

Response speeding mediates the contributions of cue familiarity and target retrievability to metamnemonic judgments

AARON S. BENJAMIN

University of Illinois at Urbana-Champaign, Urbana, Illinois

Metamnemonic judgments are influenced by the retrievability of the target memory in question, but also by the familiarity of the cue used to elicit such judgments. However, there have been few suggestions as to what factors mediate the influence of these different sources of information on metamnemonic judgments. In this experiment, I examined the interactions between prediction time pressure and variables that promote either cue familiarity or target retrievability. The data reveal that target retrievability plays a larger role than does cue familiarity in fostering predictions of future recall made under unpressured conditions, but that cue familiarity influences predictions that are speeded. This pattern is interpreted by analogy with recognition memory: Mnemonic evidence based on familiarity is more impervious to the demands of time pressure than are the products of deliberative retrieval. Several explanations for this effect are suggested.

The effective use of one's memory depends on many factors above and beyond its fidelity. We constantly make decisions about what aspects of a situation or an object to commit to memory, how to do so, and how successful we think we were in doing so. Metamemory research focuses on such monitoring and control processes that interact with memory (e.g., Nelson & Narens, 1990) and holds promise not only to help import memory research into applied and educational settings, but also to provide novel approaches to the understanding of memory phenomena. This article revisits a debate on the nature of the information that is incorporated into metamnemonic judgments and, in particular, on whether *cue familiarity* or *target retrievability* influences metamemory judgments. This debate faded from mainstream metamemory research, in large part, because the evidence marshaled in support of each of these views was sufficiently convincing to persuade most researchers in the field that both factors played a role. This evidence will be briefly reviewed below, but it is to an extension of this issue that I wish to turn: What mediates the influence of these factors across tasks, individuals, or stimuli?

Several pieces of evidence suggest that judgments about the future retrievability of a target word can be influenced by the familiarity of the cue term. Schwartz and Metcalfe (1992) showed that subjects predicted superior recognition for unrecalled target words (*feelings of knowing*) paired with cues that had been previously exposed, even though actual rates of recognition were not affected

by that prior episode. Cue familiarity also plays a role in tasks in which subjects predict future recall. In a semantic retrieval paradigm, Reder (1987) showed that speeded predictions of answer retrievability to general information questions increased when the words from the question had been primed in an earlier task. Begg, Duft, LaLonde, Melnick, and Sanvito (1989) showed that cue concreteness increased the magnitude of predicted future recall (*judgments of learning*) as well.

Other evidence has suggested that target retrievability serves as a basis for metacognitive judgments. Blake (1973), for example, showed that predictions of recognition for unrecalled trigrams increased with the amount of partial recall, as did actual recognition probability. Koriat (1993) showed that predictions of recognition for unrecalled answers to general information questions depended on the amount of information that came to mind in response to those questions. Thus, he suggested that general accessibility of information, and not target retrievability specifically, fosters certain metacognitive judgments. Schacter and Worling (1985) showed that people can retrieve affective characteristics of words for which they cannot retrieve the appropriate word form and that metamnemonic judgments about the impending recognition of those words reflected access to such characteristics.

As has been noted by Koriat and Levy-Sadot (2001), these two views have often been held in opposition to one another, tacitly implying that one or the other more accurately reflected the "true" nature of metacognitive judgments. Koriat and Levy-Sadot focused on feeling-of-knowing judgments and showed that both familiarity and retrievability influenced those judgments. They further suggested, but did not directly test, the presence of an asynchrony in the accrual of evidence for a metacognitive

Correspondence concerning this article should be addressed to A. S. Benjamin, Department of Psychology, University of Illinois, 603 E. Daniel St., Champaign, IL 61820 (e-mail: asbenjam@psych.uiuc.edu).

judgment, such that evidence for cue familiarity arrives earlier than evidence from retrieval. This proposal parallels one made in the recognition memory literature (e.g., Hintzman & Curran, 1994) and suggests a deeper analogy between metacognitive and recognition judgments (see also Benjamin, Bjork, & Hirshman, 1998).

