STRATEGIC CONTROL OF RETRIEVAL STRATEGIES

Lynne M. Reder

I. Introduction

Virtually all complex cognitive tasks can be accomplished using one of several different strategies. Not only do different people use different strategies to accomplish a given task, but the same individual will often chose to switch strategies within a short period of time even though the task or goal of this individual has not changed. Why does a person do this? How does a person decide to do this, and is it in his or her best interest to switch strategies so often? Although I have investigated these issues (e.g., Reder, 1982, 1987) within the domain of question answering, many of the arguments that hold for question answering can likewise be applied to other higher-level cognitive tasks, e.g., arithmetic, divided attention tasks, problem solving, and mental rotation. For the purposes of this article I will focus on question answering. Clearly, we use multiple strategies for question answering. Consider how differently you would approach questions such as "What country recently had a nuclear power plant disaster?" vs. "Does Ronald Reagan take naps often?" vs "How many Lithuanians live in Saskatoon?" It is most likely that attempts to answer these questions would look qualitatively different in terms of the strategies employed. So how do we decide if we know anything that is relevant to answering a particular question? How do we decide if we know anything that is relevant to answering a particular question? How do we decide the most appropriate procedure for answering a given question?

This article will try to argue for a particular type of architectural framework which accounts for basic memory phenomena and then, based on these simple assumptions, argue for the importance of assuming a strategy-selection component prior to careful memory search. These arguments will be both theoretical and empirically based. I will describe the variables that affect strategy selection and the possible mechanisms involved in selection and speculate on the generality of these mechanisms for other cognitive tasks.

II. A Two-Factor Theory of Memory Retrieval

Before describing my work on strategy selection in question answering, it will be useful to describe the type of memory framework on which I build. This framework is attractive in that it uses only two constructs to account for much of the memory retrieval literature. These concepts are *level of activation* of a memory trace and *strength* of a memory trace. This framework also makes some architectural assumptions about memory. Specifically, it assumes that memory consists of interconnected ideas (concept nodes) and that these interconnections or associations between ideas vary in strength.

Strengths of associations are determined by how often the associations are encountered. Strength builds up slowly over multiple encounters and also dies away slowly from disuse (much like a muscle). What a person is currently attending to is in a state of mental *activation*, and this activation can spread along pathways or connections to associated concepts. Concepts can become active from stimulation, either external, i.e., from the environment, or internal, i.e., something one is thinking about activates that concept. Activation spreads very quickly to other concepts but dies off very quickly, with a half-life of less than 1 sec. How much activation spreads to another concept node is partly a function of the strength of the connection between the two ideas, how many other associations or connections also share the activation, and how much activation there is to share.

Activation determines the ease of access of information—the more active the information, the easier it is to access. However, there is only a finite amount of activation that can be shared among concepts, so only a small portion of memory can be active at any one time. The rate of processing of an idea depends on the level of activation of the relevant concepts, i.e., those concepts involved in the procedures to be executed by memory.

An easy way to explain these constructs is by using some examples that we have probably all used to explain terms such as "short-term memory" to a nonpsychologist. In this situation, we remind the layperson of a past experience where a telephone number was obtained from directory

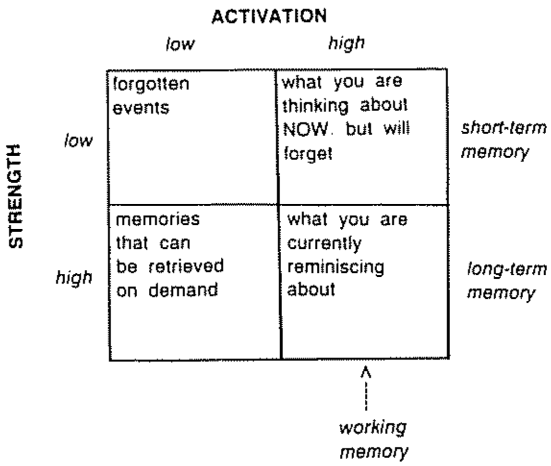

Fig. 1 Two-factor theory of memory retrieval

assistance and forgotten prior to its use because someone else asked him or her a question before the number was dialed. This situation and others are represented in Fig. 1. This figure shows the four possible states of a memory trace assuming binary values on these two dimensions.

The right-hand column describes information that is currently in a highly active state, namely, those thoughts of which we are currently aware. In other words, working memory consists of thoughts that are in a highly active state, and nothing can be in working memory that is not active. The top row refers to information that is low in strength, i.e., the connections between the elements (concepts) are weak because they have not been practiced together very much. So, for example, the phone number we get from directory assistance would be described by the upper right-hand box as long as we continued to rehearse it [and no one interrupted us to ask us a question, thereby blocking further rehearsals—See Baddeley (1986) for an elegant discussion]. We know that this phone number is low in strength because otherwise we would not have asked directory assistance for the number, but would instead have retrieved it from our long-term memory.

The bottom row refers to those items that are high in strength and, thus, have been retrieved from memory. The lower right-hand cell describes the information that we are currently *reminiscing* about. Once we stop thinking about that information, it falls from its active state, but it is not forgotten since the connections are still strong. Later they can be retrieved when wanted. The phone number, however, is long gone once we dial it and talk to the party on the other end. Of course, we all know that the phone number could become strengthened from repeated trials of using it (e.g., Hebb, 1961) or learned easily if the number were especially mnemonic (e.g., Prytulak 1971). In the latter case, associations are built to prior structures that are already especially strong, and one only needs to rehearse the pointer or connection to the prior strong set of connections.

A. Complicating the Picture: Competition for Activation

Unfortunately, we can not simply say that our ability to retrieve from memory depends on our ability to activate the relevant memory traces and that the strength of a trace or association affects how easily it can be activated. This picture needs to be complicated because the *amount of activation* is influenced by the *relative strength* of the associations, that is, the strength of an association in comparison to the other *competing* links also associated with the activated concept.[1] Indeed, many phenomena concerned with ease of learning, probability of forgetting, or time to respond can be explained in terms of competing associations to the same stimulus. The paired-associate learning paradigm (e.g., Postman 1971; Postman & Underwood, 1973) is an obvious example, as is the "fan" paradigm (e.g., Anderson, 1983). The latter paradigm shows that the more facts that are committed to memory about a particular concept, the slower a person is to recognize or reject (as not studied) any statement sharing that concept. Recently, there has been much research associated with "overwriting memory" (Loftus, 1979; Miller, & Burns, 1978; Loftus, Schooler, & Wagenaar, 1985; McCloskey & Zaragoza, 1985; Tversky & Tuchin 1987). This research is probably best thought of as instances of interference. This is shown both in the reaction time results of Donders, Schooler, and Loftus (1987) and in the research on probability of recognition by Tversky and Tuchin (1987).

These two variables, strength and interference, have opposite effects on response time: practice makes retrieval faster and competition makes it slower; however, the effects can be combined in a regular fashion. For example, Anderson (1983) showed that the time to recognize a sentence could be fit by a power function but that sentences with greater fan were fit to a different power function with longer retrieval times (RTs) for the same amount of practice. Similarly, Peterson and Potts (1982) and Lewis and Anderson (1976) showed that real-world facts are verified more quickly

[1]One metaphor that might make this easier to understand is to equate activation with the amount of water that flows through an irrigation channel. The strength of a connection maps onto the depth of the channel: the deeper the channel, the more water (activation) can pass through. The more competing irrigation channels that share one water source (more links emanating from one activated node), however, the slower the water will go down a given channel (the less activation for any particular path).

than new facts (due to their greater strength) but even real-world facts are verified more slowly when more new facts are learned that are related to them.

B. The Paradox of the Expert

It seems quite sensible, a priori, that strength should be associated with greater efficiency in retrieval; however, it seems counterintuitive to believe that knowing more about a concept should produce interference and weaker performance (Smith, Adams, & Schorr, 1978.) Nonetheless, subjects are slower to verify a fact when additional facts are known about that person or topic. When thought of as "interference," the fan effect makes sense, but when thought of in terms of expertise of knowledge, it seems paradoxical.

Perhaps the best way to understand this phenomenon is to assume that experts do not typically try to retrieve any one specific fact even to verify an assertion; rather, they frequently draw on their general knowledge and *compute* whether or not an answer seems plausible. There are several lines of research that are consistent with this point of view. For example, Reder (1976, 1979) found that verification times for statements about a story were strongly influenced by the statement's plausibility even when that statement had been explicitly mentioned in the story. This suggests that we do not always first search memory to see if a statement has been presented when we are asked to decide on a statement's plausibility.

