Benjamin, A. S. & Bjork, R. A. (1996). Retrieval fluency as a metacognitive index. In L. Reder (Ed.), Implicit Memory and Metacognition. Mahwah, NJ: Erlbaum.

Chapter 14

Retrieval Fluency as a Metacognitive Index

Aaron S. Benjamin
Robert A. Bjork
University of California, Los Angeles

How readily information "comes to mind" is one index humans use to assess the accuracy of that information and, more generally, the adequacy of their own state of knowledge in a given domain. Such *retrieval fluency* provides, in fact, a useful heuristic: In general, information that is better learned, more recent, and more strongly associated to the cues guiding recall (or any combination of the three) will tend to be more readily retrievable. Fluent retrieval can, however, reflect factors unrelated to accuracy or degree of prior learning. In that sense, making appropriate use of retrieval fluency (or the lack thereof) as a metacognitive index becomes a problem of inference.

There are both practical and theoretical reasons why it is important to understand the metacognitive assumptions that underlie such inferential processes. On the practical side, understanding these processes can enable us to better construct regimens of training and practice that educate the learner's *subjective experience* as well as objective performance. As Jacoby, Bjork, and Kelley (1994) have emphasized recently, the reading an individual takes of his or her level of comprehension and competence can be as important as his or her actual comprehension or competence. The kinds of tasks for which we volunteer, whether we seek additional instruction, and whether we see ourselves "fit" for a difficult task, all rest on our reading of our state of knowledge or efficiency. To the degree that such readings may be in error, we can become a liability or hazard to ourselves and others, particularly in real-world settings such as air-traffic control and police work, where mistakes can be devastating.

From a theoretical standpoint, understanding the metacognitive assumptions that underlie the interpretation of subjective indices, such as retrieval fluency, has the potential to reveal how our mental model of ourselves as learners and rememberers succeeds and fails in capturing the complexity of our own memory. Because such an interpretation guides the selection and execution of control processes relevant to learning and remembering, understanding this mental model may provide insight into traditional failures in self-paced learning and memory use.

In the next several sections we summarize how fluency—perceptual fluency as well as retrieval fluency—influences (and sometimes misleads) a variety of metacognitive judgments. The latter half of the chapter presents a series of assumptions that appear to underlie the occasional misreliance on retrieval fluency as a predictive index. We demonstrate that subjects utilize retrieval fluency in their metamnemonic assessments, not only in circumstances in which retrieval fluency is an unreliable index, but also in circumstances where it provides for *completely backwards predictions.* That is, the reliance on retrieval fluency can apparently supercede other metamnemonic strategies to the point that subjects predict high future retention for items which they are unlikely to retain, and low future retention for items which they indeed are likely to retain.

USING ONE INDEX TO PREDICT ANOTHER

In general, we face an inferential problem when making judgments about the objective nature of a perception or a memory. In some sense, this problem may be dubbed the "New Look" at metacognition: We realize that our perceptions and memories are heavily influenced not just by the objective event or knowledge, but also by the host of previous experiences we bring to bear on that mental event. Judgments thus become inferences of how we think the world might be or how extensive we think our knowledge might be, given subjective data concerning the ease with which we perceive a stimulus or retrieve a memory. This task is accomplished by utilizing subjective indices—such as ease of perception or ease of retrieval—to predict the objective nature of events or memories.

Perceptual Fluency

The arguments we present in this chapter have been motivated in part by recent advances in the understanding of misattributions in memory (e.g., Jacoby & Kelley, 1987) and metamemory (e.g., Reder & Ritter, 1992; Schwartz & Metcalfe, 1992). This body of work has demonstrated that perceptual fluency—the sense of familiarity that a stimulus evokes, or ease with which a stimulus can be *perceived*—affects subjects' metacognitive judgments of their current level of knowledge and their future performance. In that sense, it may be said that perceptual fluency serves as a metamnemonic heuristic. Before we turn to an analysis of retrieval fluency as a heuristic, it is useful to summarize the influence of perceptual fluency on metacognitive judgments, because those influences illustrate inferential processes analogous to those we argue apply to retrieval fluency.

Misattributing Perceptual Fluency to Stimulus Characteristics. Jacoby and his colleagues (e.g., Jacoby, Allan, Collins, & Larwill, 1988) have documented a number of ways that the ease with which a subject perceives a stimulus affects metacognitive judgments. Although it is fairly useful to assume that the clarity with which a stimulus that is perceived reflects an underlying property thereof, such as accuracy or clarity, such may not always be the case. The oversimplicity of such a heuristic is clearly evidenced in the following demonstrations.

In an experiment by Jacoby et al. (1988), subjects were presented a series of trials on each of which a spoken sentence was played against a background of white noise. Some of the sentences (the *primed* sentences) had been exposed earlier in an ostensibly unrelated exercise and others had not. Subjects were asked to evaluate the level of background noise during each sentence presentation. Given the prior exposure, we can assume that subjects heard the primed sentences more clearly than the unprimed sentences. They attributed that subjective difference, however, not to its actual cause, but rather to a difference in the level of background noise—they rated the background noise as lower for the previously heard sentences than for the novel ones.

In another experiment, Witherspoon and Allan (1985) had subjects perform a perceptual identification task in which words were presented at very rapid rates and subjects were asked to identify the words if they could. The subjects were also asked to estimate the exposure duration for each word. Consistent with noise-judgment findings, subjects rated previously exposed words as having had a longer exposure duration than did nonprimed words. Again, the subjects misattribute an influence of prior priming to a stimulus characteristic—in this case, prolonged exposure time.

Such demonstrations illustrate the inferential nature of introspections. Attributions concerning the nature of one's performance are made to some degree "on the fly," and are easily misled in contrived experimental situations. However, these misattributions are even more dangerous when subjects make judgments concerning their own state of learning. Some evidence suggests that just such a reliance on perceptual fluency may underly errors in metacognition.

Misattributing Perceptual Fluency to One's State of Learning. Recent work (e.g., Reder & Ritter, 1992; Schwartz & Metcalfe, 1992) has demonstrated clearly that *feeling-of-knowing* (FOK) judgments are influenced by perceptual fluency. This phenomenon has been termed the *cue-familiarity effect.* Consider the following example from Reder (1987).

In Reder's "game-show" paradigm, subjects make rapid judgments as to whether or not they think they will be able to answer a given question. Such judgments can be made more quickly than actually producing the answer, yet are often of considerable accuracy. Reder found that priming words (like *golf* and *par*) that were to appear in questions such as "What is the term in golf for one under par?" led to increased subjective estimates of knowing the answer, despite the fact that such prior exposure did not alter actual rates of producing the correct answer at all!

Schwartz and Metcalfe (1992) manipulated both cue and response term accessibility in a paired-associate recall task. Targets were either generated or read during learning, and cue words were either preexposed in a previous pleasantness-rating task or not. Their results demonstrated that only cue priming (affecting perceptual fluency) and not target retrievability affected FOK judgments. Reder and Ritter (1992) found a similar effect using arithmetic problems as stimuli—prior exposure to the terms (but not necessarily the operand!) in a problem increased estimates of the retrievability of the answer from memory.

These examples illustrate that using perceptual fluency as a basis for judgments about the objective nature of perceptual events, or about the objective state of one's knowledge, is subject to inferential errors. Because we often fail to understand that prior exposure facilitates perception, we misattribute such perceptual fluency to other factors, such as physical aspects of the stimulus (e.g., loudness or duration) or knowledge that we do not actually have. The remainder of this chapter focuses on how *retrieval fluency*—the ease with which information is accessed from memory—affects metamemory in similar ways. Attributing retrieval fluency to one's state of knowledge when, in fact, that fluency arises from sources unrelated to one's knowledge state, can be at least as insidious—and in applied settings, at least as dangerous (Jacoby et al., 1994).

Retrieval Fluency

From a phenomenological perspective, we are all aware that some things seem to "come to mind" more easily than others. We are inclined to relate this ease to our level of knowledge—we know our own phone number better (usually) than we know our friends' numbers; thus its retrieval is marked by greater facility. Not surprisingly, this relationship appears in controlled experiments: Overlearned information is retrieved more quickly and with greater accuracy than less well-learned information (e.g., Nelson, Leonesio, Shimamura, Landwehr, & Narens, 1982).

