneous* images were examined, children might show more confusion than adults. Foley et al., compared 6-year old children and adults in two tasks, one in which there were explicit imagery instructions and one in which images would be likely to be spontaneously generated (the function condition). Children, like adults, were more likely later to confuse spontaneous than purposeful images with perceptions; and, as in our previous research, the amount of confusion children showed was no more than that of adults. These results suggest that by the age of six, children are quite sensitive to the value of cognitive operations as a cue to the origin of memories.

THE SELF

As we have seen, self-generated or reflective activity produces a record of cognitive operations that can be used in reality monitoring. Let's consider for a few minutes this idea of "self." Many philosophers and psychologists have tackled the concept of the self. Though not easy to define (e.g., Bandura, 1982; Greenwald, 1982), the self is the target of a collection of provocative questions such as: Where does the idea of a self come from (e.g., Bem, 1972; Gergen, 1977)? Is the self a concept like any other concept or is it a special organizational structure (e.g., Markus & Sentis, 1982)? How stable is our self-concept (Gergen, 1982)? What are the consequences of self-awareness (Wicklund, 1982) or self-conceptions (Snyder & Campbell, 1982)? What role does self play in consciousness (Kihlstrom, 1987; Kunzendorf, 1987) Is the self a unified entity, or, in fact, made up of dissociated selves (Beahrs, 1982; Hilgard, 1977) or independent subsystems (Greenwald, 1982)? I think that one potentially interesting feature of MEM is that a number of ideas about self emerge from the architecture of the system; furthermore, which aspects of self are in the foreground depends on which aspects of MEM we are considering.

Self and Other

First consider the fundamental distinction between self and other. MEM postulates subsystems corresponding to perceptual experience and reflective activity. Reality monitoring is a set of processes that differentiate perceptually derived from reflectively generated information. How critical reality monitoring is for normal functioning becomes obvious when it breaks down markedly as in delusions (Johnson, 1988a) and confabulation (Johnson, in press). It seems reasonable to conclude that reality monitoring processes evolved along with reflective creativity and helped offset potentially dysfunctional confusion between the perceived and the imagined. The self (as originator), as distinguished from that which is perceived, arises naturally as a byproduct of this reality monitoring activity. Of course, the nature of one's experiences, especially one's interactions with others, will greatly influence how this idea of self becomes elaborated (e.g., Higgins, 1987; Markus & Nurius, 1986).

Agency and Self-Regulation

Next consider the important concept of agency. Bandura (1982) has suggested that most of the fundamental questions regarding the nature and functions of the self are concerned with the problem of human agency. "The matters of major interest center on whether, and how, people exert some influence over what they perceive and do (p. 3)." In Bandura's model, self-regulation operates through a set of subfunctions which include self-observation and judgmental processes. That is, to change how one is behaving one must observe one's own behavior and apply internal standards for evaluating it. The kind of self-regulation Bandura describes is possible in MEM because (1) the reflective system can access and evaluate perceptual information in terms of standards activated in the reflective system, and (2) the reflective system has two interacting subsystems, R-1 and R-2 that can evaluate and call on each other. Often R-1 and R-2 simply work toward the same goal, as in many problem solving situations. However, we seem to get a special sense of self-control when an R-2 agenda functions to counteract some ongoing agenda from R-1 as when, for example, an R-2 agenda to lose weight attempts to counteract an R-1 agenda to fix a sandwich. The phenomenal experience from interactive reflective activity is so powerful that, in effect, we often identify ourself with it. That is, we feel our own agency through intention, purpose, volitional action, planning, control, and so forth. These words are short-hand ways of referring to the fact that R-2 recruits and monitors R-1 and vice/versa; at the same time, they refer to the phenomenal experience resulting from this R-1/R-2 interaction.

The Self as Both Subject and Object (Self-Awareness)

Another classic problem of the self is how the self can be both subject and object (or knower and that which is known, James, 1915, cited in Markus & Sentis, p. 44). How such dual functioning is achieved in MEM is not hard to see. One R subsystem can "know" another (i.e., monitor its activities), just as R subsystems can monitor some activities of the perceptual subsystems. The overlapping supervisor and executive functions give us access to our own cognitive operations, from which we derive the phenomenal experience of thinking about ourselves thinking. Such self-awareness, like self-regulation, thus arises from the special relation between R-1 and R-2 subsystems.