Reder and Anderson (1982) found that people will often use a plausibility or consistency strategy even when asked to make recognition judgments. They had people study sentences such as the following: (1) *the teacher went to the train station;* (2) *the teacher bought a ticket for the 10:00 train;* and (3) *the teacher arrived on time at Grand Central Station*. The fan effect was not found when the foils (statements to be rejected as not studied) were not thematically related, e.g., *the teacher called to have a phone installed*. On the other hand, the traditional fan effect reappears when the foils preclude using an inference strategy to make recognition judgments, e.g., the test probe was *the teacher checked the Amtrak schedule*. In addition, the density of errors in the recognition task increased with the fan of the foils, suggesting that with more consistent information, the plausibility of the foil increased. This result supports the view that we often prefer to adopt a plausibility strategy even when asked to make recognition judgments and that sometimes we even adopt such a strategy when it will cause us to make errors (Reder, 1982).

The paradox is not completely resolved by showing that the interference effects of additional knowledge are dissipated when "experts" can adopt a plausibility-like strategy to make recognition judgments: knowing more should actually facilitate judgments. Experiments by Reder and Ross (1983) and Reder and Wible (1984) actually did find situations where there is facilitation in verification due to increases in fan. In these studies, all subjects learned thematically related sets of information and were tested about this information in different conditions. In Reder and Ross, subjects were tested in two types of recognition blocks similar to Reder and Anderson; however, they were also tested in a consistency judgment task where they were to say "yes" to both studied facts and facts that were not studied but were thematically consistent. For that block of trials, subjects were only to say "no" to facts that were thematically inconsistent with the material studied. Subjects were much faster to verify statements with a greater fan if they could verify them using a plausibility (or consistency) strategy. In the Reder and Wible study, subjects were also tested at a 2-day delay.[2] They replicated the findings of Reder and Ross (1983) and Reder and Anderson (1982) in the comparable conditions; however, they found greater facilitation with fan for the stated probes at longer delays, suggesting a greater reliance on the plausibility or consistency strategy at longer delays.

These facilitation effects can be explained by assuming that the semantically related facts are organized into a subnode attached to the main concept (or person) node. The more facts emanating from the subnode, the less activation that will go to any one fact; however, if the subnode itself is sufficient to determine that a fact is plausible, then the link from the main node to the subnode will be stronger when more facts fan out of it.[3]

C. The Role of Elaborations in Memory: When Do They Help and When Do They Hurt?

Thus far, we have discussed the variables that affect retrieval, e.g., the amount of activation that a memory trace receives. Related facts can compete for activation, thereby slowing response times (causing the fan effect) or hurting recall probabilities; however, we saw that, in certain situations, related facts can actually cause subjects to respond faster. The phenomenon that subjects verify a statements's plausibility more readily when there are more consistent facts would be of limited interest if it could only be demonstrated in fan-effect paradigms where the experimenter systematically varies the number of relevant facts. However, the idea that

[2]Actually, in the Reder and Wible study, different subjects made recognition vs. consistency judgments; making the judgments at two different delays was, however, a within-subject variable.

[3]A comparable model would be that facts are retrieved from the person node at random and the greater the number of consistent facts, the sooner will any one relevant fact be found.

[redacted]

people "recognize" statements by inferring semantically consistent facts can be shown in domains other than memorization of sets of facts.

The use of plausible reasoning to make recognition judgments has been shown for episodically learned materials, such as stories (e.g., Owens, Bower, & Black, 1979; Reder, 1976, 1982, 1987). It has also been shown for passages that refer to famous figures (e.g., Sulin & Dooling, 1974; Dooling & Cristiaansen, 1977). Owens *et al.* showed that people remembered more of a passage when there was *additional* information making the passage more interesting, viz., that the protagonist was concerned that she had become pregnant by her professor. In that situation, more veridical facts were recalled, but also more intrusions were made. Similarly, Sulin and colleagues found that passages about Helen Keller or Adolf Hitler were better remembered than identical passages about nonentities; however, there were also significantly more false alarms to statements that were true of the famous characters but had not been stated. Arkes and Freedman (1984) replicated the well-known finding (e.g., Chiesi, Spilich, & Voss, 1979) that high knowledge subjects have better memory for the information given; however, they showed that when the task forces the high-knowledge subjects to discriminate facts actually presented in the passage from plausible inferences, their performance was significantly *worse* than novices. This, too, suggests that we normally elaborate and that in most situations these elaborations facilitate rather than interfere.

We have thus seen that whether elaborations facilitate or interfere with memory retrieval depends on the situation. When we are able to answer a question by using a reconstructive or plausible reasoning strategy, more elaborations help our ability to answer questions; however, when we are required to use "direct retrieval," elaborations are interfering. In my own studies (e.g., Reder, 1979, 1982, 1987), I have found that people are faster to verify highly plausible statements than moderately plausible statements when they had been presented as part of the story, but also more prone to falsely recognize them when they had not been presented.

Given that there are situations where using reconstructive memory (or plausible reasoning) is the preferred way to answer a question and that there are other situations where direct retrieval is preferred, the question then becomes this: How does a person determine which procedure is preferable in a given situation? Some of my past research (e.g., Reder, 1982, 1987; Reder & Wible, 1984) suggests that the tendency to adopt one strategy or another varies with the situation and that different question-answering procedures are not simply executed in parallel with the faster procedure winning. Instead, these experiments support the view that people make a *decision* as to which procedure to prefer, i.e., that it can be under a person's conscious control.

Before reviewing the specifics of some of these experiments, let me first describe the general paradigm that characterizes them. Subjects read a series of simple stories and are then asked to make judgments about statements based on these stories. The stories are presented sentence by sentence on a computer screen, and the subject presses a space bar to see the next sentence of the story. The questions are asked after each story, after every 2 stories, after 5 stories, after all 10 stories, or a few days later. Subjects are typically asked either to make recognition judgments (i.e., "Did you see this sentence when you read the story?") or to make plausibility judgments (i.e., "Is this sentence plausible given the story you read?"). Statements that should be judged as plausible fall into two plausibility categories: they were rated by other subjects either as highly plausible or as moderately plausible. A subset of these plausible statements are randomly selected to be presented in the story as part of the story even for subjects asked to make plausibility judgments.[4] For subjects asked to make plausibility judgments, there were an equal number of implausible statements also asked at the time of the test so that subjects would respond positively to 50% of the items. An example of the type of story used and its questions are listed in Section VII. Reaction times for correct responses are used to infer which strategy subjects use to answer the questions. The difference in response times between previously stated and not-stated test probes measures the degree of use of the direct retrieval strategy, while the tendency to employ the plausibility strategy is operationalized as the difference in RT between the moderately plausible and highly plausible statements.

In earlier studies (Reder, 1982) it was found that subjects became faster in making plausibility judgments as the delay increased between reading the stories and the test of the information, while they became slower to make recognition judgments at comparable delays. The speedup for plausible reasoning judgments was due almost exclusively to the improvement in verifying statements that had not been presented as part of the story; those statements that had been asserted in the story were verified as quickly when tested immediately as when they were tested at a delay. One explanation of this result is that, at short delays, subjects prefer to search memory for the specific fact, while at longer delays they do not bother to first search memory before trying to infer the answer. In fact, the error rates for the recognition judgments went up dramatically from the short to the long delay, but only for those plausible statements that had not been presented. This also suggests a shift away from the use of a direct retrieval strategy toward a plausible reasoning strategy.

[4]This random selection was done separately for each subject so that effects due to materials were part of the subject error term.

III. Influencing Strategy Selection: Extrinsic Variables and Intrinsic Variables

A. STRATEGIC VARIABLES

In this section, I describe some recent work (Reder, 1987) that further supports the existence of an independent strategy-selection stage. These experiments use the paradigm described previously, namely, subjects read a series of stories and are asked to make judgments about them.

1. *Are Subjects Sensitive to the Probability of Finding a Statement Stored in Memory?*

In one of these experiments, subjects were asked to make plausibility judgments about statements immediately after reading a story. We varied the probability that the statement to be judged had been asserted in the story. For the first six stories that subjects read, the ratio of presented to not-presented statements was either 20:80 or 80:20, with different subjects randomly assigned to one of these two ratios. The "stated" implausibles were *contradictions* of statements from the story: a statement from the story had one word replaced with its opposite to create the contradiction. After subjects judged statements for six stories with this uneven ratio of presented to not-presented statements, they read four more stories where the ratio was an even 50:50 of presented to not presented (and an equal number of implausibles and exact contradictions).