As stated earlier, one goal of the remainder of this analysis is to outline cases in which humans seem to use retrieval fluency as a metacognitive index. A parallel goal is to consider those cases in which monitoring the dynamics of retrieval provides the metacognizer with misinformation. Using one index to predict another illustrates a failure to appreciate differences between the predictive task and the to-be-predicted behavior and illuminates aspects of human memory, the complexities of which are clearly often misunderstood by its users.

First, however, it is important for the purposes of the present analysis to more precisely define what we mean by *retrieval fluency*. Moreover, we review some important *objective* determinants of retrieval fluency.

THE DEFINITION AND DETERMINANTS OF RETRIEVAL FLUENCY

Aspects of Fluent Retrieval

When we speak of "fluent" retrieval of information in response to a cue, three characteristics of access are relevant.

Latency. First, we can speak of how quickly information is accessed by a certain cue. Given the cue *early American presidents*, for example, we typically experience greater fluency of access for George Washington and Thomas Jefferson than for John Quincy Adams and Andrew Jackson. The latter names can indeed be recalled, but typically only after a greater delay than their more accessible counterparts.

Persistence. Second is the persistence of the information that does come to mind. When asked to recall wars in which the United States has been involved in the current century, a Vietnam veteran may find that particular war coming to mind more often than would a veteran of the Second World War. We would thus infer greater fluency of access to the Vietnam War for the Vietnam vet than by a World War II vet.

Amount. Third, retrieval fluency involves the raw amount of information accessed by a given retrieval cue. A teenager may thus have greater retrieval fluency for rock-and-roll artists than composers of baroque music, whereas that relationship might reverse for one or both of his or her parents.

High retrieval fluency thus generally refers to retrieval which is quick, persistent, abundant with information, or any combination of those three effects. We turn now to a cursory discussion of the determinants of retrieval fluency.

Objective Determinants of Fluent Retrieval

We began our discussion of the retrieval-fluency heuristic by noting that reliance on such an index is, in general, not a bad idea. In fact, the ease with which we retrieve information is often closely related to the levels of knowledge and surety appropriate for metamnemonic judgments and the selection and execution of control processes. The following paragraphs delineate some of the major determinants of retrievability and retrieval fluency.

Degree of Learning. In general, speed and the reliability of access to information will, indeed, be correlated with levels of learning. How elaborated a learned representation is in terms of linkage to other concepts and schemas will also influence the reliability of access to that representation and the amount of information retrieved.

Frequency and Recency of Usage. The act of retrieval itself makes the retrieved information more retrievable in the future, as we stress later in this chapter. Information that is accessed frequently, especially in multiple contexts, will tend to be retrieved fluently in the future. The current fluency of access to information in memory will also be influenced by the recency of prior access.

Episodic Distinctiveness. A factor that influences ease of access to episodic traces is episodic distinctiveness. Uniqueness, salience, emotionality, temporal isolation, and other factors all play a role in the extent to which an event is "distinct." The relatively ready access to "flashbulb memories," such as where one was and what one was doing when one heard that the space shuttle Challenger exploded, is one example of the influence of such episodic distinctiveness.

Cue Sufficiency. Retrieval processes are driven by cues, and fluency of access to information is heavily influenced by the effectiveness of those cues. The effectiveness of a cue or configuration of cues depends, in turn, on such factors as strength of association of the target information to the cue(s) in question, and the extent to which that information is uniquely associated to that (those) cue(s). When there is competition; that is, when cues are associated to multiple items in memory, retrieval of any given item is slower and less reliable. This principle is consistent with the notion of *cue overload* (see e.g., Watkins & Watkins, 1975; cf. fan effect, Anderson, 1974)—that a cue loses its potency in evoking any given response when associated with a greater number of potential responses.

Priming. Finally, prior presentation of some target item, even if that presentation occurs in a context nominally unrelated to a current task of some kind, has been shown to increase the speed and frequency with which that item, among other possible items, is accessed in response to a cue that is part of the task in question.

It is clear that many of the characteristics which make retrieval fluent are ones that are entirely appropriate to rely on as an index of what we know, or how well we know it. Because, however, spurious factors do influence retrieval fluency, it is also possible that metacognitive predictions which are driven by subjective assessments of retrieval fluency can be misled. The following section reviews evidence which appears to support the proposition that retrieval fluency is used as a heuristic tool—and that it often fails in circumstances where retrieval fluency derives from a confluence of factors, some of which are related to degree of learning, and some of which are not.

RETRIEVAL FLUENCY AS A HEURISTIC

Confidence in Retrieved Answers

Costermans, Lories, and Ansay (1992, Experiment 1) provide evidence for a strong relationship between the latency to answer a question and the confidence one has in that answer. Subjects in their experiment answered a series of general-information questions and rated their confidence in those answers. Costermans et al. found a strongly negative gamma correlation (G) between response latency and confidence ($G = \sim -.5$). Open to interpretation, of course, is the nature of the relationship: Are response speed and judged confidence *both* a function of actual accuracy of the answer, are we more confident in answers *because* we produce them quickly, or are we just more quick to provide answers in which we have high confidence? Several pieces of evidence bear on this question.

First, Costermans et al. (1992) found that the relationship between response latency and confidence held for both correct and incorrect answers. If, indeed, a third variable, such as response accuracy, wholly mediated the observed correlation, then we would expect a greatly attenuated effect for incorrect answers. In fact, in their data, Costermans et al. found the relationship to be somewhat stronger for incorrect answers.

Even more convincing is a demonstration by Kelley and Lindsay (1993). They had subjects answer questions such as "What was Buffalo Bill's last name?" For each such question, during a prior phase of the experiment, subjects were preexposed to either a correct answer (Cody), a plausible but incorrect answer (Hickock), or an unrelated term (Letterman). With respect to the former two conditions, they found that subjects who provided the primed name were more confident than if they had not been primed—*regardless of whether that answer was correct.* In other words, increasing the retrieval fluency for an answer via priming increased confidence in that answer independent of the correctness of that answer. In a subsequent experiment, the authors actually made explicit that the preexposed words contained answers, some correct, some incorrect, to the questions to be asked. Alerting subjects to that fact did not change the basic pattern of results. Such demonstrations clearly indicate that subjects use the speed of retrieving an answer to a question as a basis for confidence in the accuracy of that answer.

A further demonstration by Shaw (1996) illustrates the manner in which such misguided confidence may have real-world consequences. Subjects in Shaw's experiment were exposed to the aftermath of a mock crime presented on a series of slides. After viewing the "crime scene," they were questioned about objects present or absent in the scene. In the interval separating the two questioning sessions, half of the subjects were asked to think about the event and carefully consider the answers they had previously provided and whether or not those answers were indeed correct. Both those subjects and the control subjects (who engaged in various unrelated distractor tasks during the same period) were then asked the same questions again. Both groups chose the same answers again most of the time, but the subjects asked to "mull over" their answers were more confident in those answers at the time of the second test, even though the actual accuracy of those answers was not different. We argue that the "mere thought" manipulation increased the retrieval fluency of the answers chosen earlier, thereby increasing confidence. The similarities between Shaw's mere-thought condition and the common experience of an actual crime eyewitness who is typically subjected over and over to interrogations and discussions about a crime further substantiates the importance of understanding the heuristic inferential nature of such metacognitions.

Koriat, Lichtenstein, and Fischhoff (1980) provided a theory for the basis of confidence that is relevant to such findings. They attempt to explain the perpetual overconfidence subjects exhibit (a) in the accuracy of their beliefs (e.g., Fischhoff, Slovic, & Lichtenstein, 1977) and (b) the degree to which those beliefs are representative of the beliefs of others (as in the *false consensus effect,* e.g., Marks & Miller, 1985). In their theory, people are overconfident because they selectively come up with reasons consistent with the answer they have provided. Support for such an interpretation derives from several findings that forcing subjects to provide an equal number of reasons for and against a proposition attenuates the overconfidence bias. By the present reasoning, we would argue that reasons that support answers that had been provided are retrieved more fluently than reasons to the contrary. Whether such fluency is the basis for confidence, or arises as a function of committing oneself to a decision, is unclear.