In the literature on recognition memory, a number of authors have suggested that judgments derive from a combination of match-based familiarity and retrieved specific evidence (e.g., Kelley & Jacoby, 2000). These two components map semantically, as well as nominally, onto the two proposed sources of evidence in metacognition. This mapping is in some ways self-apparent: In both cases, the task for the subject is to make a judgment about the likelihood of an event (either past or future) inferentially, using only presently available cues. The cues either can be matched to memory, to yield familiarity, or can be used as a retrieval probe (Humphreys, Wiles, & Dennis, 1994); in either case, the accuracy of the judgment is determined by the quality of the match and of the retrieved information, as well as the degree to which familiarity or retrieval is diagnostic of the event in question.

In recognition, speeded responding limits the degree to which subjects can retrieve contextual information associated with the test word, thus increasing false alarm rates to studied but contraindicated list members (Benjamin & Craik, 2001; Jacoby, 1999) and to high semantic associates of studied stimuli (Benjamin, 2001). More important, speeded responding changes the relationship between learning variables and false alarms to such stimuli: Whereas under unspeeded conditions increased learning led to a lower false alarm rate, it led to a higher false alarm rate under speeded conditions. This interaction suggests that speeded responding relies more predominately on familiarity than on retrieved context evidence.

In either case, retrieved products provide more appropriate evidence than does familiarity. The ease and quality of retrieval now likely predict effective retrieval in the future (Benjamin & Bjork, 1996; Benjamin, Bjork, & Schwartz, 1998), but the familiarity of the retrieval cue is only indirectly related to the likelihood of successful retrieval. Likewise, the successful retrieval of an encoding context allows the subject to accurately classify an item in a recognition task, even one that should be rejected despite its high familiarity. However, the fact that such retrieval unfolds over a longer timescale than does the assessment of familiarity (Hintzman & Curran, 1994) indicates that manipulations of time pressure on either metacognitive or recognition tasks should reveal a trade-off between timely responding and the type of evidence used in the decision.

The goal of the present experiment was to evaluate how predictions of memory performance vary with a prediction response deadline. If, in accord with the analogy drawn above between recognition and judgments of learning and the suggestion made by Koriat and Levy-Sadot (2001) and others (e.g., Nhouyvanisvong & Reder, 1998), imposing time pressure on the prediction alters the basis for the judgment, we should be able to elicit effects of cue

familiarity under speeded prediction conditions and of target retrievability under unspeeded prediction conditions.

In this experiment, exposure durations to the cue and target terms of paired associates were manipulated and were assumed to roughly affect cue familiarity and target retrievability, respectively. On each study trial, the subjects viewed a cue term that remained on the screen for a variable duration. After that interval, the target term appeared, and the pair remained on the screen for another variable duration. The cue term did not disappear when the target term appeared, because the goal of the task was to encourage integrative processing of the two terms. This fact does compromise our manipulation of cue familiarity somewhat, but given that the subjects' goal was to learn to recall the second word when presented with the first, it is unlikely that they continued to process the cue term exclusively during the second interval.

After the learning phase, the subjects made both speeded and unspeeded predictions of their ability to recall a target word when presented with the cue. These predictions were made with only the cue present, since the provision of the target term itself would disrupt the subjects' opportunities to engage in a diagnostic retrieval of that term (see Dunlosky & Nelson, 1992; Jacoby & Kelley, 1987). I chose to use the traditional judgment-of-learning task, rather than a feeling-of-knowing task, because of difficulties associated with differential item selection across different conditions in the latter. Methodologically, the speeded condition in this experiment was similar to the task used by Reder (1987), in which subjects provided rapid estimates of their ability to immediately recall an answer to a question. The critical differences between her task and this one are that the criterion test did not occur immediately and, also, that the nature of the relationship between stimuli was episodic rather than semantic.