Subjects were never explicitly told of the ratio, nor were they told that the answers to some of the questions had been stated in the story. One question of interest is whether a person's probability of employing a particular strategy varies as a function of the prior history of success with that strategy. If so, then several conclusions can be drawn about the question-answering process. For example, if subjects prefer a strategy that is sensitive to the ratio of presented to not-presented statements, then they are not always searching first for an exact match in memory before trying the plausible reasoning strategy. Second, they would not simply be letting both the search process and the inference process run in parallel, with the faster process providing the answer.

The results of this experiment are displayed in Fig. 2, which plots a difference measure of reaction times for not-stated plausible probes minus stated plausible probes as a function of the ratio of presented to not-presented probes. A large difference in RT suggests that subjects were using the direct retrieval strategy. A small difference suggests that subjects were employing the plausibility strategy. Note that for subjects who received a disproportionately large number of probes that had been stated in the

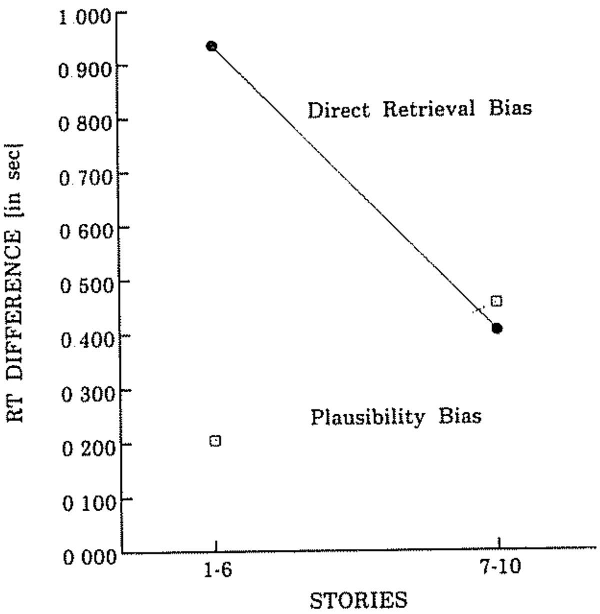

Fig. 2. Difference in reaction time between stated and not-stated statements (collapsed over plausibility).

story, there was a big difference in response time between the stated and not-stated probes. This suggests that these subjects became accustomed to searching memory first for the probe and only resorted to plausible reasoning when that strategy failed. In contrast, for those subjects who rarely found the probe stated in memory, the difference in RT was much smaller, suggesting that they tended to adopt the plausible reasoning strategy.

A second type of analysis examines the difference in response time between highly plausible statements and moderately plausible statements, collapsing over whether the statement had actually been presented as part of the story. In that case, the pattern is exactly reversed: there is a large difference in RT for subjects who rarely found a probe stated in the story, which is consistent with the view that they were using the plausible reasoning strategy as a first choice. Similarly, subjects who were biased to use the direct retrieval strategy because of the large proportion of stated probes showed a very small difference in RT between highly and moderately plausible statements, suggesting that they tended to prefer the direct retrieval strategy. The size of the differences in RT (between the stated vs. not-stated probes and the moderately vs. highly plausible probes) grew with exposure to the ratio. Analyses that compare differences after three stories and after six stories showed the increasing trend, and the differences between the two ratio groups declined during the last four stories (comparing stories 7 and 8 with stories 9 and 10).

2. *Are Subjects Sensitive to Official Task Requirements When These Do Not Matter Objectively?*

A slightly different type of experiment also provides support for the view that the type of question-answering procedure employed is under strategic control. In this experiment subjects also read stories and made judgments about statements concerning each story. Subjects were randomly assigned either to the recognition task or to the plausibility judgment task. In recognition, subjects were asked to decide whether or not a statement had in fact been seen as part of the story; the other subjects judged whether or not a statement was plausible given the story they had read. Although the official tasks were different, the experiment was constructed in such a way that subjects could follow either the plausibility strategy or the direct retrieval strategy and always get the correct answer. This effect was achieved by including all plausible statements in the story. So for the recognition task, all not-presented statements were implausible, thereby allowing subjects to use a plausibility strategy to make a recognition judgment. Likewise for the plausibility task, subjects could use a direct retrieval strategy because if the statement was plausible, it would have been previously presented. Some subjects were asked to answer these questions after each story while others were randomly assigned to the delay condition (approximately a 20-min delay) where questions were asked about each story after all ten stories had been read.

The results of this experiment indicate that subjects do not simply employ the official strategy requested of them; however, they are definitely influenced by the task demands. Figure 3 graphs the tendency to employ the plausibility strategy as a function of official task demands and delay condition. In the last experiment described, we had two converging measures of strategy use, one being the difference between the two levels of plausibility and the other being the difference between the stated and not-stated probes (of otherwise equal plausibility). In this experiment, since all plausible statements had been presented, we can only use the former measure. Note that the difference in RT is much greater for subjects who were asked to make plausibility judgments than for those subjects who were asked to make recognition judgments; however, at the longer delay interval, the tendency to use the plausibility strategy over the direct retrieval strategy has increased for both groups such that subjects in the delay recognition condition have as great a tendency to use the plausibility

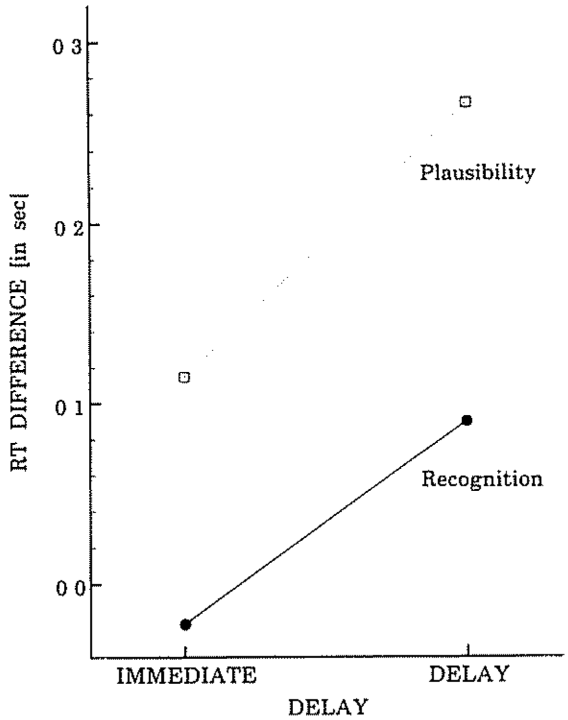

Fig. 3. Difference in reaction time between moderately and highly plausible statements.

strategy as do subjects assigned the the immediate plausibility condition. The increased use of the plausibility strategy with delay is not, by itself, an argument that subjects bias their preference for one strategy over another. However, it is the most parsimonious explanation: task instructions differentially affect the tendency to select one strategy over the other, even though subjects do not always use the strategy that corresponds to the official task. In sum, it seems clear that subjects do control which strategy they will use at any given time.

3. *Strategic Control from Trial to Trial*

The next experiment does more than support the view that subjects have strategic control of their use of question-answering processes. It also shows that subjects can adjust their strategy preference from trial to trial rather than only changing their long-term bias on the basis of prior history of success with a strategy or official task demands. In this study subjects were advised prior to each question which strategy was most likely to be effective. Again, subjects read a series of stories, each story followed by a set of statements to be rated on their plausibility with respect to the last story read. Half of the plausible statements had been included in the story and half of the implausible statements had their exact contradiction presented in the story. On half of the trials, subjects were advised that the next probe (or the exact contradiction of the probe) could probably be found in memory. On the other half of the trials, subjects were advised that they probably would not find the next test probe (or its exact contradiction) in memory. The advice was accurate 80% of the time. For the first two stories, the advice was always correct in order to motivate subjects to pay attention to the advice.

The remaining eight stories were analyzed as a function of advice appropriateness, whether or not the statement had been presented in the story, and as a function of probe plausibility. The correct RTs are displayed in Fig. 4 as a function of correctness of advice and whether the probe had been stated or not in the story. These two variables define the type of advice given to the subject. For instance if the statement had not been

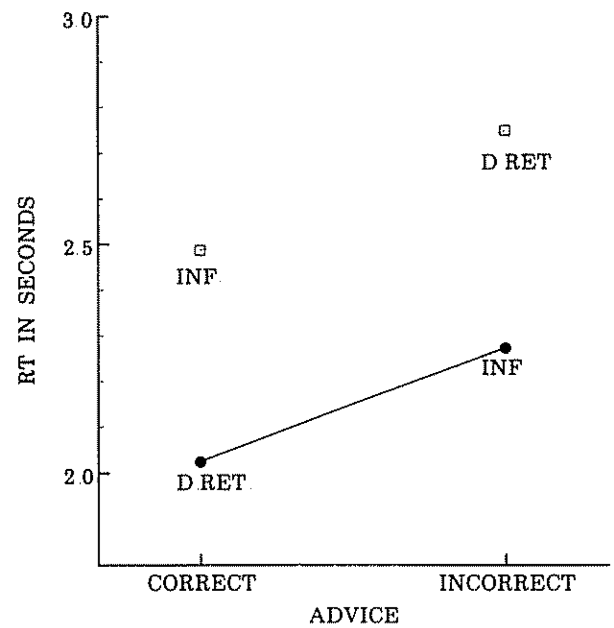

Fig. 4 Reaction times for stated vs not-stated probes (collapsed over plausibility) D. RET., Direct retrieval; INF., inference; □, not stated; ●, stated.

presented and the advice was inappropriate, then the advice must have been to try searching memory for the fact (the direct retrieval strategy). Subjects were significantly faster when they were given the correct advice as compared with the inappropriate advice, so, clearly, subjects were acting on the advice given to them.