Estimation Biases

In their seminal analysis of probability estimation, Tversky and Kahneman (1974) documented a number of ways in which such assessments may go awry. One oft-cited bias lies in the use of the *availability heuristic.* This heuristic involves using the ease with which an instance of an event comes to mind as a basis for estimating the probability of that event. Consider a case in which subjects are asked, for each of the following, to judge the number of words present in a typical four-page segment of a novel which would satisfy the fragment constraints provided (from Tversky & Kahneman, 1983):

(a) -----n-   (b) -----ing

The first fragment conjures up little, if anything, in the way of specific completions. After some effort, one may generate a few instances, but their retrieval is marked by considerable difficulty. The second example, however, evokes a great number of immediate candidates. It does so probably because it somewhat constrains the potential response set to familiar gerunds, such as *running* or *playing.* However, because the set of completions for fragment (b) is a logical subset of the set of completions for (a), it is obvious that fragment (a) allows for a greater number of completions. The merits of logical reasoning notwithstanding, subjects typically estimate a larger number of word responses for (b), thus exhibiting a preference for a suboptimal decision strategy that utilizes retrieval fluency over an infallible one that involves syllogistic reasoning.

Such misestimation is also evident in the evaluation of accident probabilities. Subjects are quite likely to underestimate the probability of being in an automobile accident; however, they typically overestimate the probability of being in a plane crash. The explanation provided by Tversky and Kahneman lies in the differential salience of the events: Plane crashes receive more media attention and are a more "glamorous" event. These kinds of events "pop to mind" more quickly and, by the present analysis, are misinterpreted on that basis as being more frequent.

Predicting Future Performance

Feeling-of-Knowing Phenomena. The feeling-of-knowing (FOK) phenomenon was first documented by Hart (1965). In his experiments, the procedure of which is quite similar to the more recent studies discussed later, subjects first attempted to answer a number of general-information questions. For those cases in which they were unsuccessful (when they provided either an incorrect answer or no answer at all), they then indicated whether they felt that they would or would not be able to recognize the correct answer among four alternatives. In fact, subjects were quite able to predict their own later recognition performance.

Explanations of the FOK effect have typically fallen into two major categories: *trace-access* theories and *inferential* theories (Nelson, Gerler, & Narens, 1984). Trace-access theories view the metacognizer as having some access to the object of judgment (i.e., the memory trace under evaluation) that serves as a basis for prediction. Such theories posit a monitor that has access to aspects of memory that cannot be accessed via conventional attempts to recall. Trace-access explanations have historically been invoked (e.g., Hart, 1965) to explain why, in the absence of recall, subjects can nonetheless discriminate between cases where they will and will not be able to recognize an answer.

Inferential theories involve the overt stipulation that no such privileged access exists. They posit, instead, an inferential basis for metamnemonic judgments. An example is the idea that retrieval-cue familiarity affects subjects' decisions about the impending retrievability of an associated target (Reder & Ritter, 1992; Schwartz & Metcalfe, 1992). Central to such theories is the heuristic nature of such judgments; because predictions are *inferred*, they are fallible and open to being misled.

It is not the goal of the present chapter to argue the relative merits of trace-access and inferential theories, but we take as a working assumption that metacognitive judgments are inferential in nature. In that context, the question of interest to us is a paradox of sorts: How is the claim to be made that retrieval fluency serves as the basis for FOK judgments when there is, by definition, no retrieval of the to-be-judged answer prior to those judgments?

One argument has been that the FOK derives from partial information retrieval. Presumably, in a paradigm like the one of Hart (1965), enough information relevant to the answer comes to mind to support a FOK judgment, but not enough to form a coherent answer. Support for such an idea was presented by Blake (1973) in his analysis of partial response production in the retrieval of nonsense trigrams. He found that higher FOK ratings were given to nonrecalled trigrams for which one or more of the correct letters could be provided. Similar results were reported by Schacter and Worling (1985) using less abstract stimuli. In their study, subjects studied unrelated paired associates, the targets for which were of either a "good" (e.g., *happy*) or a "bad" (e.g., *terrible*) connotation. Schacter and Worling found that FOK ratings were higher for those unrecalled targets for which the affective connotation could be produced.

Several other experiments provide results consistent with the proposal of Blake and Schacter and Worling. Nelson et al. (1982) showed that items learned to a higher criterion (4 correct recalls) evoked greater FOKs when unrecalled than other items learned to a lower criterion (1 correct recall). The difference between the two sets of items is not apparent in the actual production of an intact answer, but those item sets are likely to differ in the extent to which partial information is retrieved. Another example that "better" processing leads to higher FOK judgments is provided by Lupker, Harbluk, and Patrick (1991), who demonstrated that deeper levels of processing resulted in higher FOK judgments than did more shallow levels. Also consistent with the notion of partial retrieval serving FOK judgments is the finding that errors of *commission* evoke greater FOKs than errors of *omission*, even when subjects are informed of the incorrectness of their answer (Krinsky & Nelson, 1985).

A potential problem with the notion that partial information retrieval underlies FOK judgments is that it seems to blur the distinction between trace-access and inferential theories. If the partial information that does come to mind cannot contribute to the selection of an answer, how does the monitor "know" that such information is related to the answer? Do we need to posit a monitoring homunculus that has access to retrieval information that the response-production homunculus does not? A resolution of this sticky issue is provided by Koriat (1993, 1995).

Koriat's answer, put simply, is that the monitor does not, in fact, "know" anything about the relationship between the partial information retrieved and the currently unproducible correct answer. In his theory, FOK judgments are supported by the raw amount of information that comes to mind in response to a cue—whether that information is correct or incorrect. The following experiment supports this notion convincingly.

As material for his experiments, Koriat (1995) developed a series of questions that varied on two dimensions: *accessibility* (ACC) and *output-bound accuracy* (OBA). Accessibility refers to the number of responses that the questions evoked; some questions elicited many responses and others few. Output-bound accuracy refers to the proportion of answers to a given question that were correct; questions which were most often answered correctly when they were answered at all were said to have high OBA; questions principally answered falsely when an answer was given were said to have low OBA. Koriat's results are reproduced in Fig. 14.1.

As shown in Fig. 14.1, items that differ in ACC levels evoke very different FOK judgments. In general, high ACC translates into high FOK judgments. Items that have high ACC and high OBA (*consensually correct*, or CC, items) provide for accurate FOK judgments: High predicted accuracy translates into actual success on the to-be-predicted recognition test. However, items that are high ACC but low OBA (*consensually wrong*, or CW, items) yield

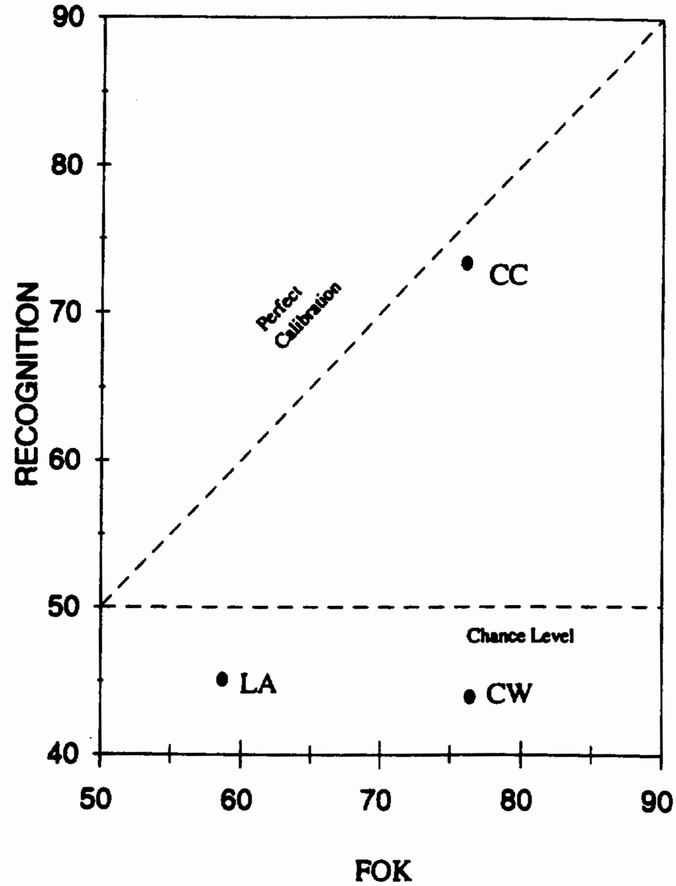

FIG. 14.1. Recognition scores and FOK ratings for consensually correct (CC), consensually wrong (CW), and low accessibility (LA) items. Adapted from Koriat (1995). Reproduced with permission.

judgments of future performance that are wild overestimates. In fact, recognition accuracy is not above chance for such items. Koriat uses such a demonstration to support the idea that the feeling of knowing derives from sources different than actual knowing. Namely, it is argued that FOK arises as a result of ready recall of some information, right or wrong.

