

In Reder, L.M. (Ed.) Implicit  
Memory and Metacognition, 1996  
Mahwah, NJ: Erlbaum

Chapter **3**

## Metacognition Does Not Imply Awareness: Strategy Choice Is Governed by Implicit Learning and Memory

Lynne M. Reder  
Christian D. Schunn  
*Carnegie Mellon University*

Metacognition means different things to different people and is generally acknowledged to include a wide range of phenomena. Nonetheless, there are two core meanings of the term metacognition to which most researchers using that label often refer: monitoring and control of cognitive processes. Monitoring of cognitive processes can include awareness of the component steps in cognitive processes as well as awareness of various features of these steps including their duration and their successfulness. For example, one might be aware of the steps one goes through in serving a tennis ball, as well as the successfulness of the serve.

Monitoring typically refers to awareness of the features of the current behavior. In contrast, control of cognitive processes refers to the processes that modify behavior, such as the selection of a strategy for performing a task. For example, deciding whether to search for a phone number in memory or search for it in a phone book, as well as deciding how long to search memory before giving up, are instances of control processes. In this chapter we focus on the relationship between monitoring and control of cognition in a special way: We argue that some aspects of metacognition typically called monitoring, and therefore implying awareness, actually operate without much awareness. Moreover, the control processes that operate to affect strategy choice are frequently influenced by implicit processes.

The relationship between the two forms of metacognition is particularly interesting. One possibility is that control of cognitive processes occurs through explicit monitoring of cognitive processes. Although this assumption is not frequently stated explicitly, it is clearly a very commonly held

assumption—the very justification for studying metacognitive monitoring is that it is believed to be necessary for control. Further, this view has been clearly articulated by several of the main researchers of metacognitive behavior. For example, Nelson and Narens (1994) described people as “systems containing self-reflective mechanisms for evaluating (and reevaluating) their progress and for changing their ongoing processing” (p. 7). Similarly, Metcalfe (1994) noted that, “Most researchers agree that the human episodic memory system requires, for its optimal functioning, a subsidiary monitoring and control system” (p. 137). These assumptions can also be found in more developmentally oriented work on metacognition. For example, Davidson, Deuser, and Sternberg (1994) stated that, “Metacognition, or knowledge of one’s own cognitive processes, guides the problem-solving process and improves the efficiency of this goal-oriented behavior” (p. 207).

In contrast to this commonly shared assumption about the relationship between monitoring and control, we believe that the opposite is true. That is, we believe that the control of cognitive processing is not achieved through explicit monitoring. Instead, we believe that the control of cognitive processing is primarily achieved through *implicit learning* and *implicit memory*.

We define *implicit learning* to be changes in behavior based on past experience for which the individual has no reportable awareness of such learning (cf. A. S. Reber, 1989; this is not to say that there was no memory or awareness of the events during which learning occurred). For example, in a probability learning experiment by A. S. Reber and Millward (1971), subjects learned to anticipate the changing probabilities of events in the experiment without any explicit awareness of these changing probabilities. In their experiment, subjects were given a training phase in which they observed 1,000 events at the rate of 2 per second. During this training phase, the probability that a certain event would occur changed gradually in a cyclical fashion every 50 trials. Subjects learned to anticipate these shifts: In the test phase they were quite successful at predicting the changing probability of these events over time. However, subjects were unaware of this shifting probability. Similar results of learning without the ability to report what was learned or any awareness that there was learning have been obtained by Berry and Broadbent (1984), and Lewicki, Hill, and Bizot (1988). In our view, this phenomenon of implicit learning extends to control of cognitive processing such as strategy selection. That is, we believe that strategy choice is influenced by past experience and that the individual is frequently unaware of these influences.<sup>1</sup>

<sup>1</sup>Our intention is not to develop a philosophical argument concerning the nature of consciousness. Nonetheless, our definition of nonconscious or implicit learning/memory does involve lack of reportability. We do not wish to imply, however, that an organism that is somehow incapable of ever reporting an event is incapable of having conscious awareness. Rather, we mean to imply that for someone generally capable of reporting events, the inability to report an event is evidence that it was not explicitly learned.

Like implicit learning, we define *implicit memory* to be changes in behavior that are based on past experience for which the individual has no awareness of such influence. The distinctions between implicit learning and memory are that (a) the experiences influencing memory tend to be discrete or even unitary (as opposed to continuous) events and there tend not to be interdependencies or contingencies among these events, and (b) the behavior that is affected in implicit memory is also something typically ascribed to declarative memory rather than procedural memory (Anderson, 1976; Cohen & Squire, 1980). For example, in an experiment by Warrington and Weiskrantz (1970), amnesics and controls were given two tests of explicit memory (free recall and recognition) and two tests of implicit memory (word fragment identification and word stem completion). As expected, they found that the amnesics performed much worse on the explicit memory tests. However, the amnesics performed at equal levels with the controls on the implicit memory tests, indicating that there could be intact memory without awareness. Subsequent experiments with normals have found other demonstrations of dissociations between implicit and explicit memory. For example, Jacoby (1983) found when subjects *generate* a to-be-remembered item (say “cold” when asked for the opposite of *hot*), their retention on explicit tests is better than when they merely study the word (read the word (COLD); however, implicit tests did not show that advantage. Testing the issue of awareness more directly in normals, Bowers and Schacter (1990) divided normal subjects by their self-reports of whether they noticed any relationship between the words they produced on the target perceptual identification task and the words studied earlier. They found that there were equal levels of priming across the two groups of subjects (although there were other differences in types of responses made) indicating that the priming effect could occur without awareness.

It is important to distinguish between the strategy-selection processes and the strategies themselves. Although we argue that people are unaware of what causes them to select one strategy rather than another, we make no claims about their awareness of the results of the strategy selections. People are often aware of the strategy that they execute. For example, people might not be aware of what led them to decide to use a calculator, but usually they are aware that they decided to use the calculator. On the other hand, when the processes or strategies are executed rapidly, it is quite possible that people are even unaware of the strategy that was selected.

Contrary to commonly held assumptions concerning metacognition, we will show that important aspects of metacognition are implicit. In particular, this chapter focuses on the implicit nature of strategy selection. Previous research has shown (e.g., Lovett & Anderson, in press; Miner & Reder, 1994; Reder, 1987, 1988) that two types of variables influence strategy selection: features contained in the question or problem, and variables

outside the problem or question. These influences are sometimes referred to as *intrinsic* and *extrinsic* variables. An example of an intrinsic variable is the familiarity with terms in a problem. Task instructions and what strategy has been working recently are examples of extrinsic variables.

This chapter is organized around the distinction between intrinsic and extrinsic variables, summarizing empirical evidence for the implicit nature of each kind of variable. In the case of intrinsic variables, we describe a computational model that implements these variables using implicit processes. We also consider the evidence that suggests that people may be completely unaware of the entire strategy selection process. Finally, we discuss why strategy selection is driven by implicit processes. This issue is especially important given the general intuition that control of cognitive processes should be the result of the monitoring of cognitive processes.

### INTRINSIC METACOGNITIVE VARIABLES AFFECTING STRATEGY CHOICE

It is generally accepted that superficial features of a problem influence the strategy that is chosen to solve the problem. The work of Hinsley, Hayes, and Simon (1977) described in an earlier Carnegie Symposium showed that people will be influenced to select a particular strategy to solve an algebra word problem when features of the problem prime the particular strategy or problem solving schema even though in fact it is an incorrect schema. More recently, Ross (1984, 1989; Blessing & Ross, 1996) has shown in many experiments that "reminders" will influence the method selected to solve statistics problems and various other tasks. Similarly, Lovett and Anderson (in press) have shown that strategy selection is influenced by features in the problem.

On the other hand, the claim that people select a question-answering strategy prior to executing any strategy, specifically prior to searching for the answer (Reder, 1982) has not been generally accepted (e.g., LeFevre, Greenham, & Waheed, 1993; Siegler, 1989; Siegler & Jenkins, 1989; Siegler & Shrager, 1984). Even more controversial has been the claim that one of the criteria used in the decision of whether to search for the answer is a rapid "feeling-of-knowing" and that this rapid feeling-of-knowing depends *only* on features of the question and not at all on partial retrievals of the answer (Reder, 1987, 1988; Reder & Ritter, 1992).

The early evidence for this position came from experiments using a "gameshow" paradigm where subjects were told to imagine that they were competing against another contestant and that in order to have an opportunity to answer questions, they had to respond as rapidly as possible with an immediate impression of whether or not they knew the answer to the

question. Subjects were able to make this assessment of their knowledge very quickly and accurately. Importantly, subjects' impressions of whether or not they knew the answer could be subverted by making pairs of terms from the question spuriously familiar—some question terms had been part of an earlier rating task that was nominally unrelated to the gameshow task. Although exposure to the terms increased the tendency to think that the answer was known, it did not influence the ability to produce the correct answer.