The RTs for not-stated probes when direct retrieval was advised were much slower than all others because this was the only condition where the advised strategy would not work. The advice to infer when the probe was stated will work; however, it is nonoptimal since, at such a short delay, direct retrieval is a faster strategy.

IV. When Does the Strategy-Selection Stage Operate?

Taken together, these studies provide evidence for the existence of a strategy-selection stage that precedes strategy execution. If subjects simply let both the direct retrieval and plausibility strategies run in parallel, then giving the wrong advice prior to receiving the question should have no effect. Likewise, official task demands and prior history of success should have no effect on behavior. On the other hand, our results do not shed light on several other related issues. First of all, these studies do not discriminate between a system where the competing procedures run in parallel with differential allocation of resources and a system where only one of the two procedures executes at one time (i.e., a serial system). In either case, however, there must first be a mechanism that decides which process to favor.

Second, these studies do not answer the question of whether or not people can adjust which strategy they will prefer based on an initial examination of the question. That is, the evidence supports both the need for and the existence of a strategy-selection phase; however, none of these data is inconsistent with the view that this decision stage occurs prior to seeing a sentence. It could be the case that subjects make specific decisions about what strategy is the best to use *in the absence of parsing a question*, yet are unable to adjust strategy selection based on seeing the question itself. Prior history of success with a strategy (base rate of presented statements, etc.), advice prior to seeing a question, official demands of the experiment (task to judge plausibility vs. recognition), and prior knowledge of the delay between reading the story and test questions could all be explained within a model that posits subjects adjusting strategy preference prior to seeing a question. Several experiments will be described that lend support to the view that strategy selection can also take place after an initial parse of the test probe.

A. Can We Select a Strategy after We Have Seen the Question?

Strategy Selection in a Mixed-Delay Design

This study (Reder & Wells, 1986) was quite similar in design to many of the studies described previously. Subjects read ten stories; half were asked to make recognition judgments and half were asked to make plausibility judgments about statements concerning the story. The critical change in design was that before a question appeared on the computer screen the subject did not know whether it would be from a story that he had just read or from a story read 2 days earlier. Subjects read five stories 2 days prior to answering questions and read the other five stories just prior to answering questions. Subjects received a random mix of questions from a story they had just read and from one read 2 days earlier.[5]

Using this design, we could determine whether subjects could use a rapid inspection of a statement in order to bias question-answering strategy use. If it is possible to bias strategy selection after seeing the question, presumably stories read 2 days earlier would be answered using a plausibility strategy while questions from stories that had just been read would be answered (when possible) using the direct retrieval strategy.

The results of this study are presented in both tabular form (see Table I) and in two figures (see Figs. 5 and 6). The data are represented in the figures as a function of the delay between reading the story and seeing the test probe (immediate vs. 2 days) and as a function of task (plausibility vs. recognition). The ordinate in Fig. 5 plots a difference measure of correct response times for not-stated probes minus stated probes. Figure 6 plots the same contrast for accuracy. The variable of plausibility is ignored in these figures, although it is reported in the table. It should be pointed out that the function for the plausibility subjects represents only "yes" responses while the correct responses for not-presented statements in the recognition task are "no" responses.

Figure 5 shows that the effect of whether an item was stated or not is much greater for subjects who were asked to make recognition judgments than for those asked to make plausibility judgments; in the immediate condition, there is an effect of stated vs. not-stated for both groups. This

[5]After reading a story on the second day, a subject was shown a title from one of the five stories from the previous session. The questions from both the story referred to by the title and the story just read were then presented. In this way, subjects knew which two stories were to be queried in any set of questions.

TABLE I

MEAN RESPONSE TIMES (SEC) TO MAKE VERIFICATION JUDGMENTS, AND PROPORTION CORRECT, MIXED-DELAY EXPERIMENT[a]

	Recognition		Plausibility	
	Stated	Not stated	Stated	Not stated
Immediate				
High	2.003	2.518	1.888	2.195
	(.82)	(.74)	(.96)	(.88)
Moderate	2.060	2.571	2.159	3.024
	(.82)	(.90)	(.92)	(.73)
Implausible/contradictory	—	—	2.529	2.676
	—	—	(.89)	(.86)
Delayed				
High	2.167	2.840	2.238	2.261
	(.77)	(.53)	(.85)	(.82)
Moderate	2.320	2.932	2.520	2.615
	(.73)	(.68)	(.83)	(.69)
Implausible/contradictory	—	—	2.693	2.695
	—	—	(.71)	(.82)

[a]Proportion correct is shown in parentheses

triple interaction is significant [$F(1,39) = 10.2, p < .01$]. This means that subjects who were asked to make plausibility judgments were adopting the direct retrieval strategy in the immediate condition but not when the statement had been read 2 days earlier. Consistent with this view, the effect of statement plausibility (in the plausibility task) is also significantly larger for stories that were read 2 days earlier if the statements had not been presented, i.e., plausibility × stated × delay [$F(1,20) = 7.23, p < .01$]. (For statements that had been presented in the story, the effect of plausibility is almost always attenuated.)

Now consider the performamce of subjects who were asked to make recognition judgments. Although the response times are considerably longer for statements that were read 2 days earlier (see Table I), the difference function (betweeen not stated and stated) does not change much with delay. This is because if subjects adopt the plausibility strategy, they are very likely to make an error on the not-stated probes, and those RTs are not reflected in these means. On the other hand, the error rates do show an increased tendency to use the plausibility strategy for the statements that were read earlier (see Fig. 6). Subjects make significantly more errors to not-stated probes that were read 2 days earlier [$F(1,19) = 8.53$,

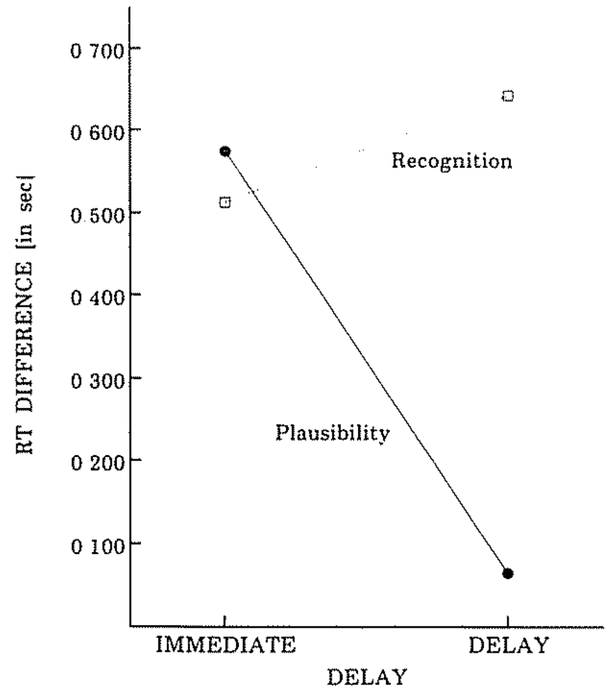

Fig. 5. Difference in reaction time between not-stated and stated probes (collapsed over plausibility).

$p < .01$], suggesting that they used the plausibility strategy for "older" statements instead of using direct retrieval and, thus, erroneously said "yes" to not-stated probes.

The reliance on a strategy that produces errors for subjects who were asked to make recognition judgments produced a main effect of task on errors ($p < .01$.) Recognition subjects were much more likely to make errors on highly plausible statements (plausibility × task) ($p < .01$). They were also much more error prone for not-stated items at a delay than were the plausibility subjects (stated × delay × task), $p < .01$).

All subjects were more likely to make errors to not-stated probes ($p < .01$), but the effect was most marked for subjects who were asked to make recognition judgments on highly plausible, not-stated items (plausibility × stated × task, $p < .01$). Here, subjects are erroneously accepting the highly plausible statements which they are supposed to reject in that condition. As one would expect, performance was also more error prone for older memory traces for all subjects, i.e., there was a main effect of delay on errors.