Finally, in the recall phase, the subjects were provided the cue terms and were asked to recall the targets. Because of the difficulty of the task, the subjects were provided with an increasing number of letters from the target word until they could recall it on their own. The dependent measure of proportion of letters necessary for successful recall has certain advantages over a traditional measure of cued recall. It increases the reliability and decreases the variability of individual subjects' performance, because unlike standard recall, it is not based on dichotomous events. In addition, it allows the experimenter to increase the number of repeated measures without introducing floor-related variability truncation. However, it does preclude the possibility of looking at the differential effects of errors of commission and omission (see Koriat & Levy-Sadot, 2001).

METHOD

Subjects

Thirty-six undergraduate students from the University of Illinois participated in partial fulfillment of course requirements.

Design

Recall accuracy was measured in a 2×2 design (cue duration \times target duration), and mean predictions were measured in a $2 \times 2 \times$

2 (cue duration \times target duration \times prediction speed) design. All variables were manipulated within subjects.

Materials

The words were nouns drawn from the MRC Psycholinguistic Database (http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm) and ranged in length between four and eight letters. All of the following randomization and counterbalancing procedures were implemented for each subject.

Eighty word pairs were created by randomly combining words within the set. These pairs were divided into four sets of 20, and each set was assigned to a unique cue-duration/target-duration combination. The pairs were quasirandomly intermixed to create the study list and were rerandomized within four blocks of prediction trials, alternating between speeded and unspeeded conditions. Whether the first block was speeded or unspeeded was counterbalanced across subjects. Half of the items from each of the four study duration conditions were assigned to the speeded prediction condition, and half to the unspeeded prediction condition. The word pairs were randomized one more time to yield the test list.

Procedure

The subjects were instructed to learn each pair of words such that they would be able to recall the second when prompted with the first. The word pairs were then presented individually. The cue duration was either 1 or 5 sec, and the target remained on the screen for another 1 or 5 sec. There was a 1-sec interstimulus interval between pairs. At the end of the study phase, the subjects took a short break (~5 sec) before receiving the instructions for the prediction phase. Those instructions informed them that they would be shown a cue term and would be asked to predict the probability of recalling the corresponding target word in about 10 min. They were also told about the different response speed conditions and how they should make their responses as rapidly as possible in the two speeded blocks of trials. The response scale ranged from 1 (*I am sure I will NOT remember this word*) to 4 (*I am sure I WILL remember this word*). In the speeded condition, the subjects were admonished to "Please SPEED up!" if the response time on a trial exceeded 1 sec. All the data were recorded and used in the analysis, regardless of whether they exceeded the deadline.

After the prediction phase and another ~5-sec break, the recall test was administered. On that test, the cue word was presented along with an empty box in which the subjects were instructed to type the target word. If their response was incorrect or they did not enter a response before pressing the "Enter" key, the first letter of the target word was presented as a clue. The subjects then attempted to report the answer again, and an incorrect or missing response elicited the next letter in the word, and so on, until the subjects provided the correct answer or the entire word was presented. Because the assignment of words was random, any differences in correct guessing rates across items were equally distributed across conditions. After the test, the subjects were debriefed and given credit for their time.

RESULTS

Figure 1 depicts the mean accuracy scores across study conditions and reveals that target duration [$t(35) = 4.50$], but not cue duration [$t(35) = 0.87$], reliably affected cued recall accuracy. The overall three-way interaction of cue duration, target duration, and test type on predictions (shown in Figure 2) was not reliable [$F(1,35) = 0.821$], nor was the two-way interaction of test and target duration [$F(1,35) = 0.23$], but the interaction of test and cue duration was [$F(1,35) = 3.99$]. An analysis of the unspeeded test condition revealed a reliable main effect of target duration [$F(1,35) = 7.05$], but not of cue duration

