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

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

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

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 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 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

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 in-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

usible but
th respect
vided the
imed—*re-*
asing the
ce in that
ibsequent
sed words
ons to be
attern of
speed of

fidence 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

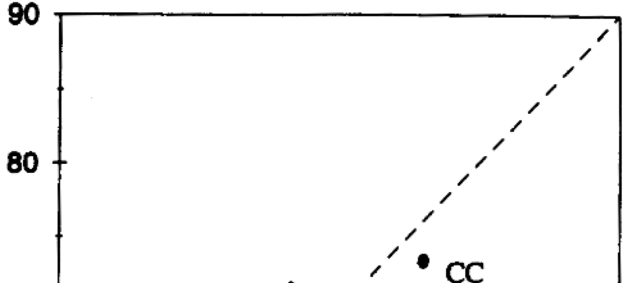
In their seminal analysis of probability estimation, Tversky and Kahneman (1974) documented a number of ways in which such assessments may go awry. One oft-cited bias lies in the use of the *availability heuristic.* This heuristic involves using the ease with which an instance of an event comes to mind as a basis for estimating the probability of that event. Consider a 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

:ions. For
either an
r they felt
er among
own later

wo major
Gerler, &
ving some
valuation)

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



Jame
hanc
studi
parti
prese
of be
men

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 given situation, however, reflects the tacit belief that certain assumptions have been met. Those assumptions include:

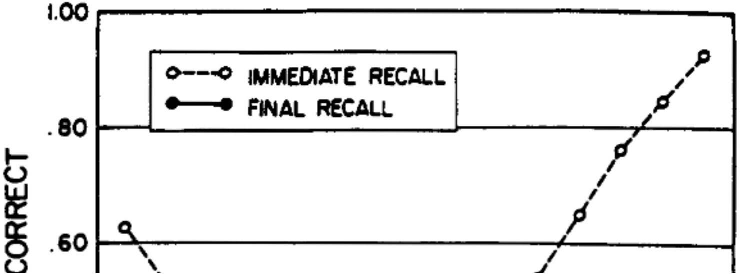
1. That the retrieval cues at the to-be-predicted time, and the retrieval task itself, will not differ substantially from the current cues and task, OR that any such differences will not measurably affect performance.
2. That events and the passage of time between now and the later task

From a r
to retrieve l
a current l
McGeogh (
one factor
factor respc
ditions at t

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

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 tion 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!



bottom cu
a function
of the exp
of the pric
still recalle
lists are nc
subjects wh
their own

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.

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

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.

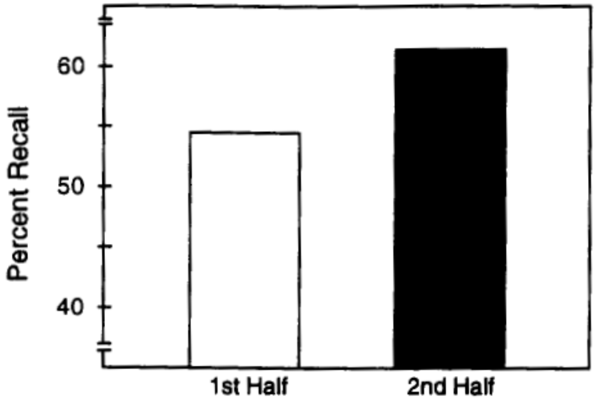
The Recency-to-Primacy Shift

Earlier in this chapter we discussed the manner in which end-of-list (resubjects, 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

demon-
(1970)
ual sub-
formed

etrieval
hem to
nember
eading)
possible

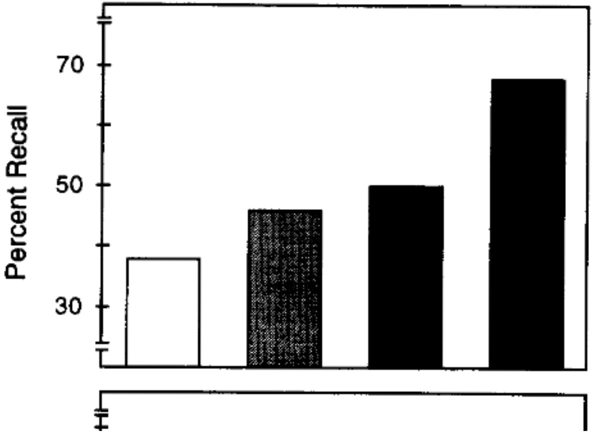


than those items that come out later. Again, there is no evidence that subjects appreciate such a relationship—predictions tend to decrease with output position, not increase. For a thorough treatment of the failure to recognize effects of retrieval practice in the judgment-of-learning paradigm, see Spellman and Bjork (1992).

Distinctions Between Episodic and Semantic Memory

lence that
rease with
failure to
ıing para-

examined
ial ease of
probability
Bleasdale



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

Subjec
between
memory
al., 1996
homoger
predictiv
in memo
There

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

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

Costerman
questio
Learnin
Craik, F.
Learnin
Dunlosky,
(JOL)
Dunlosky,
to the
and L

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*

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*

Herber
Carnegie

I am g
"closir
all of t
and p
exper
implic

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.