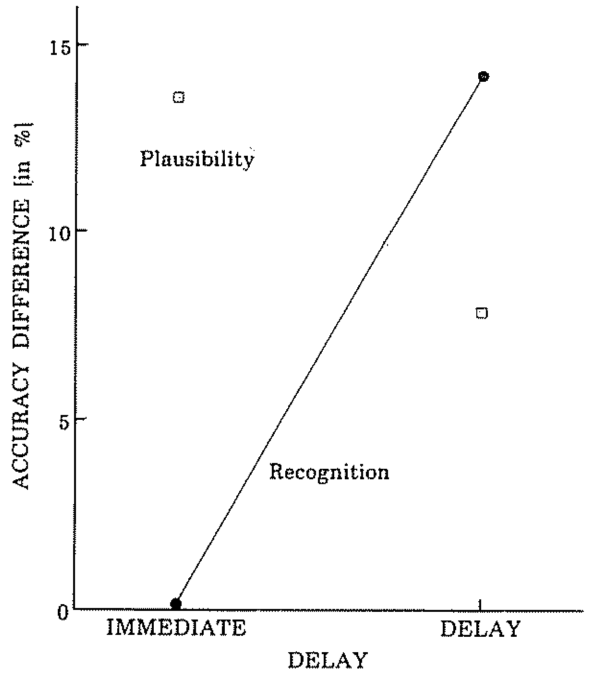

Fig 6 Percentage difference in accuracy between stated and not-stated probes (collapsed over plausibility).

TABLE II

Examples of Types of Questions Used in Game Show Experiments

Easy	
	What month follows September? (October)
	How many letters are there in the alphabet? (26)
Moderate	
	Who wrote "Romeo and Juliet"? (William Shakespeare)
	Who invented the phonograph? (Thomas Edison)
Hard	
	What was Mark Twain's real name? (Samuel Langhorne Clemens)
	Which well-known artist painted "Guernica"? (Picasso)
Impossible	
	What size collar does Lassie wear?
	What is George Bush's telephone number?

In sum, there was evidence of use of strategies other than the prescribed strategy for both tasks, viz., use of the direct retrieval strategy by subjects who were asked to make plausibility judgments at short delays and use of the plausibility strategy by subjects who were asked to make recognition judgments at long delays. This replicates other findings by Reder (1987), but in a paradigm where subjects cannot know prior to viewing the question how old the memory trace is. One can, therefore, conclude that subjects are able to adjust their strategy choice after seeing a question. It is not something than can only be adjusted in advance of processing the question.

V. Influencing Strategy Selection: Intrinsic Variables

Now that we have established that there must be a strategy selection mechanism that biases strategy use, there remains the issue of how this strategy selection is done. Given that factors such as prior history of success with a strategy, official task demands, and advice prior to seeing a question can all influence strategy selection, we know that there is a mechanism that is independent of question parsing that can influence this selection. Earlier work (Reder, 1982) had shown that delay between reading a story and test of the story's content influenced strategy use. It was unclear if such strategy preference was a conscious decision based on the subject's knowledge of when the story had been read or was influenced by the parsing of the question itself. The previous experiment described showed that people can ascertain the age of the information by inspecting the question itself and adjust their strategy use accordingly. Therefore, we know that not all mechanisms associated with strategy selection are concerned with cues that are external to the probe such as prior history of success. Thus, we need to specify a mechanism that can help people to determine, on the fly, what is the best strategy to use in a given situation.

A. Feeling of Knowing as an Intrinsic Strategy Selection Mechanism

One possible mechanism that can help people determine the best strategy involves "feeling of knowing" judgments. Previous work on feeling of knowing has demonstrated that when people are unable to answer a question, they still can assess their ability to recognize the answer (e.g., Nelson, Gerler, & Narens, 1984). Conceivably, the same mechanism that is involved in assessing our feeling of knowing when we cannot answer a question is also involved in assessing our ability to answer those questions that we can answer.

To address this hypothesis, we developed the "game show" paradigm (Reder, 1987) where subjects first estimate whether they think they will be able to answer a question before they actually attempt to answer it. We call it this because in some television game show formats, contestants are motivated to press a buzzer indicating their willingness to attempt an answer to a question even though they have not yet heard the full question. We asked half of our subjects to respond in this game show fashion, giving a rapid first impression or best guess as to their ability to subsequently answer a given question. The other half of the subjects were asked to immediately answer the questions. The experimental question is: can people estimate their ability appreciably faster than they can generate the answer, without suffering a loss in accuracy? Table II gives examples of the types of questions subjects might be asked to answer.

The results were very clear: subjects were able to *estimate* that they could answer a question significantly faster than they could answer it. The data are presented in Table III for time to estimate an answer or give an answer, proportion attempted and proportion correct. There is clearly a sizeable RT advantage for subjects in the estimate group. This effect is even larger if the first 25% of the trials are excluded. Presumably, this practice effect reflects the fact that subjects are not accustomed to overtly estimating their ability to answer a question.

One potential confounding factor in this experiment is that subjects in the answer condition had to articulate a response that they had not said before in the course of the experiment while subjects in the estimate condition were either saying "yes" or "no"—short, one-syllable words they had practiced a lot during the experiment. So we conducted a control experiment where the answer group pressed a button when ready to give the answer or when they had decided that they could not answer the question. The estimate group pressed a button if they thought they would be able to answer a question or the other button if they thought they could not. The difference in the two conditions, apart from the instructions, was that the question disappeared from the screen in the answer condition after the subject pressed a key (indicating that the answer was in mind) while the question remained on the screen after the subject in the estimate condition indicated that he probably could answer it.

The data for this control study are displayed in Fig. 7. The time to push the button is displayed in the bottom curve. The upper curve is the total time required to generate an answer. Therefore, the distance between the bottom curve and the top curve reflects the amount of time required for the subject to generate the answer after pushing the button. Note that subjects who were asked to estimate whether they could answer a question took (significantly) less time to push the button than did subjects who were asked to push a button when the answer had been retrieved from

TABLE III

TIME TO ESTIMATE OR ATTEMPT ANSWERS, PROPORTION OF QUESTIONS ATTEMPTED, PROPORTION ANSWERED CORRECTLY, AND ACCURACY OF ATTEMPTS AS A FUNCTION OF TASK AND QUESTION DIFFICULTY; GAME SHOW CONTROL EXPERIMENT[a]

Question difficulty	Time to attempt (sec)		Proportion attempted (%)		Total correct (%)		Accuracy of attempt	
	Est	Ans	Est	Ans	Est	Ans	Est	Ans
Easy	1.650	2.512	86.02	91.62	80.62	80.35	93.60	87.54
Moderate	1.828	2.654	67.46	72.29	59.48	54.97	87.98	76.08
Hard	1.728	2.388	42.38	47.09	34.13	28.42	81.44	59.07
Impossible	1.524	2.925	7.56	25.13	—	—	—	—
All attempted[b]	1.735	2.518	65.29	70.34	58.08	54.58	87.67	74.23

	Est (sec)	Ans (sec)
RT for answered correctly[b]	1.722	2.396
RT for answered incorrectly[b]	2.069	3.009
RT for not attempted[b]	1.851	3.197

[a]Est, Estimate; Ans, answer.

[b]These means do not include the impossible items. The pattern is quite similar with their inclusion.

memory. On the other hand, they took significantly more time to generate the answer once they decided that they probably could find the answer than did answer subjects, who claimed that they already had the answer in mind. Indeed, the total time required to answer a question was effectively identical for the two groups, which suggested that the processes used in both tasks are the same, but partitioned differently.

It is important to note that the significant speed advantage for the estimate group is not associated with poorer accuracy. On the contrary, the estimate group was appreciably more accurate or calibrated (88%) than the answer group (74%). Accuracy or calibration is defined as the percentage correct divided by the percentage attempted. Although the estimate group attempted considerably fewer questions, they seemed to have an excellent sense of which questions they will not be able to answer, while people in the answer group seem to try to answer too many. The point that subjects are more accurate in the estimate condition allows us to dismiss the issue of a possible speed/accuracy trade-off. Clearly, the estimate condition is a useful and quick process that might well play a role in an initial evaluation phase used to select an appropriate strategy.