Judgments of Learning. Judgments of learning (JOLs) involve having the subject evaluate the level of his or her knowledge during the course of learning. In a typical paradigm, subjects might cycle through a list of words and, after studying each word, estimate their probability of being able to free recall that word at some future time. Subjects do so with above-chance accuracy (e.g., Lovelace, 1984), and their predictions seem to take into account more than the current recallability of the item (Mazzoni & Nelson, 1995).

Again, however, it appears as though retrieval fluency serves as a potent cue for such judgments. Lee, Narens, and Nelson (1993, as cited in Narens, Jameson, & Lee, 1994) demonstrated a case in which a temporary enhancement of retrieval fluency misleads JOLs. In their experiment, subjects studied paired associates and, 3–5 minutes after the presentation of a particular pair, engaged in the prediction task. In that task, they were presented with just the cue word and asked to estimate their probability of being able to recall the target. The clever twist in the Lee et al. experiment, however, involved a subthreshold presentation of some of the correct target words immediately prior to the prediction. In those cases, Lee et al. hypothesized that the priming would influence target retrieval in such a manner as to temporarily inflate JOLs. In fact, their results bore this idea out: JOLs were indeed higher for the primed items, yet the "transitory nature of this kind of priming effect" left the final-recall rates unaltered. Clearly, potent but fleeting changes in retrieval fluency profoundly affect JOLs.

The conceptualization of retrieval fluency as a basis for the JOL also suggests an interpretation of a phenomenon that has been seen somewhat as an enigma in the JOL literature. Dunlosky and Nelson (1992, see also Nelson & Dunlosky, 1991) described what has been called the "delayed-JOL" effect. This "effect" refers to the fact that JOLs are (a) more accurate when the subject is presented with only the cue and not the target when making their prediction and (b) more accurate at a delay than immediately after study.

An evaluation of these circumstances makes it clear that, for the prediction to be accurate, two sources of information have to be intact. First, there must be *retrieval*. When subjects are presented with the entire cue–target pair, the necessity of retrieval and thus its diagnostic value are obviated. Subjects are thus deprived of the intact subjective experience that would inform their JOL. Second, there must be a degree of *fluency* that matches the degree on the impending test. That is, immediate predictions are clouded by recency and other factors that attenuate the diagnosticity of the retrieval. If subjects are in a situation in which their retrieval fluency is in fact predictive of later retrieval probability (as is more likely to be true in the cue-only, delayed prediction case), their predictions will be more accurate. Conversely, experimentally making retrieval fluency nondiagnostic leads to poorer predictions.

USING CURRENT FLUENCY AS A PREDICTOR OF LATER FLUENCY

Clearly, it makes sense to use retrieval fluency as an indicator of what we know. More often than not, the factors that currently support retrieval will act to support retrieval in the future as well. Using such a heuristic in a given situation, however, reflects the tacit belief that certain assumptions have been met. Those assumptions include:

1. That the retrieval cues at the to-be-predicted time, and the retrieval task itself, will not differ substantially from the current cues and task, OR that any such differences will not measurably affect performance.
2. That events and the passage of time between now and the later task of interest will not alter the relative accessibility of competing representations in memory.
3. That the act of retrieval during prediction does not appreciably affect the relative ease of access to competing representations at the later time of interest.

In any predictive task, these assumptions may or may not hold; when they do, current retrieval fluency will be a good index of later performance. There is, however, considerable evidence that *none* of the preceding assumptions are always valid, and, in fact, that their being true simultaneously may be the exception rather than the rule. The following sections treat each of these assumptions in turn.

Conditions Now Versus Conditions Later

In general, our performance on a given task at a given time will be sensitive to the variety of cues available to us at that time, including environmental, social, body-state, and mood-state cues, as well as task-specific cues. The first assumption may fail to hold because the later task setting, although nominally the same as the current setting, may actually differ significantly in the cues that are or are not available. That assumption may also fail because the task itself, although nominally the same or similar to the current retrieval task, may actually differ in ways that fundamentally alter the nature of the retrieval process.

Overlap of Cues. It has been pointed out by a number of writers (for a recent example, see Christina & Bjork, 1991) that every test of retention is actually a test of transfer. The point is that the conditions at the time of any later test will necessarily differ from present conditions, if for no other reason than that the performer himself or herself will not be exactly the "same" person in terms of body state, mood state, and mental set. In addition, superficial and not-so-superficial cues may also differ. A retention test therefore becomes a test of whether performance will transfer to a situation that is functionally new, at least to some extent, even if nominally the same.

From a research standpoint, we have known for decades that the ability to retrieve learned information is sensitive to the overlap of cues between a current learning or test environment and some later criterion time. McGeogh (1932), for example, cited "altered stimulating conditions" as one factor in his three-factor theory of forgetting. He argued that one factor responsible for forgetting is the extent to which the stimulus conditions at test differ from those present at learning. The experimental findings that gave rise to Tulving and Thomson's (1973) *encoding specificity principle* are one clear illustration of such effects.

From the standpoint of our own experience, however, there are indications that we fail to appreciate the extent to which our access to skills and knowledge may falter as conditions change. In real-world training environments, it is a continuing surprise and frustration to training personnel that their trainees, who may have passed every criterion test with ease at the end of training, frequently make on-the-job errors under new or unanticipated conditions, even when those conditions seem to differ only superficially from the conditions of training (for a discussion of such training issues, see Christina & Bjork, 1991, and Reder & Klatzky, 1994). Trainees themselves—with hindsight—are often in fact dismayed that they were unable to access the appropriate procedure or knowledge in some new situation. In effect, as metacognizers, we do not fully realize the extent to which our present performance is tied to present cues.

Overlap of Tasks. We may also fail to realize the extent to which some current task—our performance, which serves as the basis for predicting our future performance—may differ from some later task in the demands it places, or does not place, on our memories. In general, when using current retrieval fluency as an index, we are subject to a certain misconception about memory—namely, that performance on different tasks derive from the same underlying memory traces, which vary in strength on a unidimensional continuum.

Contemporary memory research has produced a variety of results that discredit any such construct of strength unidimensionality. Associationist notions of "strength" fail to explain important phenomena in both human and animal learning (e.g., blocking, Kamin, 1969). The following phenomena are among those that demonstrate the multidimensionality of human memory.

The *word-frequency paradox* is one such result. When subjects study a list of both high- and low-frequency words, they are more able to produce the high-frequency items on a test of free recall. On a recognition test, however, low-frequency words are better recognized than high-frequency words. If recall and recognition both reflect the strength of a unidimensional trace, such a pattern could not obtain.

More recently, much interest has been devoted to understanding dissociations between *explicit* and *implicit* tests of memory. Explicit tests, such as recall or recognition, overtly stipulate that subjects should access their memory for a given prior episode. Implicit tasks involve the measurement of performance on a task that is nominally unrelated to some prior episode of interest. For example, subjects may be asked to resolve a lexically ambiguous sentence. The measure of "performance" is the degree to which the "primed" group—the group that was exposed to a study episode of some sort—has a greater proclivity than a control group to disambiguate the sentence in a manner influenced by their prior study. Such influences typically happen in the absence of the subject recognizing any overt contingency between the study and test episodes; in fact, often, in such experiments, the two phases are deliberately contrived so as to appear unrelated.

Clear dissociations have been found in which manipulations affect performance on an explicit test in one way and, on an implicit test, in an opposite way. For example, the *generation effect* (Slamecka & Graf, 1978)—that recall or recognition is typically better for items that are generated rather than read—may turn to a real advantage on indirect tests such as perceptual identification (Jacoby, 1983).

It is outside the mission of this chapter to document in any detail the variety of dissociations that have been demonstrated on direct and indirect tests of memory. The important point is that there is abundant evidence that memory is multidimensional. Although there is no clear consensus on the number of dimensions necessary to fully explain the wealth of data on memory, estimates of such a number have ranged as high as 50 (Tulving, 1983). It should suffice to say that performance on one test does not necessarily predict performance on another.

Changes in Fluency With Time and Intervening Events

To argue that we *forget* is to engage in egregious understatement. A fundamental quality of human memory (and potentially an adaptive one; see Bjork, 1989) is that the ease with which we access information from long-term memory decreases with disuse of that information over time. As is evident from Ebbinghaus' classic work on nonsense syllable learning, our ability to retrieve learned information can drop dramatically with the passage of time. Classic work on interference processes, and more recent work on retrieval dynamics, make it clear, however, that time itself is not the crucial factor responsible for changes in retrieval access over time.