In a subsequent series of arithmetic gameshow experiments, we explored this phenomenon more carefully, controlling for prior knowledge, and tracked how learning and feeling-of-knowing changed as a function of exposure to problems (Reder & Ritter, 1992). Arithmetic problems that people were unlikely to know before the experiment (e.g.,  $37 * 23$ ) were used as stimuli. Subjects were exposed to problems over and over (up to 28 times in one experiment; in others, up to 20 times) and each time made a very rapid assessment of whether they knew the answer. The subjects indicated their assessment by pressing one key if they thought they could retrieve the answer, and a different key if they thought they had to calculate the answer. This quick assessment took place in about a half second and was too little time to actually retrieve the answer. Subjects were allowed to study the answer to the problem after each trial and were given incentives to learn the answers and to select retrieve. However, there were disincentives for selecting retrieve if the answer was not known. *Know* was operationally defined as giving the correct answer within one second *after* selecting retrieve.

Evidence that these first impressions were based on aspects of the question, and not a partial retrieval of the answer, came from several results of the experiments. First, time to give the rapid first impression (where subjects choose retrieve or calculate) was affected by practice with the task, but not by practice with a specific problem. In other words, the practice with a specific problem that led to faster answers did not lead to faster preliminary judgments. The second result came from operator switch problems—some of the problems were distorted such that the operator was switched so that the two operands had appeared together before but not with that operator. For example, if  $21 + 35$  was presented earlier, then  $21 * 35$  could appear as a special operator switch trial. For such a problem, the subject could not know the answer because it had not been previously presented, and its answer was unrelated to the answer of the original problem. However, such a problem would still look familiar as a first impression. In fact, it was operand co-occurrence that predicted retrieve judgments, not how often the exact problem had been seen. In other words, subjects were as likely to think they knew the answer to a problem that they had never seen before as one that they had seen 20 times, provided that only the

operator had been switched. By contrast, problems where the order of operands was inverted (e.g.,  $28 + 13 \rightarrow 13 + 28$ ), to which subjects consequently still knew the answer, were much less likely to be selected for retrieve than novel, operator-switch problems, for which they did not know the answer (Reder & Ritter, 1992).

A third result that supports the view that a rapid feeling-of-knowing is a function of exposure to the problem and not of knowledge of the answer comes from Experiment 1 of Schunn, Reder, Nhouyvanisvong, Richards, and Stroffolino (in press). In this experiment sometimes subjects were exposed to problems without getting a chance to actually answer the problem (either by calculation or retrieval) and without having the opportunity to study the answer. This manipulation was done for only a subset of the problems, called *infrequently-answered problems*. As one would expect, speed and accuracy in producing the answers were affected by how often subjects studied and answered the questions; on the other hand, tendency to select retrieve was not affected by those variables, but was affected by exposure to the problem itself. In other words, when we controlled for exposure time to the problem and decoupled exposure to the problem with exposure to the answer, the former predicted feeling-of-knowing and the latter predicted actual knowing. Figure 3.1 displays the percentage use of the retrieval strategy as a function of the total exposure to the problem.<sup>2</sup>

### An Activation-Based Model for Rapid Feeling-of-Knowing

There have been a number of other results, both ours and others' (e.g., Metcalfe, Schwartz, & Joaquim, 1993; Schwartz & Metcalfe, 1992), that support the claim that feeling-of-knowing is influenced exclusively by terms in the question. Our claim, however, focuses on feeling-of-knowing as it affects strategy selection, and is restricted to the rapid feeling-of-knowing that precedes execution of the retrieval strategy. Given that feeling-of-knowing, like strategy selection, tends to be thought of as the essence of a metacognitive strategy, it is important to defend our claim that this rapid feeling-of-knowing is actually an *implicit process* rather than an *explicit process*.

There are several arguments that can be made for this claim. One of these is that the process model that we use to account for this phenomenon has an implicit quality to it. The decision-making process involves rapid and automatic flow of activation rather than slow and controlled decision making about discrete features in the environment. This type of model is arguably just the type of model that can also account for other memory phenomena that are typically described as aspects of implicit memory (e.g., false fame judgments, subliminal priming causing spurious word recognition [Reder &

<sup>2</sup>The last three points for the infrequently-answered problems are unstable because they have a mean of 19 observations per point, whereas the preceding points have a mean of 150 observations per point.

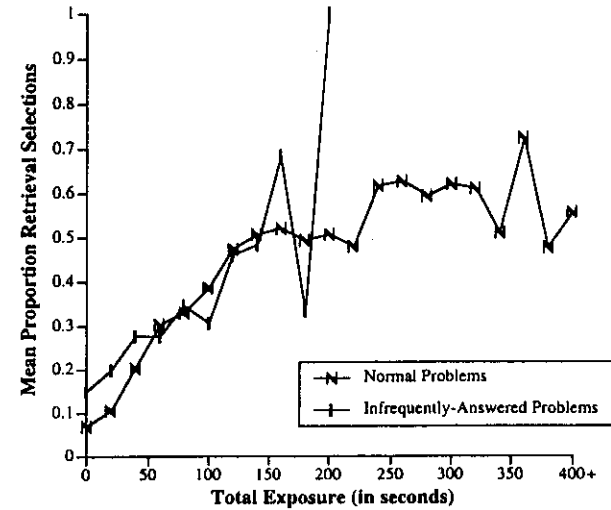


FIG. 3.1. The mean percentage of retrieval strategy selections as a function of total exposure to the problem (in seconds) for Experiment 1 of Schunn et al. (in press). Adapted from Schunn et al. (in press). Adapted with permission.

Gordon, in press]). Below, we describe an activation-based model that can account for these results.

### The SAC Model

Reder's model which stands for *Source of Activation Confusion* was developed to account for misattributions and cognitive illusions (Reder & Gordon, 1996; see also Kamas & Reder, 1994). The model bears a family resemblance to Anderson's (1983, 1993) ACT model, although SAC is concerned primarily with declarative memory, and makes slightly different representational and processing assumptions. The representation used by SAC can be thought of as a rather generic semantic network model of declarative memory. The representation consists of interassociated nodes representing concepts and varying in long term strength. For the simulations described here, there are nodes that represent numbers (e.g., 5, 17, 31), nodes that represent operators (e.g., +, /, \*), and nodes that represent whole problems (e.g.,  $17 + 31$ ). The nodes representing whole problems connect the operands and operators to the answers. Nodes that represent numbers may serve as operand nodes for some problems and answer nodes for other problems (e.g., 31 is an operand in the problem  $23 * 31$ , and is also the answer to  $14 + 17$ ). Examples of this representation scheme can be seen in Fig. 3.2.

The strength of a node represents the history of exposure to that concept, with more exposure producing greater strengthening. Nodes that

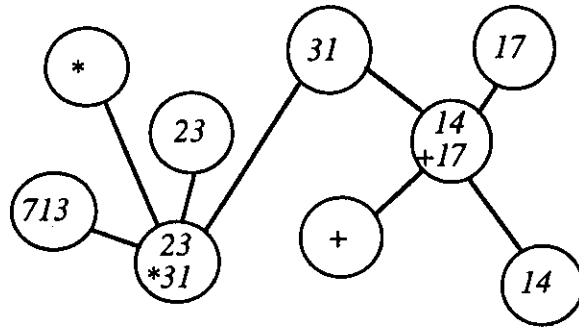


FIG. 3.2. Example nodes in the SAC network.

represent arithmetic problems such as  $27 * 34$  would start out weak at the beginning of the experiment, as these problems were initially unfamiliar to the subjects. By contrast, nodes for familiar problems such as  $12 * 12$  would be strong even at the beginning of the experiment. However, we did not use problems that were likely to have pre-experimental familiarity, and the simulations presented here assume that all problem nodes are created for the first time during the experiment.

Strength can also be thought of as the baseline or resting activation level of a node. Increases and decreases in this baseline strength change according to a power function:

$$B = c \sum t_i^{-d} \quad (1)$$

where  $B$  is the base-level activation,  $c$  and  $d$  are constants, and  $t_i$  is the time since the  $i^{\text{th}}$  presentation. This function captures both power law decay of memories with time, and power law learning of memories with practice.<sup>3</sup> The central feature of power law decay is that memories decay quickly initially and then much more slowly at increasing delays. Similarly, the central feature of power law learning is that first exposures to an item contribute more than do subsequent exposures. That is, the incremental contributions of each new exposure decreases with increasing numbers of exposures.

The base or resting activation level of a node should be distinguished from the *current activation* values of a node. The current level of a node will be higher than its baseline whenever it receives stimulation from the environment; that is, when the concept is mentioned or perceived, or when the concept receives activation from other nodes. Whereas baseline

<sup>3</sup>See Anderson and Schooler (1991) for a discussion of the evidence for this function in learning and retention phenomena.

strength decays according to a power function (i.e., first quickly and then slowly), *current* activation decays rapidly and exponentially toward the base level. Let  $A$  represent the current level of activation and  $B$  represent the base level of activation. Then, the decrease in *current* activation will be:

$$\Delta A = -\rho (A - B) \quad (2)$$

such that, after each trial, the current activation will decrease for every node by the proportion  $\rho$  times that node's current distance from its base-level activation. In all of our simulations, we used a value of 0.8 for  $\rho$ . To present a concrete example, suppose after a trial, a node's base-level activation was 60. Then after just one trial, the current activation was 20 and its current activation was 28; that is,  $60 - .8 * (60 - 20)$ . After three trials, the current activation would have dropped to 20.3, not significantly different than the resting activation of 20. Thus, current activation drops quite rapidly, and only has noticeable effects on the trial on which it became activated, and perhaps the trial immediately thereafter.