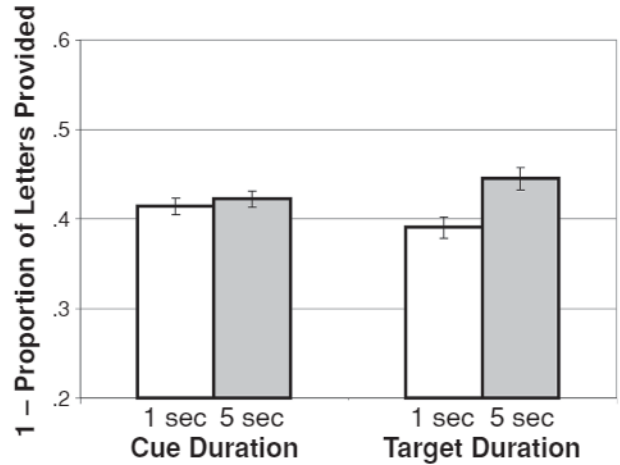


Figure 1. Mean cued recall accuracy as a function of cue and target duration. Error bars represent the standard error of the difference scores for the two variables (~.02).

[$F(1,35) = 0.52$], whereas the speeded test elicited an effect of cue duration [$F(1,35) = 5.15$], but not of target duration [$F(1,35) = 3.49$].¹ The prediction results were additionally analyzed with a contrast constructed to the form of the predicted three-way interaction, in which an effect of target duration, but not of cue duration, was expected under unspeeded conditions but an effect of cue duration, but not of target duration, was predicted under speeded conditions. This contrast was evaluated on the residualized data after removing the (uninteresting) main effect of test speed and revealed a reliable effect (mean $\psi = 1.04$, $SEM = 0.33$).² The evidence that unspeeded prediction conditions foster more accurate judgments by allowing for retrieval is further supported by the reliably higher correlation between judgments and performance in that condition ($r = .40$), as compared with the speeded condition [$r = .23$; $t(34) = 3.99$].³ Correlations between predictions and performance for the individual conditions are shown in Figure 3.⁴

DISCUSSION

This experiment revealed that imposing a deadline on the prediction changed the relationship between study variables and that prediction. Under unspeeded conditions, predictions were influenced by target duration, but not by cue duration, thus suggesting the greater influence of retrievability than of familiarity. However, under speeded conditions, a different result obtained: Predictions quite dramatically reflected cue familiarity, but target retrievability considerably less so. These effects are consistent with the claim that response deadlines limit the degree to which retrieved information is used in inferential metacognitive tasks and are analogous to results in the literature on recognition memory that implicate an increased role for familiarity under speeded recognition conditions. In addition, they reveal a condition under which the ef-