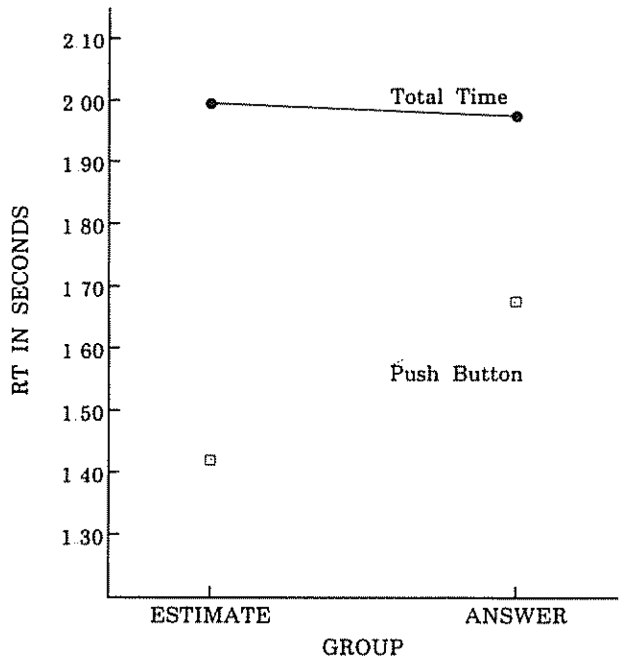

Fig. 7. Time to push button and total time to initiate response

B. Variables That Affect Feeling of Knowing and Ability to Answer Questions

The previous section provided support for the idea that people's feeling of knowing operates fast enough to be a viable mechanism involved in strategy selection. Now we can ask what variables *intrinsic* to the question affect our feeling-of-knowing judgments. One possible variable would be familiarity or recency of exposure to the terms in the question. Presumably, a feeling of knowing occurs for a question when there are recent memory traces of words used in the question and there is no concomitant association of *not knowing* the answer. That is, the question topic might well seem familiar to a person if most of the content words in the question had recently been talked about.

1. *Spurious Feelings of Knowing*

In another experiment, we attempted to give subjects a spurious feeling of knowing by giving them prior recent exposure to some of the terms of some of the questions. We did this by having subjects make frequency-of-occurence estimations on words which would later be used in some of the questions that they were asked to judge and answer. We did not want subjects to be aware of our attempt to give them a spurious feeling of knowing, so we only primed one-third of the questions. We were not certain whether subjects could be given a feeling of knowing about the questions without their also being aware that terms used in the question had also been seen a few minutes earlier in the same experimental context.

We selected questions that varied in the probablity that a subject could answer them correctly. The pool of questions that we used had been normed by Nelson *et al.* (1984). For each question, we selected two words or expressions from the question that seemed *central* to it. For example, for the question "What is the term in golf for scoring one under par?" we selected the words *golf* and *par*. For the question "What was the name of the clown on the 'Howdy Doody' television show?" the selected terms were *clown* and *Howdy Doody*. Although a pair of words was designated as central for every question, only one-third of the questions were randomly selected for priming for any given subject. All subjects were tested on all questions, but each subject received a different random selection of priming terms.

After the word-frequency estimation task, subjects were randomly assigned to either the estimate condition or the answer condition of the game show experiment. We expected different effects of priming for the two tasks even though priming was hypothesized to affect feeling of knowing in all cases. In the estimate task, we expected subjects to be more inclined to think that they could answer a question and, hence, have an elevated propensity to *attempt* to answer questions. On the other hand, we did not expect this effect to show up in the answer condition since answer subjects should not be more likely to actually retrieve the answer. The spurious feeling of knowing should manifest itself in longer search times before saying "can't say" for that group.

The results essentially supported our hypothesis. Figure 8 plots the effect of priming on percentage of questions attempted for the two groups as a function of question difficulty. The effect of priming is the difference in proportion attempted between the primed and unprimed questions. In the estimate condition, subjects are more inclined to attempt a difficult question if it was primed. Presumably, the reason priming did not matter for easier questions is that subjects already had a feeling of knowing for these questions.[6] For the answer condition, there is essentially no effect of

[6]In fact, the situation is slightly more complex than this. With very easy questions, subjects attempted even fewer if they were primed. We believe this is because subjects are sensitive to the priming manipulation and try to adjust their threshold for feeling of knowing if they recognize that the question has been primed. The adjustment tends to overcorrect for easy questions and undercorrect for very difficult ones. See Reder (1987) for additional discussion.

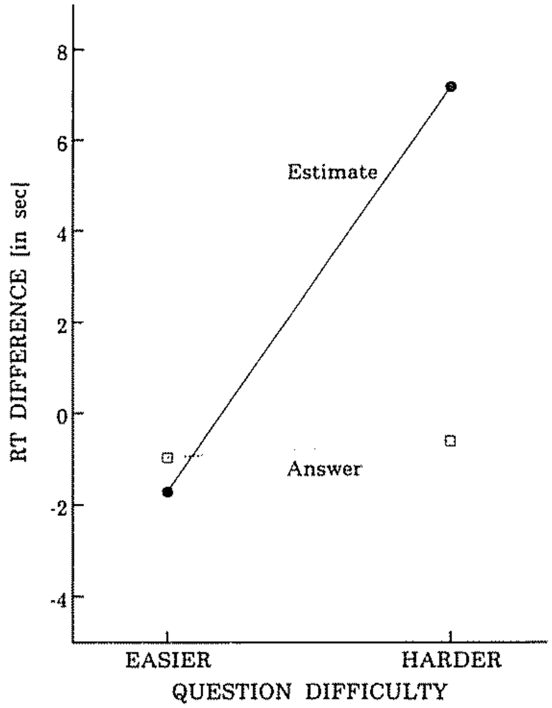

Fig 8 Difference in percentage attempted between primed and unprimed statements

priming. The effect of priming is manifest for the answer group in the time to say "don't know" shown in Fig. 9. These subjects also have a spurious feeling of knowing, but in this case it causes them to search longer before giving up.

2. *Are Feelings of Knowing Influenced by Self-Assessments of Topic Knowledge?*

In the following study (Reder & Fabri, 1984), we were interested in seeing how one's self-assessment of expertise in an area affected one's willingness to attempt to answer a question. We developed trivia questions in four domains: movies, sports, geography, and U.S. history. We varied whether the questions contained a lot of topic-relevant terms or were fairly short in length and contained few topic-relevant terms. Our interests here were in determining whether extra concepts helped to "prime" an answer, in discovering if this priming manipulation behaved differently for feeling-of-knowing judgments as compared with just answering questions, and, finally, in determining whether extra priming or topic terms interacted with one's own assessment of expertise in an area.

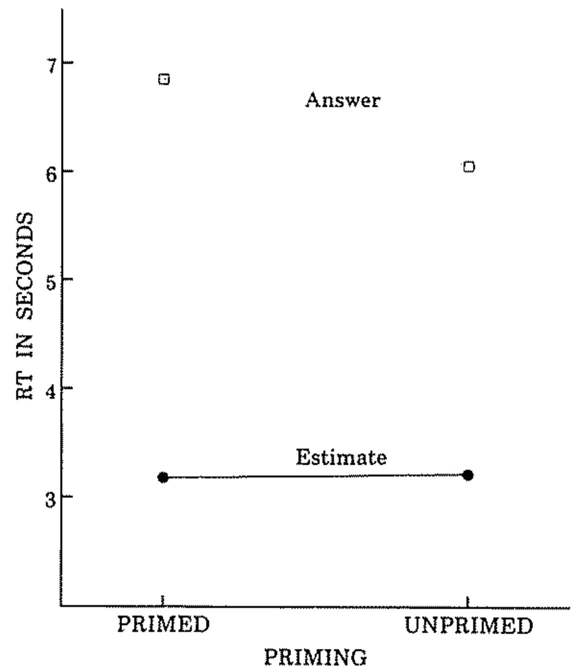

Fig 9 Time to say "don't know" instead of attempting to answer

Our initial intent was to control length of question while varying the number of terms in the question that pertained to one of the four categories mentioned previously; however, it proved virtually impossible to control length while varying number of concepts without significantly changing the style and structure of the questions. So, instead, we made sure that all additional priming terms occurred at the end of each question.[7] Examples of the short and long versions of a question are "What is the name of the longest river in the United States?" and "What is the name of the longest river in the United States whose mouth is at New Orleans?"

After subjects rank ordered their knowledge of the four topics, they were randomly assigned to either the answer condition or the estimate condition. In the latter condition, they were asked to estimate their ability to answer the presented question prior to actually answering it. These subjects were told to say "yes" or "no" as quickly as they could because we were interested in their first impression of their ability to retrieve the answer to the question. Only after "yes" responses were subjects asked to give the answer.

[7]If questions with more primes are faster, then it could not be due to getting a good clue early.

Subjects in the answer condition spoke their answer to a question directly into the microphone. They were instructed to respond only when they were sure they had an answer in mind. If subjects did not know the answer to a question, they were told to respond "don't know." In both conditions, RTs were measured from the time a question appeared on the screen until the subject began to vocalize the answer.