Activation spreads between nodes via links. Links connect nodes that are associated through conceptual relations. For example, links connect nodes that represent the components of a problem to the node that represents the entire problem. Links also connect the nodes that represent the entire problem with nodes that represent the answer. These links will vary in strength depending on how often the two concepts have been thought of at the same time. Strength of links also depends on the delay between exposures. Specifically, we assume a power function given by:

$$S_{s,r} = \sum t_i^{-d_i} \quad (3)$$

where  $S_{s,r}$  is the strength of the link from the node  $s$  to node  $r$ ,  $t_i$  is the time since the  $i^{\text{th}}$  coexposure, and  $d_i$  is the decay constant for links.

We said earlier that the current activation level of a node can rise from environmental stimulation or from associated nodes that send activation to it. How much activation is sent depends on the activation level of the source (sending) node and on the strength of the link from the source node to the receiving node, relative to all competing links out of the same source node. The change in activation of some node  $r$  is computed by summing the spread of activation from all source nodes  $s$  connected to node  $r$  according to the equation:

$$\Delta A_r = \sum (A_s * S_{s,r} / \sum S_{s,r}) \quad (4)$$

where  $\Delta A_r$  is the change in activation of the receiving node  $r$ ,  $A_s$  is the activation of each source node  $s$ ,  $S_{s,r}$  is strength of the link between nodes

s and r, and  $\Sigma S_{s,i}$  is sum of the strengths of all links emanating from node s. The effect of the ratio  $S_{s,r}/\Sigma S_{s,i}$  is to limit the total spread from a node s to all connected nodes to be equal to the node s's current activation  $A_s$ .<sup>4</sup> For example, if a node had three connections emanating from it with link strengths of 1, 2, and 3, then the activation spread along those links would be, respectively, 1/6, 1/3 (i.e., 2/6), and 1/2 (i.e., 3/6) of the node's current activation level.

With these few assumptions, we can provide an account of the rapid feeling-of-knowing responses if we also assume that feeling-of-knowing monitors intersection of activation from two source nodes. Specifically, when two terms in a question send out activation to associated concepts and an intersection of activation is detected by bringing an intermediate node over threshold, a person will have a feeling-of-knowing response (cf. Doshier & Rosedale, 1989, 1990; Glucksberg & McCloskey, 1981; Ratcliff & McKoon, 1988; Reder, 1979, 1987, 1988, for related treatments of intersection of activation).

For present purposes, we assume that the nodes corresponding to the operands and the operator in the problem are activated when the problem is presented (e.g., 17, 31, and +). Activation spreads from these nodes to the node that represents the entire problem (e.g., 17 + 31). The extent of activation that accumulates at the problem node affects the likelihood of selecting retrieve as the strategy of choice.

#### Implementational Details: Generating Predictions From the Model

This model can be used to predict feeling-of-knowing decisions (i.e., deciding between retrieval and computation). It can also predict which answers are retrieved from memory, and the speed with which the answers are retrieved. In this chapter, we focus on the feeling-of-knowing, or retrieve/compute, decisions.

In this spreading activation model, feeling-of-knowing judgments are based on the activation level of a problem node in memory after a problem has been presented. When a problem is presented, all the nodes representing the components are activated. For example, in the problem 23 \* 37, the nodes representing 23, \*, and 37 are all activated. Activation then spreads to the node representing 23 \* 37, in addition to other problem nodes connected to the nodes representing 23, \*, and 37. Feeling-of-knowing judgments are based on the activation of the most activate problem node.

<sup>4</sup>This feature gives the model the ability to simulate fan effects (Anderson, 1974, 1976, 1983; Reder & Anderson, 1980; Reder & Ross, 1983), an important feature of declarative memory.

The computer simulation is given as input the same problems presented to the subjects. Because each subject received a different set of problems in a random order, a separate simulation was conducted for each subject. This was important because on a given trial the expected activation level for a problem would vary depending on the exact sequence of trials: For any subject on a given trial, the number of links, the current activation, and strengths would be different from any other subject's values. The simulation output is a probability of selecting to retrieve on each trial. We now step through the process by which that probability is determined.

At the start of the experiment, the representation of memory for the simulation is identical regardless of the experimental stimuli to be seen. Nodes for the operands and operators are assumed to already exist, whereas nodes for the problem components are assumed not to exist (i.e., the problems are novel). For simplicity, the initial base level strengths of the operand nodes (numbers used in the problems) and operator nodes are set to a constant amount. When problems are seen for the first time, a problem node is created, as are the links from the component operand and operator nodes to the novel problem node. The initial base level strengths of the problem nodes and of the links is simply determined by the equations determining power-law growth and decay—the selection of initial strength values requires no extra constants.

We assume that on each trial, all the nodes representing the problem components are activated to some constant amount. We assume that a basic perceptual process activates these nodes, and that all the problem components (e.g., the operators and the operands) used in these experiments were familiar entities. For example, when the problem 23 \* 37 is presented, the three nodes representing 23, \*, and 37 are activated. Activation then spreads along the links emanating from nodes representing each of the problem components to other associated nodes. Activation only spreads to directly connected nodes at this point, and is not yet carried forward beyond the first layer of receiving nodes.

Once the activation has spread across these links, activation of the problem nodes can be used to make a strategy selection between retrieve and calculate (a feeling-of-knowing judgment). The activation value of the most active problem node is used. Again, these values are affected by aspects of the simulation such as those represented in Equation 4. Rather than making a binary choice, the simulation predicts a probability of choosing retrieve based on this activation value. This means that if the activation value of the most active node is low, the probability of selecting retrieve is very low; conversely, when the activation value of a node is very high, the probability of selecting retrieval is high, but not necessarily unity. This probability of choosing retrieve is calculated by assuming a normal distribution of activation values with a fixed variance and activation threshold for selecting retrieve. This probability is reflected in the formula:

$$P = N[(A - T)/Sd] \quad (5)$$

where  $A$  is the activation of the highest problem node,  $T$  is the subject's threshold,  $Sd$  is the standard deviation, and  $N[x]$  is the area under the normal curve to the left of  $x$  for a normal curve with mean = 0, and standard deviation = 1.

After each trial, all the strengths and activations are updated. It is at this point that if a new problem has been presented for the first time, that a new node representing that problem is created, and links are created connecting the component nodes to the problem node.

The simulation just described involves seven parameters, enumerated in Table 3.1. The table lists the values that were used and reviews the equations mentioned earlier. With the exception of the threshold parameter, which varied by subject, all of these parameter values were held constant across subjects and across experiments that we tested (4 data sets in all). The threshold parameter was allowed to vary across subjects because it was obvious from the data that subjects differed greatly in how often they decided to select retrieve. Although the subjects might have differed on other dimensions as well, there were no other obvious differences (with the exception of the one mentioned later), and so, for parsimony's sake, the other six parameters were held constant across subjects.

We fit the experiment using these seven parameters and also using an eighth parameter. This eighth parameter was simply a binary value by subject reflecting whether the subject had a predilection not to choose retrieve for a particular operator. We decided to include this parameter

TABLE 3.1  
SAC Model Parameter Descriptions and Values, and the Model Equations

Parameter Name	Function	Value
input-activation	Input current activation for component nodes	50
$r$	Exponential decay constant for current activation	0.8
$c$	Power-law growth constant for base level activation	5
$d$	Power-law decay constant for base level activation	0.175
$d_l$	Power-law decay constant for link strength	0.12
$T$	Retrieve/compute decision threshold	30-200
$Sd$	Retrieve/compute decision standard deviation	45
never-retrieve	Does subject decide to never retrieve for one of the operators?	T/F
<b>Equations</b>		
(1) $B = c \sum t_i^{-d}$		
(2) $\Delta A = -\rho (A - B)$		
(3) $S_{i,j} = \sum t_i^{-d}$		
(4) $\Delta A_{i,j} = \Sigma (A_{i,j} * S_{i,j} / \Sigma S_{i,j})$		
(5) $P = N[(A - T)/Sd]$		

because we found that some subjects had a strong aversion to choosing retrieve for a particular operator. For example, a few subjects never chose retrieve for problems involving the operator *sharp* (a novel operator that involved a combination of addition and multiplication). Perhaps they did not want to memorize problems that involved a fake operator. A few other subjects were found to never retrieve for multiply although they chose retrieve for *sharp* problems. These subjects may have been bothered by the modular arithmetic that was used in Reder and Ritter (1992) and in Schunn, Reder, et al. (in press) and did not want to memorize the wrong answers to multiply problems. Whatever subjects' reasons for choosing to never retrieve for an operator, we found that the eighth parameter was useful for simulating these subjects—those that seemed to use a metarule for making their decisions, in which they refused to retrieve for one of the operators.