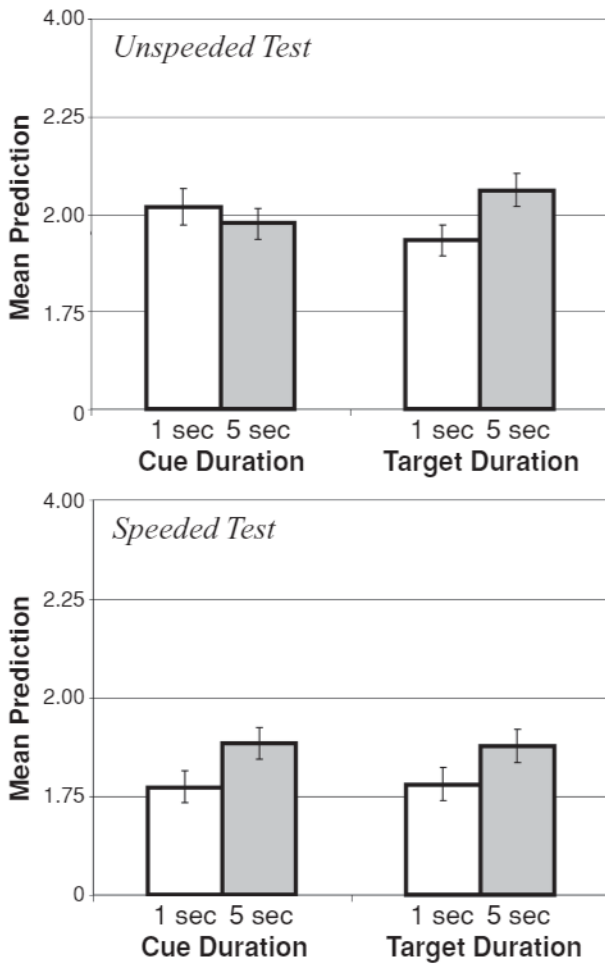


Figure 2. Mean prediction as a function of cue duration, target duration, and response deadline condition. Error bars represent the standard error of the difference scores for each pair of scores (~ 0.04).

fects of cue familiarity on judgments about impending semantic knowledge (Connor, Balota, & Neely, 1992; Reder, 1987) or answer retrievability (as opposed to computability) in mathematical problems (Reder & Ritter, 1992) are likely to be evident. The fact that only relatively unpressured conditions allow other retrieved sources of evidence to influence metacognitive decisions also helps explain why the influence of target retrievability is not always apparent.

So what is the mechanism whereby speeding recognition and metacognitive assessments lead to responding biased toward familiarity and away from retrieval? One possibility is simply that the computations underlying the two processes necessarily imply different time courses. Models that characterize familiarity assessment as a match between a probe and all of memory (for a review, see Humphreys, Pike, Bain, & Tehan, 1989) provide a somewhat natural account for such a distinction: Familiarity-based evidence is computed in parallel across

all traces in memory, whereas retrieved evidence involves pattern completion at the level of individual traces. By this argument, deadlines directly decrease the contribution of retrieved evidence by making it probabilistically less likely that the information will be available at the time that the judgment needs to be made. In addition, deadlines may affect decisions indirectly by inducing subjects to change their strategies on the basis of their assessments of both useful and timely available information. That is, if subjects appreciate the differential time courses whereby familiarity and retrieved evidence become available, they may (wisely) adopt a strategy that relies more heavily on familiarity, which is more rapidly accessible.

Another account for the effect of deadlines is more consistent with what have been termed *single-process* accounts of recognition. Such theories presume that when a cue is submitted to memory for a recognition judgment, it consists not only of the probe word, but also of elements of the sought-after context. It thus elicits a higher match to memory if both the context and the word match than if only the word does. The judicious placement of a decision criterion will thus allow subjects to discriminate between

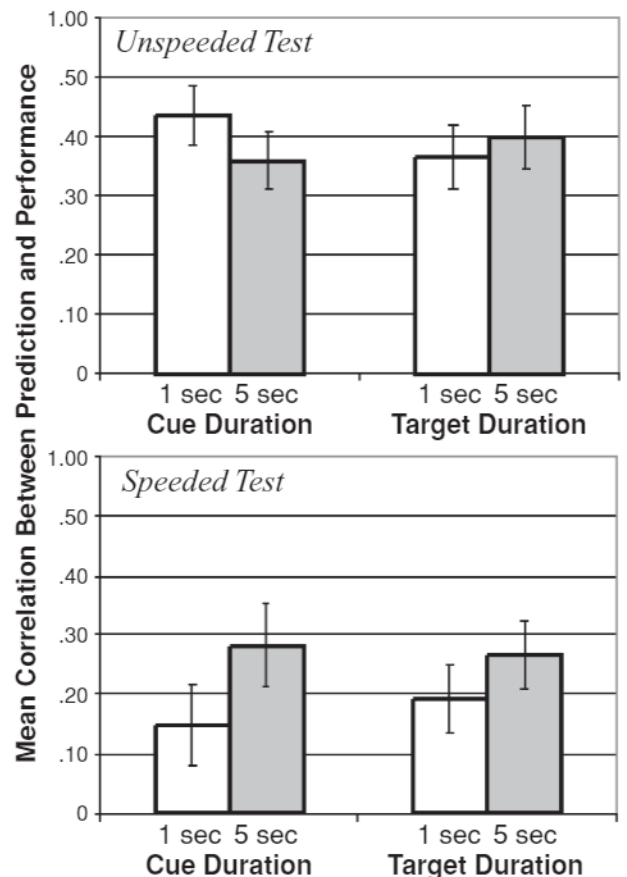


Figure 3. Mean correlations (r) between predictions and performance for the individual study/test conditions. Error bars represent the standard error of the difference scores for each pair of scores (~ 0.05).