Aspects or Styles of Self

We might also suppose, looking at MEM's architecture, that there would be different aspects of self or styles of self. I have been emphasizing that a sense of self arises from reflective control and monitoring of cognitive operations. This self as actor or agent might be called an *instrumental* self. I do not mean to imply, however, that this is the only way a sense of self comes about and is maintained. In addition to an instrumental self, there is another type of self that is suggested by MEM's architecture. This is an *experiential* self that arises from focusing on information from the perceptual subsystems. (The experiential self is, perhaps, something like Hume's idea that the self is nothing but a bundle of perceptions.) Especially important should be those perceptions that give rise to familiarity responses; my relatives, clothes, house, office, friends, car. These familiarity responses contribute to our sense of a temporally extended existence and, thus, to our idea of an ongoing self. A person's "autobiography" arises from both experiential activity and instrumental activity. One difference in the consequences of these two types of activities is that a primarily experiential focus might produce a sense of *fate* about what happens to us and who we are, whereas a primarily instrumental focus might produce a sense of *choice*. In any event, the idea of instrumental vs. experiential focus could characterize reasonably stable differences in self concepts across individuals, or differences in aspects of the self that are salient within the same individual at different times.

Styles of Self and Reality Monitoring

Returning now to reality monitoring, we might speculate that experiential and instrumental styles of self would be associated with characteristic differences in primary mode of reality monitoring, that is, whether R-1 or R-2 reality monitoring processes are more likely to be used. An experiential focus might be associated with using qualitative characteristics of mental experience, particularly perceptual detail or vividness as the major criteria for veridicality. An instrumental focus, in contrast, might be associated with relying more on the plausibility relations between what is remembered and what is otherwise known. Thus we might find different types of reality monitoring errors associated with differences in the types of self concept people have since such differences in self concept presumably arise from habitual modes of exercising one's cognitive architecture.

SUMMARY

According to MEM, the overall architecture of the cognitive system is an integrated configuration of perceptual and reflective processes that permits interaction as well as some degree of independence among subsystems (P-1, P-2, R-1, R-2) defined in terms of these processes. Normally, subsystems work together and account for our ability to engage in complex tasks, many of which (e.g., playing tennis, writing research papers, and responding appropriately to the feelings of others) require learning in more than one subsystem (Johnson, 1983). At the same time, the potential for MEM's subsystems to work without reference to each other (i.e., from different cues) helps explicate phenomena such as dissociations among memory measures (Johnson, 1983), variations in cognitive contributions to emotional responses (Johnson & Multhaup, in press), and selective disruption of memory such as occurs in anterograde amnesia patients (Johnson, 1990; Johnson & Hirst, in press).

Here, I have focused on one aspect of cognition that arises from the MEM architecture, namely on the possibility that information from perceptual and reflective subsystems might be confused, and on the reality monitoring processes that help offset this potential confusion (Johnson, 1988, in press; Johnson & Raye, 1981). There are two major classes of reality monitoring processes. Processes controlled by R-1 make relatively quick attributions about the origin of information based on appraising the qualitative characteristics of memories such as perceptual and contextual detail and information about cognitive operations that is available. Reality monitoring processes controlled by R-2 engage in retrieval and evaluation of additional information and consider such things as plausibility in light of antecedents, consequences, and general world knowledge. To illustrate empirical work on reality monitoring, I described research indicating that records of reflective cognitive operations later serve as evidence that we were the source of information in memory.

Finally, I suggested that the idea of a self derives from the MEM architecture via several mechanisms. A self is a byproduct of reality monitoring processes that distinguish perceptually-derived from reflectively-generated information. That is, self-as-source is a category that emerges from the reflective/perceptual dichotomy. Second, self as a phenomenal experience is associated with the mental activity of reflective control and monitoring, especially, from interactions between R-1 and R-2 in which they recruit and regulate each other. Third, this capacity for agency is embedded within an overall cognitive system that does not depend on reflective control and monitoring for all its critical activities. Consequently, in addition to an instrumental self that arises from records of reflective activity, an experiential self arises from records of what we have perceived. Thus, as in the case of other cognitive products, selves may vary in the relative contributions that perceptual and reflective processes make to their character.

REFERENCES

Bandura, A. (1982). The self and mechanisms of agency. In J. Suls (Ed.),*Psychological perspectives on the self, vol. 1*. Hillsdale, NJ: Erlbaum.

Beahrs, J. O. (1982). *Unity and multiplicity: Multilevel consciousness of self in hypnosis, psychiatric disorder and mental health*. New York: Brunner/Mazel.

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology, vol. 6*. New York: Academic Press.

Bentall, R. P. (1990). The illusion of reality: A review and integration of psychological research on hallucinations. *Psychological Bulletin, 107*, 82-95.

Cermak, L. S. (Ed.). (1982). *Human memory and amnesia*. Hillsdale, NJ: Erlbaum.