To model these subjects, we had the simulation put the probability of choosing retrieve at zero for the operator in question. The criterion for a subject to be put in this class of having a metarule was choosing retrieve less than 5% of the time for a given operator. It should be noted that this rule was invoked only 8 times out of 58 subjects modeled. We felt it was better to use this metarule than to assign separate thresholds for problems of each operator type. Not only would this give us too many degrees of freedom, it was hardly necessary: Except for these few subjects using this metarule, the correlation between the rate of selecting retrieval for problems involving of each operator type was quite high across subjects. Finally, it is important to note that although we believe that some subjects actually employed this metarule, this feature of the simulation is not necessary to fit the data. Therefore, the fits to data without the use of this feature are also presented.

To compare the model's predictions to subjects' actual retrieve/compute decisions, we used an aggregation procedure developed by Anderson (1990). For each trial, for each subject, the model produced a probability of choosing retrieve based on the calculated activation values resulting from the trial history for that subject. That is, the probability reflected the model's experience with the exact same problems given to the subject. This probability was also based on the particular subject's threshold. Because subjects made binary decisions on each trial and the simulation produced probabilities, it was necessary to aggregate trials. That is, all trials for a given subject in which the simulation predicted that the probability of selecting retrieve would fall between 0 and 10% were grouped together; all trials where the probability fell between 10 and 20% were grouped together and so on. Next, we tabulated the actual proportion of retrieval strategy selections that were made by that subject for the exact same trials in each probability range. This was done for all probability ranges. The ranges were made sufficiently

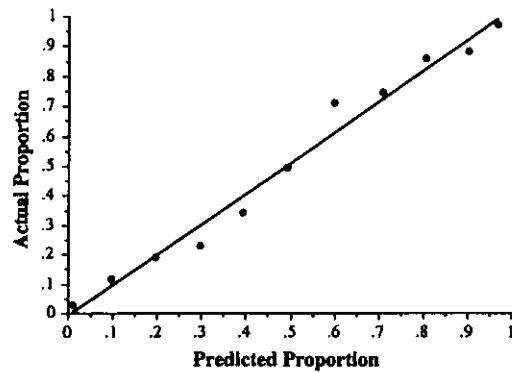


FIG. 3.3. Predicted vs. actual proportion of retrieval strategy selections in Reder and Ritter, grouped by predicted proportion for all problems.

large such that at least 5 data points would be collected in each range—thereby ensuring stable proportions. The fit of the model was tested by plotting mean actual proportion of retrieval strategy selections against mean expected percent retrievals. A perfect fit would be a straight line with a slope of 1 and a y-intercept of 0 (i.e., predicted = actual).

The model fit the data quite well, producing a Pearson's  $r$  of .990 (see Fig. 3.3).<sup>5</sup> The slope of the best fitting line was not significantly different from 1 (slope = 1.03,  $t(9) = 0.61$ ,  $p > .5$ ), nor was the intercept significantly different from 0 (intercept =  $-0.004$ ,  $t(9) = 0.15$ ,  $p > .8$ ). Without the never-retrieve rule, the model's overall fit was just as good (Pearson's  $r$  of .990).

A key result of Reder and Ritter (1992) was that subjects were as likely to select retrieve for operator-switch problems as for training problems. The model predicts this effect because operators are associated with a large number of problems (i.e., they have a very large fan) and the amount of activation spread from a node along any one link is inversely related to the number of links emanating from that node (see Anderson, 1983, for a more complete discussion of the fan effect). Thus, the model predicts that there will be little impact of switching operators on retrieve/compute decisions because the activations of the problem nodes are not significantly affected. Verifying this prediction, the fit of the model to the operator-switch retrieve data is quite good ( $r = .981$ ). Figure 3.4 presents this fit.<sup>6</sup> Again, the slope of the best fitting line was not significantly different from 1 (slope = 1.15,  $t(4) = 1.35$ ,  $p > .25$ ), nor was the intercept significantly different from 0 (intercept = 0.014,  $t(4) = 0.2$ ,  $p > .8$ ).

<sup>5</sup>Plots of individual subject values can be found in Schunn et al. (in press).

<sup>6</sup>Fewer groupings were used in this analysis because there were relatively few operator-switch problems.

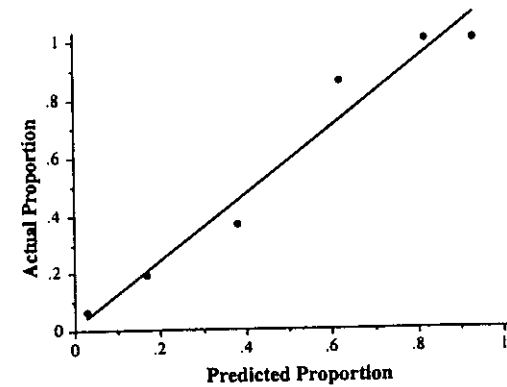


FIG. 3.4. Predicted vs. actual proportion of retrieval strategy selections in Reder and Ritter, grouped by predicted proportion, for the operator-switch problems only.

### Fitting the Model to Other Data Sets

The extraordinarily good fits of the model to the data might be partially caused by searching the parameter space to find the best values of the parameters to fit the data. For this reason it seemed important to use the same parameter estimates on several other data sets and see whether comparable fits could be obtained without searching the parameter space. We therefore fit three other data sets using the same values.

Earlier in the chapter, we referred to an experiment (Experiment 1 of Schunn et al.) that showed that rapid first impressions (i.e., rapid decisions to retrieve) were based on familiarity with the problem rather than strength of association to the answer. To review, in Experiment 1 we did not switch the operator to show this effect; rather, for some problems exposure to the problem was decoupled from learning the answer by frequently denying the subject an opportunity to solve the problem or study the answer with the problem. That experiment confirmed that the conclusions of Reder and Ritter (1992) were not based on a less than perfect encoding of the problem. That is, a possible alternative interpretation of Reder and Ritter was that subjects failed to encode the operator and just assumed that it was the operator usually seen with the operands. Because there were no operator switch problems, that possibility was ruled out. Subjects still based their decision on whether or not to retrieve on the amount of exposure to the problem, not the answer.

As noted earlier, the same parameter values were used across all the fits of the individual subjects in the Reder and Ritter (1992) experiment, with the exception of the individual subject thresholds. These same parameter values were used to fit the subject strategy choices in Experiment



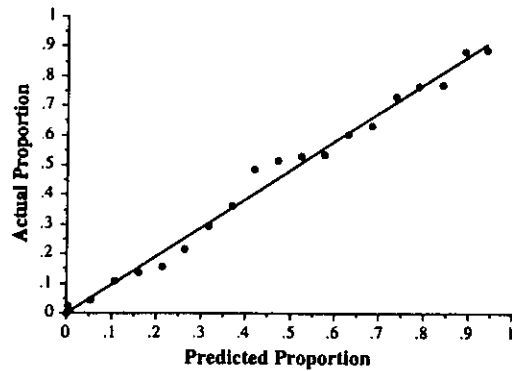


FIG. 3.5. Predicted vs. actual proportion of retrieval strategy selections in Experiment 1, grouped by predicted proportion, for all problems.

1 of Schunn et al. (in press) where the only parameter values to change were again the subject thresholds. The fit of the simulation's predictions to the subject performance is shown in Fig. 3.5. The fit to each individual subject's data is comparable and may be seen in Schunn et al. We accounted for as much experimental variance as in the first experiment that we attempted to fit. The correlation between predicted use of retrieve and actual use was  $r = .994$ . Without the never retrieve rule, the correlation was also excellent,  $r = .971$ .

In another experiment (Experiment 2 of Schunn et al.), we performed a replication of the Reder and Ritter experiments (with the operator swap) except that we had subjects come back 24 hours later. On the first day, there were no operator swaps; the second day was the same experiment as the first day, but with all new problems (novel combinations of operands used on Day 1) and some operator swap problems. The operator swap problems used operand pairs from the second day and from the first day. In other words, there were operator swap problems for both Day 1 and Day 2 problems, but all operator swap problems occurred on Day 2. Again we used the same parameter values as we had in the other model fitting efforts. In this case, we also used the same individual threshold value for a given subject on both Day 1 and Day 2. We used the same decay rate parameter to account for forgetting over 1 hour and to account for forgetting over 24 hours.<sup>7</sup> The fits for the Day 1, Day 2 data and the operator swap data for Day 1 and Day 2 problems are shown in Fig. 3.6. The correlation between the observed and predicted use of the retrieval strategy was .994, .986, 1.0, and .971 for these respective data sets.

<sup>7</sup>Specifically, we estimated passage of time by number of intervening trials in the original experiments. To remain consistent, we estimated the number of intervening trials for 24 hours. There were 300 trials in 90 minutes, so we estimated 4,800 trials in 24 hours and used that value for the 24-hour delay condition.

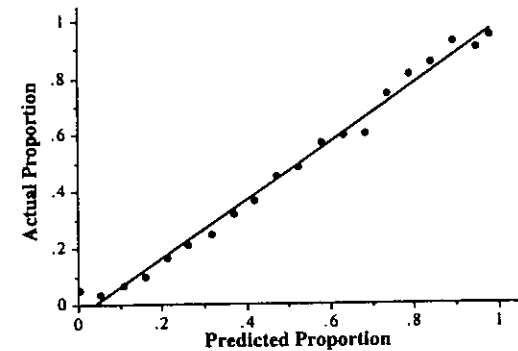


FIG. 3.6a. Predicted vs. actual proportion of retrieval selections in Experiment 2 of Schunn et al. (in press), grouped by predicted proportion, for Day 1 problems only.