Unlearning and Spontaneous Recovery. The research of Postman, Underwood, and others using paired-associate paradigms, such as the A–B, A–C paradigm, demonstrates the importance of cue uniqueness in the elicitation of a paired target term (for a review, see Postman, 1971). The pairing of a cue with a new target term impairs cued recall of the original response, thus producing retroactive interference. Clearly, however, memory does not work so simply as to "overwrite" the old information—with the passage of time and the forgetting of more newly paired responses, *spontaneous recovery* of the original target terms is evident. Thus, any completely accurate prospective evaluation of the retrievability of a target memory must incorporate information about interfering learning and the retrievability of that learning—a tall order for any metamnemonist, particularly when the learning has yet to take place!

Retrieval-Induced Forgetting. Recent work by Anderson, Bjork, and Bjork (1994) demonstrates that retrieving some information in response to a given cue can inhibit retrieval access to other information associated with that cue. Competitive dynamics at any point in time influence the retrievability of a target memory, even when the ongoing activity may bear little superficial relation to the original encoding episode. These dynamics may be predictable at the time of metamnemonic evaluation, as they may be intrinsic to the task on which performance is to be predicted. Incorporation of such information into metamnemonic judgments is hardly an obvious strategy, but it is—at least in principle—available at the time of prediction. For example, knowledge of these competitive dynamics might allow a subject to predict lowered recall of remaining-list items when half of the members of an originally studied list are presented as "cues" at the time of test (as in the *part-list cuing* effect, cf. Brown, 1968).

Shifts From Recency to Primacy. Other changes in the retrieval environment from assessment to test may violate the assumption that the relative accessibility of different items will remain the same across the retention interval. Some such changes require the subject to recognize that short-term influences are affecting current performance (during prediction), but that those influences are likely to be absent at the later time in question. For example, the *recency* effect in traditional serial-position curves of immediate list recall has been shown to disappear if the test is delayed even by 30 seconds, whereas recall of the primacy and middle portions of the list is essentially unchanged (Glanzer & Cunitz, 1966; Postman & Phillips, 1965). Figure 14.2 demonstrates an even more striking change from positive recency in immediate free recall to "negative recency" in end-of-experiment free recall (Bjork, 1975; Craik, 1970).

The top curve in the top panel of Fig. 14.2 shows the probability of recall of a word as a function of the position of that word in its original input list, averaged over a series of lists presented in succession. Note that immediate recall exhibits the typical primacy and recency effects. The

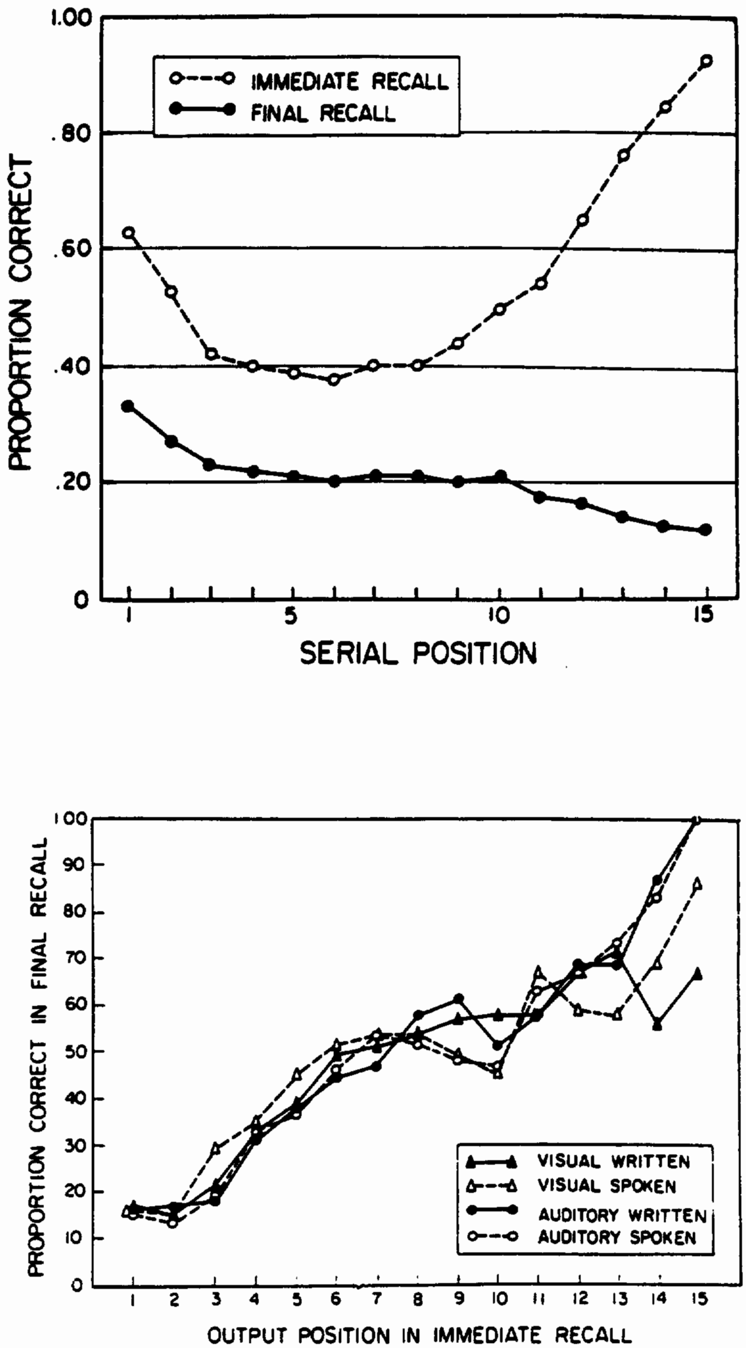

FIG. 14.2. Top panel: Probability of recall on immediate and delayed tests of retention as a function of initial input position. Bottom panel: Probability of recall on an end-of-experiment test as a function of an item's original output position on an immediate test of recall. Adapted from Craik (1970). Reproduced with permission.

bottom cu
a function
of the exp
of the pric
still recalle
lists are nc
subjects wl
their own
only the te
fluency), b
term mem

Episodic
retrieval flu
method, f
remember
distracting
beyond th
were interp
1986; Hitc
terms of te
given set
recall, def
items fron
interval, r
interval in
items in m

In gene
isodically
stant. To
standing t
be suppor
are now r
will not n

In the
another.
and unles
is unpred
and poter
one migh
target "ki
cue "fruit
incidental
facilitate

bottom curve in the top panel shows the probability of recall of a word as a function of its initial input position when subjects are asked, at the end of the experiment, to free recall any and all words that they can from any of the prior lists. Note that words that were at the beginnings of lists are still recalled with greater frequency, but words that came at the ends of lists are now suppressed as compared to their original recall levels. Thus, subjects who, at the time of immediate free recall, are asked to predict their own later recall performance on individual items must evaluate not only the temporary accessibility of a particular word in that list (i.e., retrieval fluency), but also adjust such estimates for the misleading effects of short-term memory.

Episodic distinctiveness, another factor cited earlier as a determinant of retrieval fluency, can also change with time. With the continuous-distractor method, for example, in which the presentations of successive to-be-remembered items are separated in time by interpolated periods of a distracting activity of some type, there are recency effects that extend well beyond the range of short-term memory. Such "long-term recency effects" were interpreted by Bjork and Whitten (1974; see also Glenberg & Swanson, 1986; Hitch, Rejman, & Turner, 1980, as reported in Baddeley, 1990) in terms of temporal distinctiveness. To the extent that the presentation of a given set of items remains temporally distinct in memory at the time of recall, defined by whether the temporal interval that separated that set of items from adjacent sets is above some fraction of the current retention interval, retrieval of that set of items will be facilitated. As the retention interval increases, however, the functional separation of successive sets of items in memory becomes insignificant, and such recency effects disappear.

In general, as memories become more and more distant, the less episodically distinct they become in memory, other factors being held constant. To predict future access reliably, then, requires not only an understanding that retrieval currently facilitated by short-term memory will not be supported by short-term memory in the future, but also that items that are now readily accessible because they are temporally distinct in memory will not necessarily enjoy that same advantage in the future.