There were a number of interesting results associated with this experiment. Figure 10 displays the percent of questions attempted as a function of task and topic familiarity. The percentage of questions that subjects attempted to answer differed as a function of the question topic, such that questions from the topics that subjects rated as their best topic were attempted more often than questions from the topics that subjects rated as their worst topic. This result in itself is not surprising; what is more interesting is that subjects' self-classifications of relative knowledge of these topics seems to have a greater impact in the estimate condition than in the answer condition. That is, subjects seem to use their own impressions of what they know fast enough to *estimate* whether or not they could answer a question. Unfortunately, this effect did not reach significance.

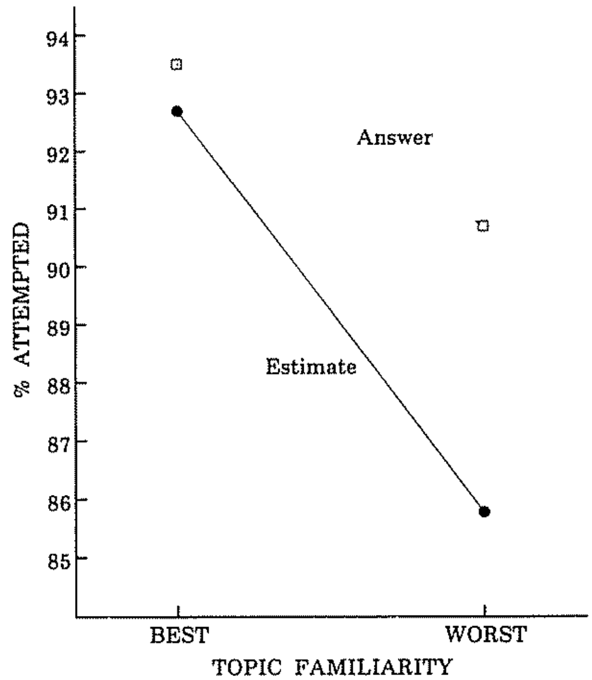

Fig. 10. Percentage attempted as a function of self-rated topic knowledge and task

A related finding was that there was more of an effect of rated topic knowledge in the answer condition upon the time to say "don't know" (see Fig. 11). This is quite similar to the effects in previously described experiments which concerned spurious feelings of knowing where primed questions took longer to reject. Here, subjects in the answer condition took longer to say "don't know" to a fact if it concerned an area they knew more about.

C. An Unexpected Result from Manipulating Feeling of Knowing

A side result from the work designed to investigate which variables affect feeling of knowing was the phenomenon that priming the terms of a question seemed to actually increase the probability that the answerer could answer the question. That is, in one of the game show studies described previously, there seemed to be a tendency for subjects to respond more accurately in the answer condition when the question had been primed than when it had not been primed. Conceivably, priming the terms

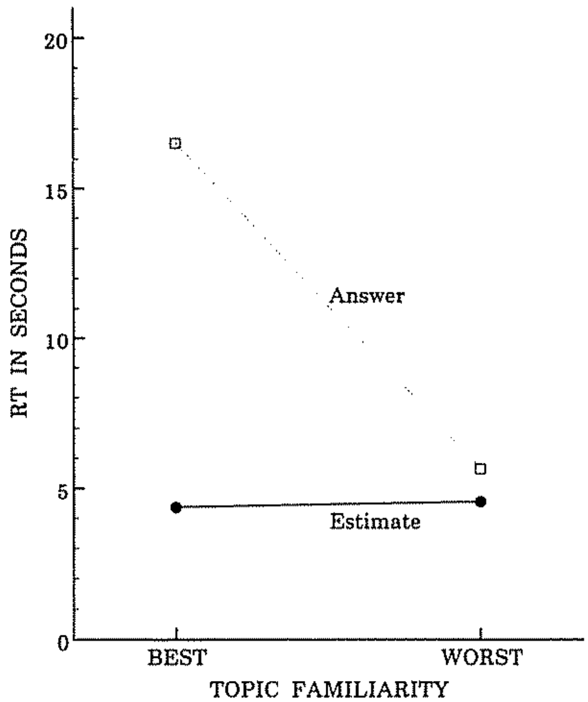

Fig. 11 Time to say "don't know."

of a question not only gives one a feeling of knowing but actually raises the level of activation for relevant information such that the answer is more likely to pass over some kind of threshold necessary to elicit an answer.

This result led us to do further tests to confirm the result. We conducted another study (Reder, Dennler, & Wells, 1985) where subjects were asked to generate sentences using terms that would later appear in questions they had to answer. In other words, just as in the "spurious priming" study described previously, questions were randomly selected to be primed for each subject. In the earlier study, subjects rated the terms for familiarity. Here they were required to generate original sentences using two terms. Again, subjects were not told that these terms would later appear in questions. Of those questions that were selected for priming, some were assigned to the conjoint and some to the disjoint condition. In the conjoint condition, subjects composed a sentence using two terms that came from the same eventual question; in the disjoint condition the two terms came from two different questions. For example, if the test questions were "What is the name of the clown on the Howdy Doody show?" and "In what town did Lady Godiva make her famous ride?" the primed terms might be recombined into the pairs, *clown–ride* and *Howdy Doody–Lady Godiva*.[8]

The results are displayed in Table IV. Subjects were significantly faster to answer a question in the conjoint condition, as compared with the disjoint or the unprimed condition, ($p < .05$). Subjects were also more accurate for primed than for unprimed questions and slower to say "don't know" to primed than to unprimed questions, but these latter effects did not reach significance. It appears that the same manipulation that affects feeling of knowing may also influence a person's ability to answer a question, namely, level of activation of the relevant structures. Note that this effect cannot be explained in terms of a simple lexical encoding effect. If the result were simply due to ease of encoding the words in the question, then the effect should be as large for the disjoint condition as the conjoint condition; on the contrary, the disjoint condition is at least as bad as the unprimed condition.

VI. Conclusions

This article has presented some of the reasons for the importance of considering the strategic components of memory retrieval when develop-

[8]Of the terms that were selected for priming, half were assigned to a three-sentence condition and half to a one-sentence condition. "Three" or "one" referred to the number of sentences that subjects were required to write for a given priming pair. This variable had no effect on the data.

TABLE IV

ANSWER GROUP ONLY: CORRECT ANSWER RTs (SEC). CONJOINT/DISJOINT SENTENCE-GENERATION EXPERIMENT

		Primed	
	Unprimed	Disjoint	Conjoint
Easy	3.05	3.16	2.33
Hard	3.33	3.73	2.75

ing a model of memory and question answering. The experiments described here help to identify the variables that affect this strategy selection, both when the selection appears to be under conscious control and when it appears to be an automatic by-product of processing the question. The variables *extrinsic* to a memory probe that influence strategy selection include prior history of success with a strategy—when the strategy is successful, the subject stays with it; when it fails the subject tends to adopt an alternative as the preferred strategy. Other situational variables, such as explicit advice about successful strategies, task instructions and knowledge of the age of the memories tested can also influence this bias in strategy selection.

Data were presented that support the view that strategy selection can occur while or after the question is understood, i.e., it does not have to be decided before reading the question. Evidence was reviewed that supports the idea that our feeling of knowing may be involved in this strategy selection. Several variables were shown to affect our feeling of knowing as well. Both recent exposure to concepts involved in the question and knowledge of the general topic referenced by the question affect one's estimates of ability to answer a question.

Further research is needed to uncover the mapping between feeling of knowing and strategy selection. Feeling of knowing must affect strategy selection. It seems obvious that if nothing is known about a topic, no question-answering strategy is used at all. In this case, people quickly say "don't know." Certain manipulations have been shown to influence feeling of knowing, and others have been shown to influence strategy selection. What we have not shown is that variables that manipulate feeling of knowing also influence the type of strategy attempted.

Surprisingly little work has been done on strategy selection in question answering. Work on strategy selection mechanisms is limited in general; however, there have been a few investigations in other domains (e.g., Dixon & Just, 1986; Payne, Bettman & Johnson, 1988; Siegler & Shrager, 1984). Siegler & Shrager looked at strategy use in children's arithmetic. They also believed that the competing strategies are carried out serially (as opposed to a parallel race with the faster strategy forcing a response). However, they also believed that for this domain, children would always first attempt the direct retrieval strategy and only try the computational strategy if the correct answer was not stored. More recently, Siegler (1987) believes that children do, in fact, choose between the competing strategies and that they can elect to use a computational strategy even when they have an answer stored for an arithmetic problem. He posits similar types of variables that might affect strategy selection, e.g., frequency of problem presentation, knowledge of related problems, and difficulty of using the computional strategy.