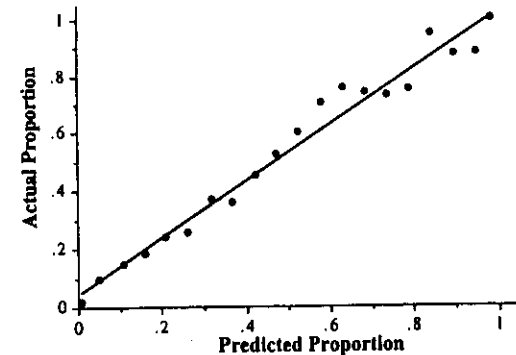


FIG. 3.6b. Predicted vs. actual proportion of retrieval selections in Experiment 2 of Schunn et al. (in press), grouped by predicted proportion, for Day 2 training problems only.

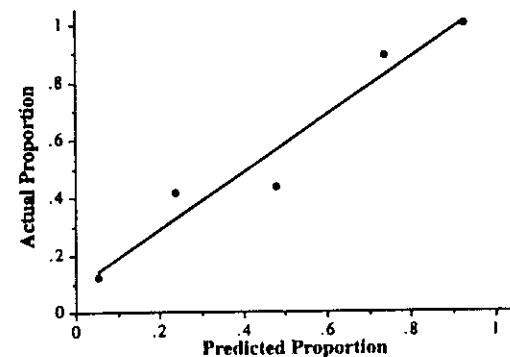


FIG. 3.6c. Predicted vs. actual proportion of retrieval selections in Experiment 2 of Schunn et al. (in press), grouped by predicted proportion, for same day operator-switch test problems only.

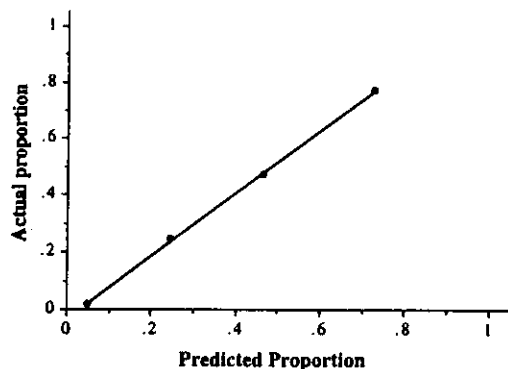


FIG. 3.6d. Predicted vs. actual proportion of retrieval selections in Experiment 2 of Schunn et al. (in press), grouped by predicted proportion, for previous day operator-switch test problems only.

### Summary of Model Fitting Enterprise

The experiments clearly showed that rapid feelings of knowing are made based on familiarity with pairs of terms from the question and are not based on partial retrieval of the answer. The SAC model seems to do an excellent job of capturing this process, fitting four data sets extremely well using only a few constant parameter values. This enterprise involved 58 subjects, each with between 250 and 500 trials to predict. Given that there were 6 parameter values that remained constant across all experiments and were estimated for only the first experiment, these fits are especially impressive. It is worth noting that the same parameter value for decay was retained even when we started fitting data that involved a 24-hour delay from acquisition to test.

Reder's SAC model also answered the computational conundrum of how people can make a rapid preliminary evaluation of whether or not it is fruitful to search memory before actually engaging in such a search: In our model, activation converges on the problem node from elements in the question; this logically precedes retrieval of the answer. Therefore, a person would easily have the opportunity to make a decision about strategy preference before execution on retrieval could begin.<sup>8</sup>

How general are the conclusions? It is conceivable that the heuristic used to make a rapid feeling-of-knowing is idiosyncratic to the task employed. That

<sup>8</sup>Although beyond the scope of the simulation, we believe that when people choose to retrieve, they simply wait for activation to continue to spread from the activated problem node to the node representing the answer. When people choose to calculate, they then shift their focus of attention to the calculation process. However, even in this case, the spread of activation to the answer nodes is not stopped; it is merely attenuated, and the activation of nodes involved in the calculation process will dominate in terms of overall activation levels.

is, we used artificial arithmetic problems and a forced deadline. On the other hand, this same finding has been shown before with general knowledge questions for which subjects are not given a specific deadline for responding. In that case, too, subjects made feeling-of-knowing judgments (estimates of whether they could answer the question) faster than they could retrieve the answer and those feelings of knowing were equally susceptible to distortions from prior exposure to terms in the question.

There are other reasons to think that the rapid feeling-of-knowing is part of the normal question-answering process. In that general knowledge experiment, there was a second group of subjects who were asked to give the answer as quickly as they could (rather than estimate whether or not they knew the answer). These subjects were affected by the spurious priming just like the estimate group, but for these subjects, the priming led to longer search times before responding "don't know." In other words, the answer group was not asked to give a rapid feeling-of-knowing and therefore the manipulation that inflated feeling-of-knowing judgments for the estimate group instead caused them to initiate more or longer searches. Another reason to think that this rapid first impression is part of the normal question answering procedure is that the total time to estimate whether they knew the answer and then retrieve the answer (measured separately and then summed) was equivalent to the total time to indicate that the answer was found (without first providing an explicit estimate) and then to speak it.

### THE IMPLICIT CHARACTER OF FEELING-OF-KNOWING

Now that we have shown how the SAC model can account for rapid feeling-of-knowing, we are left with the task of justifying why we believe rapid feeling-of-knowing to be a largely implicit process. Of course, part of the argument depends on the definition of implicit process, and by that we mean processes that are "not reportable," "not open to inspection," or that are "involved in tasks generally considered to be implicit memory tasks." Kihlstrom, Shames, and Dorfman (chapter 1, this volume) make an eloquent case for the similarity between implicit processes and feeling-of-knowing. We agree with their view that "the person's experience, thought, and action is affected by . . . [aspects] . . . which he or she has not remembered" (p. 3).

Although we agree that this rapid feeling-of-knowing is based on something unreportable, we differ as to the nature of that which is unreportable or implicit. Kihlstrom et al. are referring to subthreshold activation of the answer, and for their task, that may well be the case. However, for the task that we explore (i.e., rapid feelings of knowing), impressions are not based on a *subthreshold* activation of the answer. Rather, the judgments are based

on a *suprathreshold* activation of the *problem node*. We call the process implicit because the subject does not know why the answer seems available. The subject uses a heuristic that generally works, but high activation of a problem node does not guarantee that there is an answer associated to it nor does it guarantee that the *correct* problem node was activated.<sup>9</sup> That subjects can have their heuristic subverted by prior exposure to terms in the question or familiarity with critical parts of the problem is evidence for the implicit nature of the process. It is a seat-of-the-pants heuristic, a gut feeling that cannot be explained by the subject. The important point is that the subject is *unaware* of why the answer seems available. If the subject realized that it was because pieces of the question were familiar, the subject would not make these positive feeling-of-knowing judgments.

Indeed, one of the hallmarks of manipulations that affect implicit memory is that when subjects become aware of the manipulations, they make different attributions about their memories. For example, Jacoby and Whitehouse (1989) have found an increase in the tendency to say "old" to a word when, just prior to it, there was a brief exposure to the same word. However, this influence of a brief flash only occurred when the subjects were *unaware* of the flash. Likewise, an earlier exposure to non-famous names will tend to increase the probability that these names will later be judged famous, provided that the subject does not remember the context in which the name was seen earlier. When the person is aware of the brief flash or remembers studying the name earlier, the attribution of familiarity changes and the subject no longer has an increased tendency to respond "old" or to judge the name as famous.

In a similar vein, the gameshow experiment of Reder (1987, Experiment 6) demonstrated that subjects try to counteract experimental manipulations that affect their implicitly driven (i.e., activation-based) judgments. In this experiment, general knowledge questions had been primed by having pairs of terms from the questions rated for co-occurrence in everyday life. This earlier exposure to those terms caused subjects to have a spurious feeling-of-knowing, despite the fact that subjects seemed aware of this manipulation at least some of the time. For the more difficult the questions, the subjects were more likely to think they knew the answer if it had been primed (see Table 3.2). However, for just the easiest questions, subjects were appreciably *less* likely to judge that they knew the answer to the question if it had been primed. We take this as evidence that subjects used the strategy of raising their criterion for saying that they knew the answer when they recognized that terms in the question had been primed. However, this attempt to counter the spurious feeling-of-knowing was only partially successful. That

<sup>9</sup>Nonetheless, this mechanism is an excellent heuristic, giving *d*'s of over 2.0 and gammas (Nelson, 1984, 1986) over .8. Those values are calculated over trials that include the operator switch problems (that lead people astray), so clearly this heuristic works well overall.

TABLE 3.2  
The Proportion of Questions Attempted at Each Difficulty Level for  
Primed and Unprimed Questions in Experiment 6 of Reder (1987)

Question Difficulty	Unprimed	Primed
Easy	0.82	0.76
Moderate	0.50	0.52
Hard	0.24	0.31

Note. Adapted from Reder (1987). Adapted with permission.

is, despite the correction in threshold, there was still a residual influence of the spurious priming when the answer was not known. The strategy to counteract the spurious familiarity only served to *reduce* the number of positive feeling-of-knowing responses when the answer *was* known.