In the most general sense, cues change from one point in time to another. Aspects of retrieval that are cue-driven are thus highly variable, and unless one has knowledge of future cues, some portion of that variance is unpredictable. In some simple cases, such knowledge is indeed available and potentially highly useful in metamnemonic prediction. For example, one might expect to see different assessments of the recallability of the target "kiwi" between a retrieval circumstance with no cues, one with the cue "fruit," and one with the cue "fruit–k__." However, the momentary incidental cues available during encoding, the reoccurrence of which may facilitate later recall (see Richardson-Klavehn & Bjork, 1988), are often not consciously perceived and are thus not available to the predictive apparatus. Recapitulating a particular bowel state similar to one experienced during learning may facilitate recall, but it is unlikely that subjects perceive such a tangential variable to be an important cue.

Retrieval as a Memory Modifier

The act of retrieving a memory has been shown to be a potent influence in increasing later accessibility of that memory (e.g., Bjork, 1975). Thus, to the degree that a predictive task involves retrieval, covert or otherwise, one must adjust one's estimates of how retrieval fluency at the time of prediction maps onto retrieval fluency at the time of interest. In fact, dissociations have already been noted between predictive tasks that incorporate retrieval and those that discourage it. Dunlosky and Nelson (1992, 1994) found that cue-only prompts were more effective in promoting JOL accuracy than intact cue-target prompts. One clear explanation is that retrieval fluency in the cue-only case was more diagnostic of later retrieval. Furthermore, it has been argued that the necessary retrieval in that case affected future recall to the degree that the act of retrieval during prediction made the prediction serendipitously accurate (Spellman & Bjork, 1992).

Moreover, not all retrievals are created equal. In fact, it appears as though a more difficult retrieval facilitates later accessibility to a greater degree than does an easy one (Bjork & Bjork, 1992). One corollary of such a hypothesis is that a successful retrieval that is delayed by an amount of time t will be more effective in fostering future retrievability than will a successful retrieval at time t_2, $t_2 < t$. Such a relationship appears to hold and is evident in the phenomena of spacing effects and the efficacy of expanding retrieval practice (Landauer & Bjork, 1978). Successful metamnemonic prediction must thus adjust for the beneficial effects of retrieval, but in a counterintuitive manner—those traces that are difficult to retrieve will become *relatively more accessible* as a result of that retrieval than will traces which are initially easy to retrieve. In order to adjust appropriately, then, retrieval fluency must thus be used in a manner counter to the way in which it has been argued to be used here!

WHEN RETRIEVAL FLUENCY IS COUNTERDIAGNOSTIC: TWO EXAMPLES

As we have argued at some length in the preceding sections, retrieval fluency on one task at one time may prove to be a misleading index of retrieval fluency at a later time on a different task, or even on the same task. Benjamin, Bjork, and Schwartz (1996) set out to examine the extent to which subjects understand some of the dynamics summarized earlier. The first experiment to be discussed here examined whether subjects understand that fluent access to items in immediate free recall may not be accompanied by fluent access to those same items in delayed free recall. The second experiment examined whether subjects understand that the fluency of access to information from semantic memory may or may not predict the later access to that information from episodic memory.

The Recency-to-Primacy Shift

Earlier in this chapter we discussed the manner in which end-of-list (recency) items tend to be recalled with high probability on an immediate test of retention, but "suffer" more than items from the remainder of the list over a delay (again, see the top panel of Fig. 14.2). Another way in which such a result can be characterized is in the change in the makeup of the recall protocol from an immediate test to a delayed test. This change has been called the "recency-to-primacy shift," reflecting the fact that, whereas recency items are recalled very well immediately, only the primacy items show such enhanced end-of-experiment recall. This result presumably reflects two major contributing factors:

1. Because recency items are so easy to recall immediately after learning, they are typically provided before other items (middle or primacy) in the recall protocol. In that sense, they are privy to greater retrieval fluency than other items, yet this fluency is short-lived. Because such easy retrieval owes to the items' immediate availability and not to particularly good long-term memory, later recall of those same items suffers relative to the rest of the list. Again, this relationship is expressed in the serial position curves displayed in the top panel of Fig. 14.2.

2. Those items that are produced with some difficulty on an initial recall derive more benefit than easily produced items in terms of later retrievability. Thus, those items that are produced quickly, with ease, on an initial recall (i.e., recency items) enjoy less retrieval enhancement as a function of that recall than items produced later in the recall protocol. This relationship is shown in the bottom panel of Fig. 14.2, which plots the probability of end-of-experiment recall as a function of the serial *output* position of a word on an immediate test of retention. It is apparent that items which are produced later, with difficulty, on an immediate test of retention are more likely to be recalled at a delay than items which are produced on the immediate retention test with ease (i.e., earlier in the output). The data from Craik (1970) demonstrate this relationship, but are subject to a possible subject selection bias; namely, items constituting the data plotted at the higher serial positions come from increasingly few

subjects, and, in fact, from a particular subset of subjects who are demonstrably better at recall than the average subject! However, Bjork (1970) demonstrated that this relationship holds even when each individual subject's output is normalized into quartiles, and the analysis is performed on output quartile and later recall probability.

If, as hypothesized here, subjects rely principally on current retrieval fluency as an index for later retrieval ability, we might expect them to make gross errors in their estimates of what they are likely to remember at a delay when the predictions are made after an immediate (misleading) test. In fact, such mispredictions would be the strongest evidence possible for the retrieval fluency heuristic presented here—not only would subjects fail to predict their later recall qualitatively, their predictions would be opposite to their eventual recall!

In fact, such errors are made. Benjamin et al. (1996, Experiment 2) showed that subjects engage in such backwards prediction in the following paradigm. As in the experiment described earlier, subjects studied a series of six lists of 13 words each and recalled each list immediately after study. After studying and recalling all six, they were given a final free-recall test for all of the items on all of their lists.

In addition, however, subjects were asked to make a prediction for each word as they recalled it on the immediate test. In fact, subjects only made such predictions for half of the lists, in order to ascertain that the prediction had no qualitative effect on the nature of the items recalled, either at immediate or final test. In fact, the prediction making had no deleterious effects whatsoever on either test. They were asked to give their estimate of the probability of being able to recall that word again in approximately 10 minutes.

As noted earlier, subjects typically recall recency items first when tested immediately. Because these items are exactly those ones that will suffer maximally at later test, items which are produced in the latter portion of each individual immediate recall protocol tend to be recalled again with higher probability than those items recalled in the first half. Furthermore, because those items produced with ease on the initial test "gain" less in terms of later retrievability from this initial recall, we expect the relationship between immediate-recall output position and later-recall probability to be magnified.

This relationship does indeed hold, and is presented in the top half of Fig. 14.3. Items which are produced later on each immediate recall test are re-recalled with higher frequency than those items which are produced earlier.

Correct prediction must thus entail the assignment of lower estimates of final free recall to items "first-out" (high retrieval fluency) and higher

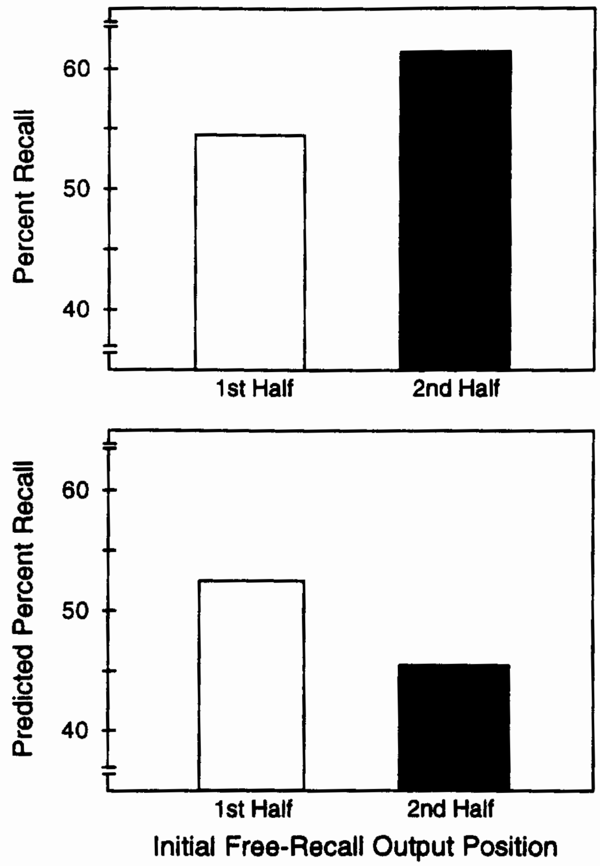

FIG. 14.3. Top panel: Probability of final recall as a function of initial recall output half. Bottom panel: Predictions of free recall as a function of initial recall output half. Adapted from Benjamin et al. (1996). Reproduced with permission.

estimates for items that are produced later on the initial-recall test (low retrieval fluency). As hinted at, however, subjects' predictions follow the opposite pattern. This relationship is depicted in the bottom half of Fig. 14.3.