Corballis, M. C., & Roldan, C. E. (1975). Detection of symmetry as a function of angular orientation. *Journal of Experimental Psychology: Human Perception & Performance, 1*, 221-230.

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities* (pp. 409-422). Amsterdam: North Holland.

Durso, F. T., & Johnson, M. K. (1980). The effects of orienting tasks on recognition, recall, and modality confusion of pictures and words. *Journal of Verbal Learning and Verbal Behavior, 19*, 416-429.

Eich, J. E. (1975). State-dependent accessibility of retrieval cues in the retention of a categorized list. *Journal of Verbal Learning and Verbal Behavior, 14*, 408-417.

Finke, R. A., Johnson, M. K., & Shyi, G. C.-W. (1988). Memory confusions for real and imagined completions of symmetrical visual patterns. *Memory &Cognition, 16*, 133-137.

Finke, R. A., & Shepard, R. N. (1986). Visual functions of mental imagery. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance, vol. 2*. New York: Wiley-Interscience.

Flavell, J. H. (1985). *Cognitive development* (2nd Edition). Englewood Cliffs, NJ: Prentice Hall.

Foley, M. A., Durso, F. T., Wilder, A., & Friedman, R. (in press). Developmental comparisons of explicit vs. implicit imagery and reality monitoring. *Journal of Experimental Child Psychology*.

Gergen, K. J. (1977). The social construction of self-knowledge. In T. Mischel (Ed.), *The self: Psychological and philosophical issues*. Totowa, NJ: Rowman & Littlefield.

Gergen, K. J. (1982). From self to science: What is there to know? In J. Suls (Ed.), *Psychological perspectives on the self, vol. 1* (pp. 129-149). Hillsdale, NJ: Erlbaum.

Greenwald, A. G. (1982). Is anyone in charge? Personalysis versus the principle of personal unity. In J. Suls (Ed.), *Psychological perspectives on the self, vol. 1* (pp. 151-181). Hillsdale, NJ: Erlbaum.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology, 108*, 356-388.

Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist, 39*, 1372-1388.

Hashtroudi, S., & Parker, E. S. (1986). Acute alcohol amnesia: What is remembered and what is forgotten. In H. D. Cappell, F. B. Glaser, Y. Israel, et. al. (Eds.), *Research advances in alcohol and drug problems, vol. 9*. New York: Plenum.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94*, 319-340.

Hilgard, E. R. (1977). *Divided consciousness: Multiple controls in human thought and action*. New York: Wiley.

Hirst, W., Johnson, M. K., Kim, J. K., Phelps, E. A., Risse, G., & Volpe, B. T. (1986). Recognition and recall in amnesics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 445-451.

Hirst, W., Johnson, M. K., Phelps, E. A., & Volpe, B. T. (1988). More on recall and recognition in amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 758-762.

Horowitz, M. J. (1978). *Image formation and cognition* (2nd Edition). New York: Appleton-Century-Crofts.

James, W. (1915). *Psychology, a briefer course*. New York: Holt.

Johnson, M. K. (1977). What is being counted none the less? In I. M. Birnbaum & E. S. Parker (Eds.), *Alcohol and human memory* (pp. 43-57). Hillsdale, NJ: Erlbaum.