Some definitions of implicit memory focus on the notion that there is no intention to remember or retrieve. However, there are many phenomena that might be considered products of implicit memory that may involve intentional inspections of memory. Some of these have been called *illusions of memory*, and contain some inaccurate components of metamemory. For example, the false-fame results of Jacoby and his associates (Jacoby, Kelley, Brown, & Jasechko, 1989; Jacoby, Woloshyn, & Kelley, 1989), and the inaccurate judgments of learning of Dunlosky and Nelson (1992) share a quality: In all cases, subjects were influenced by prior experiences and misinterpreted memory traces when making judgments. Benjamin and Bjork (chapter 14, this volume) illustrate how people can misattribute current familiarity with a memory trace to long-term learning of the trace and vice versa. Of course, showing the lack of accuracy in metamemorial judgments such as those described by Narens, Nelson, and their colleagues is not direct support for the position that strategy choice is based on implicit memory. Nonetheless, we contend that it is the current activation level of the memory trace, which can come about by recent priming or earlier exposure, that drives the inference processes that are used for these judgments. We believe that implicit memory processes are just that—the availability heuristic (Kahneman & Tversky, 1973) driven by current activation levels (cf. Kamas & Reder, 1994).

### EXTRINSIC METACOGNITIVE VARIABLES AFFECTING STRATEGY CHOICE

A primary extrinsic variable that has been shown to influence strategy selection is the strategies' past history of success. This effect was first demonstrated by Luchins (1942) in what he termed the *Einstellung Effect*,

in which recent success with one strategy caused subjects to overlook a much easier strategy. The principles of operant conditioning (Watson, 1925) describe a more general version of the influence of success on strategy use. That is, it was hypothesized that positive reinforcement (e.g., success) leads to increased occurrence of a behavior (e.g., a strategy), and negative reinforcement (e.g., failure) leads to decreased occurrence of a behavior. More recently the importance of prior history of success has been demonstrated anew. For example, Lovett and Anderson (in press) used a problem-solving task analogous to the classic water-jug task of Luchins (1942). They found that subjects were guided in their strategy preference not only by features of the problem, but also by the recent history of success of the various strategies.

Another illustration of the adaptiveness of strategy choice to the apparent success of the strategy are the data of Rudnicky (1990). His task allowed for the use of alternative strategies to perform operations within a spreadsheet: Subjects could use either a relative movement (arrow key) to get around in a spreadsheet in order to perform a required task, or they could use an absolute movement (a *go to* command). Their preference for these two commands for performing the tasks depended on the system delay (e.g., the arrow key was not popular when there was an appreciable system delay after each keystroke) and the distance to be traversed in the spreadsheet (e.g., *go to*'s were used more often for longer distance traversals). The most remarkable aspect of the data was that the tendency to adopt a given strategy was extremely well tuned to minimize the expected wait time. Subjects used the expected total duration for an action to choose the most efficient strategy; however, it is unlikely that subjects could explicitly calculate the optimal use of these strategies in "real time."

Of course, a strategy's history of success is not the only extrinsic variable that influences strategy choice. For example, although subjects do not necessarily select the strategy that corresponds with the official task demands, task instructions nonetheless have an impact on strategy preference even when accuracy does not depend on a specific strategy (Reder, 1987, Experiment 2). Likewise, explicit advice about which strategy is likely to prove successful in a given context does not guarantee that the suggested strategy will work, but it does influence preference (Reder, 1987, Experiment 3). Our interest in the success of a strategy is that we believe that subjects are frequently *unaware* of which strategy is most successful or that they are being influenced by strategy successfulness. In the case of the Rudnicky (1990) task, it seems unlikely that subjects were in fact making conscious decisions of strategy use. Hernnstein (1990) found that even pigeons were precisely tuned to the success base rate of particular choices, so there is reason to believe that subjects can be influenced without conscious awareness. Next, we review data that support this view.

### Implicit Monitoring of the Extrinsic Variables

The extrinsic variable on which we focus our attention is the history of success with the strategy. It is of interest to us because its influence has been so frequently studied and has evidence relevant to the issue of awareness. Although it has been found that people are sometimes exquisitely sensitive to the probability of success or the relative usefulness of the various available strategies, we have reason to believe that this sensitivity is not in conscious awareness.

Three quite different experiments (Lemaire & Reder, 1996; Reder, 1987; Zbrodoff & Logan, 1986) illustrate subjects' adaptive use of strategies without even any awareness of the use of one of the strategies, let alone the relative use (proportional use) of the strategies. In Zbrodoff and Logan's experiment, subjects were given addition problems to verify or reject, such as  $8 + 4 = 13$ . Subjects made many errors when given problems in which the answer was the correct answer to the corresponding multiplication problem, such as  $8 + 4 = 32$ . The subjects' tendency to make these errors depended on the relative proportion of such foils: The greater the proportion of foils, the less the number of errors. That subjects are sensitive to the relative proportion of such foils is not terribly noteworthy for this chapter except to note that subjects were totally unaware that they were especially prone to making these kinds of errors!

In Experiment 1 of Reder (1987), subjects were extremely adaptive to the relative proportion of statements to be verified that had actually been presented as part of the previously read stories, yet totally oblivious to these proportions or even the strategies that they were using. In that experiment, subjects read a series of stories and were asked to make plausibility judgments for statements about the story. For one group of subjects, there was an 80% probability that a statement to be verified had been presented earlier as part of the story. In the case of false statements, it was an 80% probability that the exact contradiction of the statement had been presented earlier in the story. This condition was called the Direct Retrieval Bias condition because direct retrieval was likely to work. For another group of subjects, the probabilities were 20% instead of 80%. This condition was called the Plausibility Bias condition because most answers could only be made using plausibility. After reading six stories and answering questions after each story, the ratio of previously-presented to not-previously-presented statements became 50% for both groups (where previous presentation was not predictive of truth because the exact contradiction also counts as a previous presentation). Figure 3.7 illustrates subjects' sensitivity to the probability of success with a strategy. The y-axis plots the difference in latency between statements that had been previously presented or not, or in other words, the relative advantage of having seen the statement

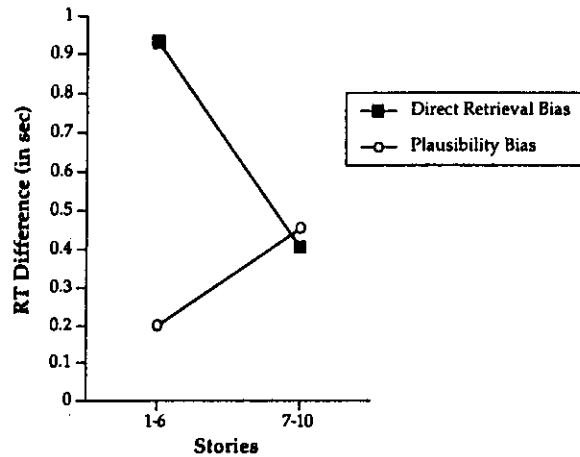


FIG. 3.7. Difference in reaction time between stated and nonstated statements (collapsed over plausibility) as a function of strategy bias, when bias imposed (Stories 1-6) and not imposed (Stories 7-10) for Experiment 1 of Reder (1987).

earlier. We take this difference as an index of the proportional use of the retrieve strategy as opposed to the plausibility strategy. Note that the size of the difference is much larger for the group where the retrieval strategy would work 80% of the time. The size of the difference becomes comparable for the two groups when the proportions become the same for the two groups.

Again, like in the Zbrodoff and Logan (1986) experiment, what makes these results especially noteworthy is that subjects were totally unaware of which strategy they were using more often, what the base-rate frequency was of presented statements in the story, whether that proportion changed during the experiment, or in which direction. Yet despite this inability to report these frequencies or to appreciate what strategies they were using, the data clearly indicate that they were very adaptive in their strategy use.

Lemaire and Reder (1996) have also found this dissociation between influence of the manipulation and awareness of the manipulation. Subjects were given multiplication problems to verify or reject, such as  $8 * 4 = 33$ . Subjects' latencies to reject incorrect answers were slower the closer in value the answer to the correct answer, with one important qualification: Subjects were faster to reject problems that violated the parity rule. For example,  $8 * 4 = 33$  violates parity because the product of operands that are even must be even. The relative RT advantage of violations compared to incorrect nonviolations of the parity rule depended on the proportion of wrong answers that were consistent or inconsistent with the parity rule. When 80% of the incorrect answers violated the parity rule, the RT *advantage* for violations over nonviolations was five times larger than when

there were only 20% violations among the incorrect facts to reject. There were also more errors (false acceptance) for nonviolation inequalities when most of the problems did violate the parity rule. However, when asked to guess what percentage of the problems were parity consistent problems, the subjects' responses did not differ for the conditions in which the parity did differ. In other words, the subjects seemed not to have explicit access to the relevant proportions. Furthermore, when asked about the parity rule, many subjects reported being unaware of the rule!