Clearly, these mispredictions reflect the tacit use of the assumption dictating intact ordinal relations of retrieval fluency over time (Assumption 2, cited earlier). Because these relations do vary and predictors have no access to the nature of such change, we see predictions severely misled.

Furthermore, we have some evidence that subjects are unable to recognize when Assumption 3 may be validly applied. With respect to this experiment, the laws of retrieval practice delineated earlier dictate that the items that are more difficult to retrieve "gain" more in terms of later recallability. By this argument, the relative increase in final recall capacity as a function of the initial retrieval should increase with the serial position of an item in the output protocol. Recency items, which are typically output first, should thus be attributed lesser beneficial effects of retrieval practice than those items that come out later. Again, there is no evidence that subjects appreciate such a relationship—predictions tend to decrease with output position, not increase. For a thorough treatment of the failure to recognize effects of retrieval practice in the judgment-of-learning paradigm, see Spellman and Bjork (1992).

Distinctions Between Episodic and Semantic Memory

In another experiment, Benjamin et al. (1996, Experiment 1) examined whether subjects could appreciate the relationship between initial ease of access to the answer to a general-information question and the probability of free recalling that answer again later. Gardiner, Craik, and Bleasdale (1973) demonstrated that the probability of free recall varied positively with the latency to answer the general-information question. In the Benjamin et al. task, subjects answered a series of general-information questions that were designed to be answerable, but varied somewhat in their difficulty. After providing an answer, they made a prediction as to the future free recallability of that answer. It was emphasized to the subjects that the future task would be different from the present one, in that no recall cues would be provided. After answering and predicting for 20 such questions, and engaging in 10 minutes of distractor activity, subjects were given a blank sheet of paper and asked to recall as many of the provided answers as possible.

As noted previously, earlier work has shown that the probability of free recalling any answer is positively related to the original latency taken to provide that answer to a question. This relationship is consistent with the following perspective on the task: The longer it takes to answer the question, the more a distinctive retrieval episode is created, thus maximizing the probability of remembering that episode during the free-recall task. In other words, questions that are answered very quickly and easily do not leave much to be remembered at time of free recall. Benjamin et al. replicated this relationship, which is presented in the top panel of Fig. 14.4.

Such a conceptualization rests on a popular distinction between *semantic* and *episodic* memory (Tulving, 1983). Semantic memory contains abstract factual and relational information stored in an associative manner. Such an architecture makes the conjecture plausible that answers to questions that are provided slowly or with some difficulty are in some sense not well known.

Episodic memory, on the other hand, stores events and autobiographical experiences from a first-person perspective and does not represent information abstractly. In fact, the free-recall task described earlier is a paradigmatic example of an episodically based task—target memories are not cued by or associated to probes presented at time of retrieval; rather, retrieval is of the episode in which the to-be-recalled information was learned. From this perspective, again, the words which took longer to

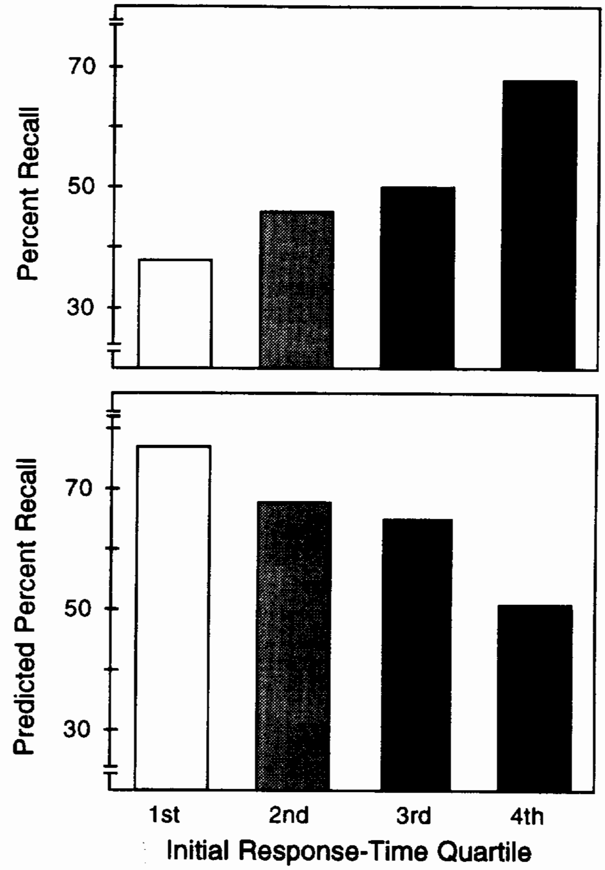

FIG. 14.4. Top panel: Probability of free recall as a function of initial response time quartile. Bottom panel: Predictions of free recall as a function of initial response time quartile. Adapted from Benjamin et al. (1996). Reproduced with permission.

produce provide more of an episode to be recalled. It is for this reason that original response latency in the Gardiner et al. (1973) task is positively related to later probability of free recall, despite the superficial counter-intuitive nature of such a relationship.

If, however, subjects fail to recognize such a dissociation of memory types (semantic vs. episodic), and instead rely on retrieval fluency as a predictor of later performance, their estimates of future recallability will again be misled. Specifically, those questions that take a long time to respond to should evoke low predictions of later recall of the answer, predicated on a phenomenological sense of "not knowing the answer well." Conversely, those answers that subjects feel they know well (i.e., answer quickly) should be assigned high estimates of later free recallability.

The bottom panel of Fig. 14.4 shows that subjects do in fact mistakenly rely on initial retrieval fluency as an indicant of later recallability, despite the fact initial retrieval fluency in the Gardiner et al. (1973) paradigm predicted *poorer* later recall. Subjects predict greater recallability for those items that are initially produced quickly and less recallability for those produced slowly.

This circumstance violates another assumption tacit within the global use of retrieval fluency as a predictive index. In particular, Assumption 1 dictating the homogeneity of tasks and their sources of retrieval fluency is violated. That is, the Gardiner et al. (1973) paradigm provides a case where retrieval fluency on different tasks clearly draws on two separate sources in memory. Because subjects fail to recognize this episodic–semantic distinction, they are subscribing to the assumption that different tasks derive retrieval fluency from the same source. This is the fallacy alluded to in Assumption 1.

MISCONCEPTIONS ABOUT OURSELVES AS LEARNERS AND REMEMBERERS

Several ways in which naive notions of memory appear to be misguided are apparent within the discussion presented here. For example, we discussed early in this analysis the fact that subjects demonstrate a failure to understand the transitory nature of priming effects, and instead attribute them as enhanced confidence (Kelley & Lindsay, 1993) or even enhanced knowing (Lee et al., 1993, cited in Narens et al., 1994).

More importantly, however, we have provided a set of three assumptions inherent within the indiscriminant use of retrieval fluency as a metacognitive index. Although these assumptions may, in practice, be met, they most definitely represent an oversimplified and incorrect view of the nature of human memory. By demonstrating that these assumptions guide predictive behavior—even when such guidance misleads metacognition in a direction opposite to true performance—we hope to have convinced the reader of several fundamental failures of the implicit models of memory held by humans:

1. The failure to appreciate the multidimensionality of memory.
2. The failure to fully appreciate the role of time and context in the preservation of memory and the manner in which they interact with the multiple dimensions of memory.
3. The failure to recognize and understand either the existence or qualitative nature of retrieval practice.

These fallacies follow directly from the assumptions listed earlier—assumptions, the truth of which in any given situation, either apparently cannot be or simply is not assessed.

Subjects demonstrate an acute unawareness of the marked difference between memory tasks, performance on which derives from semantic memory and those in which it derives from episodic memory (Benjamin et al., 1996, Experiment 1). This has been referred to as the "fallacy of homogeneous memory" and manifests itself in paradigms in which the predictive task and the to-be-predicted behavior stem from different sources in memory.