Payne *et al.* (1988; Bettman, Johnson, & Payne, 1987) have looked at adaptive strategy selection in the domain of decision making. They find that people are very adaptive at switching strategies. In this case strategy selection depends on the cognitive effort involved in a strategy, how difficulty interacts with time pressure, and the probability of error with a particular strategy. Their notion that cognitive effort and accuracy predict strategy use fits nicely with my own conception in that ease of use and accuracy of the direct retrieval strategy will vary with delay between reading and test and with task demands. Further, the probability of success with a strategy, either based on the ratio of presented to not-presented statements or based on the advice of the experimenter, is also consistent with the idea of effort/accuracy tradeoffs.

As a final note, it is important that memory researchers appreciate that people do not always search memory first for an exact match to a memory probe. Therefore, we cannot answer the question of which inferences are made during reading by examining differences in latencies among types of inferences. Some of the more recent research by McKoon and Ratcliff (1987) and Keenan, Potts, and Golding (1987) using word priming in a lexical decision task or in naming latencies are examples of promising approaches to understanding which inferences are made while reading. Question answering after a passage is read is a useful technique for understanding the strategies people use to answer questions, but not to infer the processes involved during reading.

VII. Appendix

This story, the "Riverboat Race," is an example of the type of story used in the experiments.

The following incident occurred in the river town of Napoleon, Arkansas.

A steamboat race was held on the 4th of July.
The Kentucky was favored to beat the Walter Scott easily.
The boats ran neck and neck over the first part of the course.
They were stripped of all ornamentation.
Then the Kentucky ran out of fuel.
The captain took charge of the situation.
There was no time to stop for more fuel.
The crew gathered on deck.
The Walter Scott began to pull far ahead.
The crew of the Kentucky began to lose hope.
They chopped up everything wooden on board.
They worked carefully.
The woodwork kept the fires going.
The big stern paddlewheel went on turning.
The Walter Scott was about to win.
The finish line was in sight.
The Walter Scott struck a sand bar.
Her bow was stuck fast.
All that was left of the Kentucky was the hull and engines.
The captain ordered full speed ahead.
She had no trouble with the sand bar.
The crowd was stunned as the Kentucky crossed the finish line.
The night was spent carousing.
The winners of bets bought drinks for the losers.
The next day business as usual resumed.

The following statements were to be judged for plausibility.

Moderately plausible statements:
- The race was a big event in the town.
- Ornamentation slowed down boats.
- Bets were being placed on the two ships.

Highly plausible statements
- The Kentucky's crew chopped and burned everything but the hull and the engines.
- The Kentucky's crew was going to use the wood for fuel.
- The Kentucky won the race.

Contradictory statements
- The Kentucky struck a sand bar.
- The losers of bets bought drinks for the winners.
- The first mate took charge of the Kentucky's situation.

Implausible statements
- The race occurred on a cold, windy day.

• Each boat was allotted the same amount of fuel.
• The townspeople had a very sober nature.

Acknowledgments

The work reported here was sponsored by Grant BNS-0371 from the National Science Foundation and, in part, by the Office of Naval Research, Contract No. N00014-84-K-0063, Contract Authority Identification Number NR667-529. I thank John Anderson and Gail Wells for their comments.

References

Anderson, J. R. (1983). Retrieval of information from long-term memory. *Science*, **220**, 25–30.

Anderson, J. R. (1985). *Cognitive psychology and its implications*. New York: Freeman.

Arkes, H. R., & Freedman, M. R. (1984). A demonstration of the costs and benefits of expertise in recognition memory. *Memory and Cognition*, **12**, 84–89.

Baddeley, A. (1986). *Oxford Psychology Series*, No 11. Working memory. New York: Oxford University Press.

Bettman, J. R., Johnson, E. J., & Payne, J. W. (1987). *Cognitive effort and decision making strategies: A componential analysis of choice*. Unpublished manuscript, Duke University.

Chiesi, H. L., Spilich, G. J., & Voss, J. F. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior*, **18**, 257–273.

Dixon, P. & Just, M. A. (1986). A chronometric analysis of strategy preparation in choice reactions. *Memory and Cognition*, **14**, 488–500.

Donders, K. A., Schooler, J. W., & Loftus, E. F. (1987). *Troubles with memory*. Talk presented at the 28th Annual Meeting of The Psychonomic Society, Seattle, WA.

Dooling, D. J., & Cristiaansen, R. E. (1977). Episodic and semantic aspects of memory for prose. *Journal of Experimental Psychology: Human Learning and Memory*, **3**, 428–436.

Hebb, D. O. (1961). Brain mechanisms and learning. In J. F. Delafresnaye (Ed.), *Distinctive features of learning in the higher animal*. New York: Oxford University Press.

Keenan, J. M., Potts, G. R., & Golding, J. M. (1987). *Methods for assessing the occurrence of elaborative inferences*. Talk presented at the 28th Annual Meeting of the Psychonomics Society, Seattle, WA.

Lewis, C. H., & Anderson, J. R. (1976). Interference with real world knowledge. *Cognitive Psychology*, **7**, 311–335.

Loftus, E. F. (1979). The malleability of memory. *American Scientist*, **67**, 312–320.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, **4**, 19–31.

Loftus, E. F., Schooler, J. W., & Wagenaar, W. (1985). The fate of memory: Comment on McCloskey and Zaragoza. *Journal of Experimental Psychology: General*, **114**, 375–380.

McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General*, **114**, 3–18.

McKoon, G., & Ratcliff, R. (1987). *Semantic association and inference processes*. Talk given at the 28th Annual Meeting of The Psychonomics Society, Seattle, WA.

Nelson. T. O., Gerler. D., & Narens. L. (1984). Accuracy of feeling-of-knowing judgments for predicting perpetual identification and relearning. *Journal of Experimental Psychology*. **113**, 301–323.

Owens. J., Bower. G. H., & Black. J. B. (1979). The 'soap opera' effect in story recall. *Memory and Cognition*. **7**, 185–191.

Payne. J. W., Bettman. J. R., & Johnson. E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, in press.

Peterson. S. B., & Potts. G. R. (1982). Global and specific components of information integration. *Journal of Verbal Learning and Verbal Behavior*. **21**, 403–420.

Postman. L. (1971). Organizing and interference. *Psychological Review*. **78**, 290–302.

Postman. L., & Underwood. B. J. (1973). Critical issues in interference theory. *Memory and Cognition*. **1**, 19–40.

Prytulak. L. S. (1971). Natural language mediation. *Cognitive Psychology*. **2**, 1–56.

Reder. L. M. (1976). *The role of elaborations in the processing of prose*. Doctoral dissertation University of Michigan. Available through University Microfilms. Ann Arbor.

Reder. L. M. (1979). The role of elaborations in memory for prose. *Cognitive Psychology*. **11**, 221–234.

Reder. L. M. (1982). Plausibility judgments vs. fact retrieval: Alternative strategies for sentence verification. *Psychological Review*. **89**, 250–280.

Reder. L. M. (1987). Strategy selection in question answering. *Cognitive Psychology*. **19**, 90–138.

Reder. L. M., & Anderson. J. R. (1982). Effects of spacing and embellishment on memory for the main points of a text. *Memory and Cognition*. **10**, 97–102.

Reder. L. M., Dennler. C., & Wells. G. (1985). *Conjoint vs. disjoint priming experiment*. Unpublished raw data. Carnegie-Mellon University.

Reder. L. M., & Fabri. S. (1984). *World knowledge experiment*. Unpublished raw data. Carnegie-Mellon University.

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.

Reder. L. M., & Wells. G. (1986). *Mixed delay experiment*. Unpublished raw data. Carnegie-Mellon University.

Reder. L. M., & Wible. C. (1984). Strategy use in question-answering: Memory strength and task constraints on fan effects. *Memory and Cognition*. **12**, 411–419.

Siegler. R. S. (1987). *How domain-general and domain-specific knowledge interact to produce strategy choices*. Unpublished paper. Carnegie-Mellon University.

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: The eighteenth annual Carnegie Symposium on cognition*. Hillsdale. NJ: Erlbaum.

Smith. E. E., Adams. N., & Schorr. D. (1978). Fact retrieval and the paradox of interference. *Cognitive Psychology*. **10**, 438–464.

Sulin. R. A., & Dooling. D. J. (1974). Intrusion of a thematic idea in retention of prose. *Journal of Experimental Psychology*. **103**, 255–262.

Tversky. B., & Tuchin. M. (1987). *Memory impairment by misleading postevent information: A reconciliation*. Talk presented at the 28th Annual Meeting of The Psychonomic Society. Seattle. WA.