The notion that strategy use should be optimal without conscious awareness seems controversial from the perspective of metacognition; however, there is considerable work in the field of implicit learning that has already established this point quite clearly. Research under the title of implicit learning performed by Berry and Broadbent (1984), Lewicki et al. (1988), and A. S. Reber (1989) have shown a dissociation between subjects' performance and subjects' ability to report the rules that seem to guide their performance. For example, on a given trial in the experiments of Berry and Broadbent, a subject was told the output of a hypothetical sugar factory in tons and asked to choose the number of workers for the next month so that the output of the factory would remain within a specified range. The relationship between the tonnage and the number of workers was not obvious, yet subjects became proficient at controlling the factory's output. When asked to state the rule, subjects could not do this and claimed to make their responses on the basis of some sort of intuition or because the response *seemed* correct.

### CONNECTIONS BETWEEN IMPLICIT LEARNING AND IMPLICIT MEMORY

Given the similarity of the two words—learning and memory—and both starting with the modifier “implicit,” one might think that the terms “implicit memory” and “implicit learning” are synonyms and therefore find the heading here silly or nonsensical. However, these two terms do in fact refer to two distinct subdisciplines where there exist few examples of papers where one literature cites the other literature. That is, it is almost as if the two approaches are being conducted without an awareness or interest in the other. Ignoring the obvious similarity in names, these two disciplines are still highly interrelated. Consider, for example, the work of P. J. Reber (1993; P. J. Reber & Kotovsky, 1992) on implicit learning: Similarity or even formal isomorphism between two different problems affected subjects' ability to solve the second problem, such that they were much faster to solve the second problem; however, subjects were unaware of this formal similarity and attributed the improvement on the second problem to it

having an easier solution. This result is reminiscent of the finding on implicit memory of Jacoby, Allan, Collins, and Larwill (1988), in which subjects rated the loudness of aurally presented sentences that were embedded in white noise. Sentences that had been previously presented were rated as significantly louder than those that had not. Similarly, Warton and Lange (1994) found that subjects' ratings of the comprehensibility of written passages increased when they had previously read a passage with similar underlying themes. Jacoby et al.'s and Warton and Lange's subjects failed to appreciate that the perception of loudness or comprehensibility was due to temporary effects on their memory trace, just as Reber's subjects failed to recognize that the ease of solution was actually due to learning from performing a similar task earlier.

### WHY SHOULD METACOGNITIVE PROCESSES BE IMPLICIT?

We began this chapter with the somewhat counterintuitive claim that many of the processes that are considered prototypically metacognitive are not in fact open to conscious awareness (i.e., are not reportable). This seemed counterintuitive because it is typically assumed that metacognition is the one aspect of cognition that is mostly likely to be conscious. Indeed, we have argued that even the monitoring processes are frequently occurring without conscious awareness. Given that metacognition has been long thought to be the essence of the conscious aspects of cognitive processing, why should it be that these processes too are largely occurring without conscious awareness?

Again, we should review what we mean by conscious awareness. We mean that the information is reportable (see footnote 1). So perhaps we should rephrase our question to be, *Why do the processes that are typically considered metacognitive largely occur without the ability to report or verbalize the nature of these processes?* Perhaps the answer is that strategy selection is implicit because the very task of attempting to verbalize or make conscious the processes is interfering with the task at hand. Berry and Broadbent (1984) reported that performance degraded when subjects attempted to verbalize what was going on in the implicit learning tasks, especially when the rule underlying the task was not transparent. Smith, Haviland, Reder, Brownell, and Adams (1976) found that when subjects were consciously aware of the choices involved in a perception task, performance was significantly worse. The explanation put forward there too was that the conscious process interfered with the normal, automatic process of perception. This notion that automatic processes are interfered with by verbal processing has been most clearly articulated by Schooler (e.g., Schooler & Engstler-Schooler, 1990; Schooler, Ohlsson, &

Brooks, 1993; Schooler, Ryan, & Reder, in press; Wilson & Schooler, 1991). In these papers, he has demonstrated that performance is qualitatively different, and, more importantly, quantitatively worse when subjects attempt to give verbal reports. One interpretation of these results is that the act of verbalizing or trying to come up with a verbal representation uses processing capacity in a limited capacity system. Because we do not normally attempt to verbalize these processes, the act of verbalization is especially taxing and disruptive to the relatively automatic processes.

Another possible explanation for why strategy selection occurs implicitly is that it may not be worth the extra computational effort to consider the strategy selections explicitly. In many of the experiments that we reviewed in this chapter, subjects demonstrated very impressive sensitivities to various factors, and were well calibrated in their strategy selections. Thus, if the human architecture is so adept at making strategy selections implicitly, the small room for improvement may render an explicit strategy selection process inefficient.

There is considerable evidence that some of the processes which underlie these metacognitive tasks occur automatically, and are basic cognitive processes. For example, Hasher and Zacks (1979), Hintzman (1988), and others have demonstrated that people automatically register frequency information. This ability does not degrade with age (Hertzog & Dixon, 1994), nor with multitasking; however, if subjects are asked to be conscious of how they are making their decisions, they may well not use this automatically registering frequency information and base their judgments on some other process. For example, the literature on base-rate neglect (e.g., Kahneman & Tversky, 1973; Tversky & Kahneman, 1974) has found that people often do not use base rates in their decision making when they are presented the base-rate information in a verbal summary format. By contrast, when people are allowed to experience the base-rate information in a nonverbal problem-solving context, they are quite sensitive to base-rate information (Koehler, 1993; Spellman, 1993).

### IMPLICATIONS FOR EDUCATION

In considering the implications of these findings for educational practice, one must be careful in distinguishing between strategy selection and learning of new strategies. Our main thesis is that the processes which select between strategies in the current repertoire are independent of explicit monitoring, and that attempts to make those processes explicit results in degraded performance. For this case, the educational implications are clear: The implicit strategy selection processes should be left implicit, and attempts to tune them through explicit instructions will fail.

By contrast, the case of learning new strategies may be entirely different—it may well be that the process by which new strategies are learned is an explicit one and amenable to explicit interventions. Demonstrations of improving task performance by explicitly teaching new, more effective strategies are fairly common. For example, much of primary and secondary education consists of learning new skills via explicit teaching. Of course, these demonstrations do not prove that explicit teaching methods are better than more implicit teaching methods (e.g., learning by examples or learning by doing). However, studies by Chi and colleagues (Chi, Bassok, Lewis, & Reimann, 1989; Chi & VanLehn, 1991) provide suggestive evidence on this issue. They found that the ability to provide good self-explanations was important for learning physics problem-solving skills. Similarly, Crowley (1994) found that children acquired and transferred new strategies much more effectively when forced to provide explanations.

Given these findings, one can speculate about the next task for educational researchers: to determine whether poor performance on a task is due to incorrect strategy selection among an existing repertoire or lack of knowledge of the appropriate strategy. If the appropriate strategy is missing, then explicit teaching methods are likely to be best. However, if the appropriate strategy is known but simply not selected, then more implicit methods of teaching are required. For example, one might structure the environment such that student receives much success feedback with the desired strategy or much failure with undesired strategies.

### CAVEATS

Of course, we do not claim that strategy selection cannot be done explicitly, nor that it is not occasionally influenced by explicit factors. For example, Reder (1987) has shown that subjects' strategy selections can be influenced by explicit instructions. However, it is also true, much to experimenters' chagrin, that subjects frequently ignore explicit instructions (e.g., Mynatt, Doherty, & Tweney, 1978; Wason & Johnson-Laird, 1972).

There are also other cases in which people have been shown to have explicit knowledge of strategy effectiveness. For example, Crowley and Siegler (1993) found that young children were able to judge the effectiveness of a strategy even before they started using it. However, that people occasionally have explicit knowledge of strategy effectiveness does not necessarily imply that they use this explicit knowledge in making strategy selections.

Our conclusions may seem incongruous with our readers' own introspections that they are frequently aware of the variables affecting strategy choice. We have two responses: (a) of course there are times when people are aware of all the variables that affect their choice behavior; (b) however,

sometimes people may *attribute* reasons for actions (strategy choices) that are plausible reasons that have no basis in fact. That is, people are frequently conscious of the strategy (behavior) selected, and from those observations inferences are made about why the strategy was selected. Nisbett and Wilson (1977) have made a strong case that our verbal reports for why we behave the way we do are not always accurate.

Earlier we described cases in which explicit awareness had negative impact on performance. However, awareness need not always influence performance negatively; it could also simply have no effect. For example, in Experiment 2 of Schunn et al. (in press), subjects were asked during the debriefing whether they were aware of the operator-swap manipulation. We hypothesized that if awareness had helped performance, then the aware subjects would have been less likely to be fooled by the operator-swap manipulation (like Jacoby and Whitehouse's subjects); if awareness had hurt, then the aware subject would have been more likely to be fooled by the operator-swap manipulation. However, awareness of this manipulation had no influence on the subjects' tendency to be fooled (i.e., select retrieve and then give the answer to the original problem),  $F(1, 21) < 1$ . In this case, awareness neither helped nor hurt the subjects.

### CONCLUSIONS

We conclude this chapter as we began, namely with the bold conjecture that much of the processing that is called metacognitive typically operates at an implicit level; that is, without conscious awareness. Many of the tasks that are called monitoring are also operating without conscious awareness—people cannot veridically report what they have perceived and acted on. Furthermore, when people make efforts to change the nature of the task so that they are conscious and can report what they are doing, they run the risk of fundamentally changing their performance in non-optimal ways.