There is also a failure to understand the effects of time on memory. Specifically, Benjamin et al. (1996, Experiment 2) have shown that the differential decay processes underlying recency and primacy recall are not incorporated into metamnemonic evaluation. This inability is also reflected in the examples alluded to earlier (e.g., Narens et al., 1994) in which predictions subject to temporary priming effects do not reflect the transitory nature of such effects.

Finally, Benjamin et al. (1996, Experiment 2) and Spellman and Bjork (1992) show that subjects do not duly utilize information concerning the nature of retrieval practice when making metamnemonic predictions.

CONCLUDING COMMENT

The appearance of metacognition as a viable topic of scientific inquiry reflects a return to a central issue in psychology. Early "experimental" psychologies sought to illuminate the nature of immediate experience via trained introspection. It was of particular importance that the subjects be well trained—the idea being that only after extensive practice could a phenomenological report serve as an accurate metric for the underlying mental processes. Thus, the subject matter of psychology concerned the *objective* nature of smelling a rose or imagining a wild beast with five limbs.

In the modern era, the approach is somewhat different. A general rejection of a dualist conception of mind means that the verbal report no longer holds an untarnished image. In fact, the use of verbal reports in experimental psychology fell into such disfavor as to evoke claims of uselessness—that such reports bear essentially no relation to cognitive processes (Nisbett & Wilson, 1977).

Indeed, the resurgence of introspection within psychology is marked by a fundamental difference from its usage in the era of Wilhelm Wundt. The verbal report, like signal detection ability, sleeping behavior, or bar pressing in the rat, is viewed as a behavioral product to be explained. It is not seen as a vehicle with explanatory power for greater deeper mysteries of the human psyche—its predictive capacity, as well as its fallibility and unreliability, are all aspects to be explained much in the way that any cognitive process might be.

Toward that end, it is crucial to identify systematic biases people have in the production of introspections (see, e.g., Ericsson & Simon, 1984).

The goal of the current chapter has been to outline one such bias. It has been argued that such an undertaking has not only theoretical, but also practical value.

It might be argued that the education of metamemory is, by definition, more important than the improvement of memorial processes per se. Memory has evolved to be a highly adaptive, but fallible organ of human cognition (cf. Bjork, 1989). Metamemory, which has been considered as a sort of "system manager" for the incredible complexity of memory, serves in one sense the paramount role in cognition—as a determiner of when and how to use memory. Such a metaphor, although smacking of the problem of homunculi-driven cognition and the infinite regress problem, does emphasize an important point: Assessment of how and what we know drives what we believe we can and cannot do and, furthermore, what we do and do not continue to try to learn. Perhaps, then, one underrepresented key to enhancing human performance and learning is to redirect our focus from how to improve the system to how to use the system we have with maximal efficiency.

REFERENCES

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

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1063–1087.

Baddeley, A. (1990). *Human memory: Theory and practice.* Boston, MA: Allyn and Bacon.

Benjamin, A. S., Bjork, R. A., & Schwarz, B. L. (1996). *The mismeasure of memory: When retrieval fluency is misleading.* Manuscript submitted for publication.

Bjork, R. A. (1970, September). *Control processes and serial position effects in free recall.* Symposium on Memory, Mathematical Psychology Meetings, Miami Beach, FL.

Bjork, R. A. (1975). Retrieval as a memory modifier. In R. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 123–144). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honor of Endel Tulving* (pp. 309–330). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (Vol. 2, pp. 35–67). Hillsdale, NJ: Lawrence Erlbaum Associates.

Blake, M. (1973). Prediction of recall when recognition fails: Exploring the feeling-of-knowing phenomenon. *Journal of Verbal Learning and Verbal Behavior, 12*, 311–319.

Brown, J. (1968). Reciprocal facilitation and impairment in free recall. *Psychonomic Science, 10*, 41–42.

Christina, R. W., & Bjork, R. A. (1991). Optimizing long-term retention and transfer. In D. Druckman & R. A. Bjork (Eds.), *In the mind's eye: Enhancing human performance* (pp. 23–56). Washington, DC: National Academy Press.

Costermans, J., Lories, G., & Ansay, C. (1992). Confidence level and feeling of knowing in question answering: The weight of inferential processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 142–150.

Craik, F. I. M. (1970). The fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behavior, 9,* 143–148.

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition, 20,* 374–380.

Dunlosky, J., & Nelson, T. O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory and Language, 33,* 545–565.

Ericsson, K. A., & Simon, H. A. (1984). *Verbal reports as data.* Cambridge, MA: MIT Press.

Gardiner, J. M., Craik, F. I. M., & Bleasdale, F. A. (1973). Retrieval difficulty and subsequent recall. *Memory & Cognition, 1,* 213–216.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 55,* 351–360.

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology, 56,* 208–216.

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 experience: Noise judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 240–247.

Jacoby, L. L., Bjork, R. A., & Kelley, C. M. (1994). Illusions of comprehension, competence, and remembering. In D. Druckman & R. A. Bjork (Eds.), *Learning, remembering, believing: Enhancing human performance* (pp. 57–81). Washington, DC: National Academy Press.

Jacoby, L. L., & Kelley, C. M. (1987). Unconscious influences of memory for a prior event. *Personality and Social Psychology Bulletin, 13,* 314–336.

Kamin, L. J. (1969). Predictability, surprise, attention and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279–296). New York: Appleton-Century-Crofts.

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language, 32,* 1–24.

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100,* 609–639.

Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General, 124,* 311–333.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 107–118.

Krinsky, R., & Nelson, T. O. (1985). The feeling of knowing for different types of retrieval failure. *Acta Psychologica, 58,* 141–158.

Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 625–632). London: Academic Press.

Lee, V. A., Narens, L., & Nelson, T. O. (1993). *Subthreshold priming and the judgment of learning.* Manuscript submitted for publication.

Lovelace, E. A. (1984). Metamemory: Monitoring future recallability during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 756–766.

Lupker, S. J., Harbluk, J. L., & Patrick, A. S. (1991). Memory for things forgotten. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 897–907.

Marks, G., & Miller, N. (1985). The effect of certainty on consensus judgments. *Personality and Social Psychology Bulletin, 11,* 165–177.

Mazzoni, G., & Nelson, T. O. (1995). Judgments of learning are affected by the kind of encoding in ways that cannot be attributed to level of recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1263–1274.

McGeogh, J. A. (1932). Forgetting and the law of disuse. *Psychological Review, 39*, 352–370.

Narens, L., Jameson, K. A., & Lee, V. A. (1994). Subthreshold priming and memory monitoring. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 71–92). Cambridge, MA: MIT Press.

Nelson, T. O., & Dunlosky, J. (1992). How shall we explain the delayed judgment-of-learning effect? *Psychological Science, 3*, 317–318.

Nelson, T. O., Gerler, D., & Narens, L. (1984). Accuracy of feeling-of-knowing judgments for predicting perceptual identification and relearning. *Journal of Experimental Psychology: General, 113*, 282–300.

Nelson, T. O., Leonesio, R. J., Shimamura, A. P., Landwehr, R. F., & Narens, L. (1982). Overlearning and the feeling of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 279–288.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231–259.

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

Postman, L., & Phillips, L. W. (1965). Short term temporal changes in free recall. *Quarterly Journal of Experimental Psychology, 17*, 132–138.

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology, 19*, 111–138.

Reder, L. M., & Klatzky, R. L. (1994). Transfer: Training for performance. In D. Druckman & R. A. Bjork (Eds.), *Learning, remembering, believing: Enhancing human performance* (pp. 25–56). Washington, DC: National Academy Press.

Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 435–451.

Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology, 39*, 475–543.

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.

Schacter, D. L., & Worling, J. R. (1985). Attribute information and the feeling-of-knowing. *Canadian Journal of Psychology, 39*, 467–475.

Shaw, J. (1996). Increases in eyewitness confidence resulting from postevent questioning. *Journal of Experimental Psychology: Applied, 2*, 126–146.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 592–604.

Spellman, B. A., & Bjork, R. A. (1992). When predictions create reality: Judgments of learning may alter what they are intended to assess. *Psychological Science, 3*, 315–316.

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

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352–373.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 124–131.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*, 293–315.

Watkins, O. C., & Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory, 104*, 442–452.

Witherspoon, D., & Allan, L. G. (1985). The effects of a prior presentation on temporal judgments in a perceptual identification task. *Memory & Cognition, 13*, 103–111.