to-be-endorsed and contraindicated stimuli. Because this complex cue is multidimensional—the subjects need to evaluate what characteristics best discriminate between the multiple contexts in question—it may take time and effort to construct it. Thus, deadlines may force the subjects to submit to memory a cue that is different from the one that they would have chosen had they been given more time.

To apply such an explanation to the present task, consider the fact that words are part of a massive network of information and, thus, share associations, semantic or otherwise, with a tremendous number of other words. In the context of the study phase of the experiment, of course, they are associated only with one other word, but discriminating this single episodic pairing from a lifetime of semantic knowledge is not a trivial task. Thus, the cue submitted to memory may consist of more than the cue word itself; it may contain information about the experimental or study context in addition. If assembling this complex cue takes an appreciable amount of time, relative to the total available time, the subject may forgo this strategy and simply submit the cue word, which is easily available on the computer screen. This strategy is consistent with the finding that deadlines in recognition change the relationship between previous learning and false alarm rates to semantically related stimuli (Benjamin, 2001), but there are no data that bear on whether it applies well to metacognitive tasks.

The present data do not discriminate between these two theoretical positions, but they do illustrate a deep analogy between metacognition and recognition. The sources of data for both tasks include two potential sources of mnemonic information: the degree to which the cue term matches memory (familiarity) and the specific information retrieved in response to that cue (target retrievability). Just as familiarity can mislead recognition (e.g., Benjamin, 2001; Benjamin & Craik, 2001; Jacoby, 1999), it also can mislead metacognition (Reder, 1987; Reder & Ritter, 1992; Schwartz & Metcalfe, 1992). Accordingly, both judgments are superior in accuracy when retrieved information can be brought to bear on the decision. However, something about the nature of speeded responding affects the bases for these decisions and biases both toward familiarity.

I have tacitly implied here that acts of metacognition are fundamentally acts of memory. Of course, this is not entirely true; many factors are brought to bear in metacognitive evaluation, of which only a subset are mnemonic. But even memory judgments are not entirely mnemonic in origin; for example, recognition judgments reflect situational (e.g., base rates) and idiosyncratic (e.g., response criteria) factors, in addition to mnemonic ones.

I do claim, however, that there is no reason to search outside the domain of theoretically well-established mnemonic processes for mnemonic sources of information in metamemory. Although early theories of the feeling of knowing proposed *direct access* to parameters about

even currently inaccessible memory traces (Hart, 1965), such theories failed, in large part, because they presupposed mnemonic capabilities that do not exist. However, descendant theories that postulate partial access to retrieved information without any privileged knowledge of its correctness or relevance—and are thus concordant with the fact that memory retrieval is rarely an all-or-none process—are quite successful in explaining and predicting novel phenomena in metamemory (Koriat, 1993).

Likewise, every memory task involves various metacognitive decisions. Outside the laboratory, students make decisions about their readiness for an exam and choices about how to increase their preparedness. Inside the laboratory, subjects make decisions about which stimuli to skip or give up on on a memory test (Koriat & Goldsmith, 1993), when to switch from one semantic retrieval task to another (Young, 2004), how to space multiple study repetitions of to-be-learned material (Son, 2004), or what strategies to use in order to formulate an answer to a question (Reder, 1987).

In this article, I have shown how prediction deadlines influence the type of information that is used in the generation of a metamnemonic judgment. Given sufficient time, subjects rely primarily on target retrievability. However, when time is limited, cue familiarity plays a larger role. This result is analogous to results for recognition memory and illustrates how common principles and mechanisms fundamentally underlie both memory judgments and metacognition.