### ACKNOWLEDGMENTS

Preparation of this chapter was supported by a grant from the Office of Naval Research N00014-95-1-0223 to the first author. The authors wish to thank P. Stroffolino for development of the original SAC simulations and D. Richards for conducting the original fits of the simulation to the empirical data.

### REFERENCES

- Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, 6, 451-474.

- Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 2, 396-408.
- Berry, D. C., & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. *Quarterly Journal of Experimental Psychology*, 36A, 209-231.
- Blessing, S. B., & Ross, B. H. (1996). Context effects in problem categorization and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 792-810.
- Bowers, J. S., & Schacter, D. L. (1990). Implicit memory and test awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 404-416.
- Chi, M. T., Bassok, M., Lewis, M. W., & Reimann, P. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13(2), 145-182.
- Chi, M. T., VanLehn, K. A. (1991). The content of physics self-explanations. *Journal of the Learning Sciences*, 1(1), 69-105.
- Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern analyzing skills in amnesia: Dissociation of knowing how and knowing that. *Science*, 210, 207-210.
- Crowley, K. (1994). *How children learn from examples: A developmental study of explanation-based learning*. Unpublished doctoral dissertation, Carnegie Mellon University, Pittsburgh, PA.
- Crowley, K., & Siegler, R. S. (1993). Flexible strategy use in young children's tic-tac-toe. *Cognitive Science*, 17(4), 531-651.
- Davidson, J. E., Deuser, R., & Sternberg, R. J. (1994). The role of metacognition in problem solving. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 207-226). Cambridge, MA: Bradford.
- Doshier, B. A., & Rosedale, G. (1989). Integrated retrieval cues as a mechanism for priming in retrieval from memory. *Journal of Experimental Psychology: General*, 118, 191-211.
- Doshier, B. A., & Rosedale, G. (1990, November). *Dual cues in item recognition: An ensemble model of priming*. Paper presented at the Thirty-First Annual Meeting of the Psychonomics Society, New Orleans, LA.
- Dunlosky, J., & Nelson, T. O. (1992). Importance of the cue for judgment of learning JOLs and the delayed-JOL effect. *Memory and Cognition*, 20(4), 374-380.
- Glucksberg, S., & McCloskey, M. (1981). Decisions about ignorance: Knowing that you don't know. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 311-325.
- Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108, 356-388.
- Hernstein, R. J. (1990). Rational choice theory. *American Psychologist*, 45, 356-367.
- Hertzog, C., & Dixon, R. A. (1994). Metacognitive development in adulthood and old age. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 227-251). Cambridge, MA: Bradford.
- Hinsley, D. A., Hayes, J. R., & Simon, H. A. (1977). From words to equations: Meaning and representation in algebra word problems. In P. A. Carpenter & M. A. Just (Eds.), *Cognitive processes in comprehension* (pp. 89-106). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, 95, 528-551.
- Jacoby, L. L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior*, 22, 485-508.
- Jacoby, L. L., Allan, L. G., Collins, J. C., & Larwill, L. K. (1988). Memory influences subjective confidence: Noise judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 240-247.

- Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology*, 56, 326-338.
- Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, 118(2), 126-135.
- Jacoby, L. L., Woloshyn, V., & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General*, 118, 115-125.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237-251.
- Kamas, E., & Reder, L. M. (1994). The role of familiarity in cognitive processing. In E. O'Brien & R. Lorch (Eds.), *Sources of coherence in text comprehension: A festschrift in honor of Jerome L. Myers*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Koehler, J. J. (1993). *The base rate fallacy myth* [On-line]. Available: [psyc.4.49.base-rate.1.koehler](http://psyc.4.49.base-rate.1.koehler).
- LeFevre, J., Greenham, S. L., & Waheed, N. (1993). The development of procedural and conceptual knowledge in computational estimation. *Cognition and Instruction*, 11(2), 95-132.
- Lemaire, P., & Reder, L. M. (1996). *Is the strategy selection process biased: A case study (or the example) of arithmetic product verification?* Manuscript submitted for publication.
- Lewicki, P., Hill, T., & Bizot, E. (1988). Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. *Cognitive Psychology*, 20, 24-37.
- Lovett, M. C., & Anderson, J. R. (in press). History of success and current context in problem solving: Combined influences on operator selection. *Cognitive Psychology*.
- Luchins, A. S. (1942). Mechanization in problem solving. *Psychological Monographs*, 54(Whole No. 248).
- Metcalfe, J. (1994). A computational modeling approach to novelty monitoring, metacognition, and frontal lobe dysfunction. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 137-156). Cambridge, MA: Bradford.
- Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 851-864.
- Miner, A., & Reder, L. M. (1994). A new look at feeling-of-knowing: Its metacognitive role in regulating question answering. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 47-70). Cambridge, MA: Bradford.
- Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1978). Consequences of confirmation and disconfirmation in a simulated research environment. *Quarterly Journal of Experimental Psychology*, 30, 395-406.
- Nelson, T. O. (1984). A comparison of current measure of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, 95, 109-133.
- Nelson, T. O. (1986). ROC curves and measures of discrimination accuracy: A reply to Swets. *Psychological Bulletin*, 100, 128-132.
- Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1-25). Cambridge, MA: Bradford.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, 95(3), 385-408.
- Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, 118, 210-235.



- Reber, A. S., & Millward, R. B. (1971). Event tracking in probability learning. *American Journal of Psychology*, 84, 85-99.
- Reber, P. J. (1993). *Working memory, learning and transfer of nonverbalizable knowledge: Solving the balls and boxes puzzle*. Unpublished doctoral dissertation, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA.
- Reber, P. J., & Kotovsky, K. (1992). Learning and problem solving under a memory load. In *Proceedings of the 14th Annual Conference of the Cognitive Science Society* (pp. 1068-1073). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Reder, L. M. (1979). The role of elaborations in memory for prose. *Cognitive Psychology*, 11, 221-234.
- Reder, L. M. (1982). Plausibility judgments versus fact retrieval: Alternative strategies for sentence verification. *Psychological Review*, 89, 250-280.
- Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology*, 19, 90-137.
- Reder, L. M. (1988). Strategic control of retrieval strategies. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 227-259). San Diego: Academic Press.
- Reder, L. M., & Anderson, J. R. (1980). A partial resolution of the paradox of interference: The role of integrating knowledge. *Cognitive Psychology*, 12, 447-472.
- Reder, L. M., & Gordon, J. S. (1996). Subliminal perception: Nothing special cognitively speaking. In J. Cohen & J. Schooler (Eds.), *Cognitive and neuropsychological approaches to the study of consciousness*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with questions terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 435-451.
- Reder, L. M., & Ross, B. H. (1983). Integrated knowledge in different tasks: The role of retrieval strategy on fan effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 55-72.
- Ross, B. H. (1984). Reminders and their effects in learning a cognitive skill. *Cognitive Psychology*, 16, 371-416.
- Ross, B. H. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 456-468.
- Rudnick, A. I. (1990, April). *System response delay and user strategy selection*. Invited poster given at the 1990 Conference on Human Factors in Computing System, Seattle, WA.
- Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, 17, 36-71.
- Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General*, 122, 166-183.
- Schooler, J. W., Ryan, R. S., & Reder, L. (in press). The costs and benefits of verbally rehearsing memory for faces. In D. Herrmann, M. Johnson, C. McEvoy, C. Hertzog, & P. Hertel (Eds.), *Basic and applied memory: New findings*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Schunn, C. D., Reder, L. M., Nhouyvanisvong, A., Richards, D. R., & Strohffolino, P. J. (in press). To calculate or not calculate: A source activation confusion (SAC) model of problem-familiarity's role in strategy selection. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Schwartz, B. L., & Metcalfe, J. (1992). Cue familiarity but not target retrievability enhances feeling-of-knowing judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 1074-1083.
- Siegler, R. S. (1989). How domain-general and domain-specific knowledge interact to produce strategy choices. *Merrill-Palmer Quarterly*, 35, 1-26.
- Siegler, R. S., & Jenkins, E. A. (1989). *How children discover new strategies*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229-293). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Smith, E. E., Haviland, S. E., Reder, L. M., Brownell, H., & Adams, N. (1976). When preparation fails: Disruptive effects of prior information on perceptual recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 151-161.
- Spellman, B. (1993). Implicit learning of base rates: Commentary on Koehler on base-rate [On-line]. Available: [psyc.4.61.base-rate.4.spellman](http://psyc.4.61.base-rate.4.spellman).
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.
- Warrington, E. K., & Weiskrantz, L. (1970). Amnesic Syndrome: Consolidation or retrieval? *Nature*, 228, 629-630.
- Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content*. Cambridge, MA: Harvard University Press.
- Watson, J. B. (1925). *Behaviorism*. New York: Norton.
- Wharton, C., & Lange, T. E. (1994). Analogical transfer through comprehension and priming. In *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society* (pp. 934-939). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology*, 60, 181-192.
- Zbrodoff, N. J., & Logan, G. D. (1986). On the autonomy of mental processes: A case study of arithmetic. *Journal of Experimental Psychology: General*, 115, 118-130.