Johnson, M. K. (1983). A multiple-entry, modular memory system. In G. H. Bower (Ed.), *The psychology of learning and motivation, 17* (pp. 81-123). New York: Academic Press.
Johnson, M. K. (1985). The origin of memories. In P. C. Kendall (Ed.), *Advances in cognitive-behavioral research and therapy, (Vol. 4)* (pp. 1-26). New York: Academic Press.
Johnson, M. K. (1988a). Discriminating the origin of information. In T. F. Oltmanns, & B. A. Maher (Eds.), *Delusional beliefs: Interdisciplinary perspectives* (pp. 34-65). New York: John Wiley & Sons.
Johnson, M. K. (1988b). Reality monitoring: An experimental phenomenological approach. *Journal of Experimental Psychology: General, 117,* 390-394.
Johnson, M. K. (1990). Functional forms of human memory. In J. L. McGaugh, N. M. Weinberger, & G. Lynch (Eds.), *Brain organization and memory: Cells, systems and circuits* (pp. 106-134). New York: Oxford University Press.
Johnson, M. K. (in press). Reality monitoring: Evidence from confabulation in organic brain disease patients. In G. Prigatano & D. L. Schacter (Eds.), *Awareness of deficit after brain injury.* New York: Oxford University Press.
Johnson, M. K., Finke, R. A., Danzer, A., & Shyi, G. C.-W. (in preparation). Ease of imaging and reality monitoring.
Johnson, M. K., Foley, M. A., & Leach, K. (1988). The consequences for memory of imagining in another person's voice. *Memory & Cognition, 16,* 337-342.
Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General, 117,* 371-376.
Johnson, M. K., & Hirst, W. (in press). Processing subsystems of memory. In R. G. Lister & J. J. Weingartner, *Perspectives in cognitive neuroscience.* New York: Oxford University Press.
Johnson, M. K., Kahan, T. L., & Raye, C. L. (1984). Dreams and reality monitoring. *Journal of Experimental Psychology: General, 113,* 329-344.
Johnson, M. K., & Kim, J. K. (1985). Recognition of pictures by alcoholic Korsakoff patients. *Bulletin of the Psychonomic Society, 23,* 456-458.
Johnson, M. K., Kim, J. K., & Risse, G. (1985). Do alcoholic Korsakoff patients acquire affective reactions? *Journal of Experimental Psychology: Learning, Memory, and Cognition, ll,* 22-36.
Johnson, M. K., & Multhaup, K. S. (in press). Emotion and MEM. In S. A. Christianson (Ed.), *Handbook of emotion and memory.*
Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88,* 67-85.
Johnson, M. K., Raye, C. L., Foley, H. J., & Foley, M. A. (1981). Cognitive operations and decision bias in reality monitoring. *American Journal of Psychology, 94,* 37-64.
Johnson, M. K., Raye, C. L., Hasher, L., & Chromiak, W. (1979). Are there developmental differences in reality-monitoring? *Journal of Experimental Child Psychology, 27,* 120-128.
Johnson, M. K., Raye, C. L., Wang, A. Y., & Taylor, T. H. (1979). Fact and fantasy: The roles of accuracy and variability in confusing imaginations with perceptual experience. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 229-240.
Johnson, M. K., Taylor, T. H., & Raye, C. L. (1977). Fact and fantasy: The effects of internally generated events on the apparent frequency of externally generated events. *Memory & Cognition, 5,* 116-122.
Kihlstrom, J. F. (1987). The cognitive unconscious. *Science, 237,* 1445-1452.
Kosslyn, S. M. (1980). *Image and mind.* Cambridge, MA: Harvard University Press.
Kunzendorf, R. G. (1986). Hypnotic hallucinations as "unmonitored" images: An empirical study. *Imagination, Cognition and Personality, 5,* 255-270.
Kunzendorf, R. G. (1987). Self-consciousness as the monitoring of cognitive states: A theoretical perspective. *Imagination, Cognition and Personality, 7,* 3-22.
Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist, 41,* 954-969.
Markus, H., & Sentis, K. (1982). The self in social information processing. In J. Suls (Ed.), *Psychological perspectives on the self, vol. 1* (pp. 41-70). Hillsdale, NJ: Erlbaum.
Miller, G. A., Galanter, E., & Pribram, K. A. (1960). *Plans and the structure of behavior.* New York: Holt, Rhinehart & Winston.
Moscovitch, M. (1985). Memory from infancy to old age: implications for theories of normal and pathological memory. *Annals of the New York Academy of Sciences, 444,* 78-96.
Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and some new findings. In G. Bower (Ed.), *The psychology of learning and motivation.*

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro, *Consciousness and self-regulation* (pp. 1-18). New York: Plenum.

Perlmutter, M. (1984). Continuities and discontinuities in early human memory paradigms, processes, and performance. In R. Kail, & N. E. Spear (Eds.), *Comparative perspectives on the development of memory* (pp. 253-284). Hillsdale, NJ: Erlbaum.

Rabinowitz, J. C. (1989). Judgments of origin and generation effects: Comparisons between young and elderly adults. *Psychology and Aging, 4*, 1-10.

Schacter, D. L., & Moscovitch, M. (1984). Infants, amnesics, and dissociable memory systems. In M. Moscovitch (Ed.), *Infant memory* (pp. 173-216). New York: Plenum Press.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127-190.

Snyder, M., & Campbell, B. H. (1982). Self-monitoring: The self in action. In J. Suls (Ed.), *Psychological perspectives on the self, vol. 1* (pp. 185-207). Hillsdale, NJ: Erlbaum.

Stuss, D. T., & Benson, D. F. (1986). *The frontal lobes*. New York: Raven Press.

Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.

Weinstein, A. (1987). *Preserved recognition memory in amnesia*. Unpublished doctoral dissertation, State University of New York at Stony Brook.

Wicklund, R. A. (1982). How society uses self-awareness. In J. Suls (Ed.), *Psychological perspectives on the self, vol. 1* (pp. 209-230). Hillsdale, NJ: Erlbaum.