REFERENCES

- BEGG, I., DUFT, S., LALONDE, P., MELNICK, R., & SANVITO, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory & Language*, **28**, 610-632.
- BENJAMIN, A. S. (2001). On the dual effects of repetition on false recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 941-947.
- BENJAMIN, A. S., & BJORK, R. A. (1996). Retrieval fluency as a metacognitive index. In L. Reder (Ed.), *Metacognition and implicit memory* (pp. 309-338). Hillsdale, NJ: Erlbaum.
- BENJAMIN, A. S., BJORK, R. A., & HIRSHMAN, E. (1998). Predicting the future and reconstructing the past: A Bayesian characterization of the utility of subjective fluency. *Acta Psychologica*, **98**, 267-290.
- BENJAMIN, A. S., BJORK, R. A., & SCHWARTZ, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, **127**, 55-68.
- BENJAMIN, A. S., & CRAIK, F. I. M. (2001). Parallel effects of aging and time pressure on memory for source: Evidence from the spacing effect. *Memory & Cognition*, **29**, 691-697.
- BLAKE, M. (1973). Prediction of recognition when recall fails: Exploring the feeling-of-knowing phenomenon. *Journal of Verbal Learning & Verbal Behavior*, **12**, 311-319.
- CONNOR, L. T., BALOTA, D. A., & NEELY, J. H. (1992). On the relation between feeling of knowing and lexical decision: Persistent subthreshold activation or topic familiarity? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 544-554.
- 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.
- HART, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, **56**, 208-216.

- HINTZMAN, D. L., & CURRAN, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory & Language*, **33**, 1-18.
- HUMPHREYS, M. S., PIKE, R., BAIN, J. D., & TEHAN, G. (1989). Global matching: A comparison of the SAM, Minerva II, Matrix, and TODAM models. *Journal of Mathematical Psychology*, **33**, 36-67.
- HUMPHREYS, M. S., WILES, J., & DENNIS, S. (1994). Toward a theory of human memory: Data structures and access processes. *Behavioral & Brain Sciences*, **17**, 655-692.
- JACOBY, L. L. (1999). Ironic effects of repetition: Measuring age-related differences in memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 3-22.
- JACOBY, L. L., & KELLEY, C. M. (1987). Unconscious influences of memory for a prior event. *Personality & Social Psychology Bulletin*, **13**, 314-336.
- KELLEY, C. M., & JACOBY, L. L. (2000). Recollection and familiarity: Process-dissociation. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 215-228). London: Oxford University Press.
- KIMBALL, D. R., & METCALFE, J. (2003). Delaying judgments of learning affects memory, not metamemory. *Memory & Cognition*, **31**, 918-929.
- 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., & GOLDSMITH, M. (1993). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, **103**, 490-517.
- KORIAT, A., & LEVY-SADOT, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 34-53.
- MORRISON, D. F. (1967). *Multivariate statistical methods*. New York: McGraw-Hill.
- NELSON, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, **95**, 109-133.
- NELSON, T. O., & NARENS, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125-141). San Diego: Academic Press.
- NHOUYVANISVONG, A., & REDER, L. M. (1998). Rapid feeling-of-knowing: A strategy selection mechanism. In V. Y. Yzerbyt, G. Lories, & B. Dardenne (Eds.), *Metacognition: Cognitive and social dimensions* (pp. 35-52). London: Sage.
- REDER, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology*, **19**, 111-138.
- 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, & Cognition*, **18**, 435-451.
- SCHACTER, D. L., & WORLING, J. R. (1985). Attribute information and the feeling-of-knowing. *Canadian Journal of Psychology*, **39**, 467-475.
- SCHWARTZ, B. L., & METCALFE, J. (1992). Cue familiarity but not target retrievability enhances feeling-of-knowing judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 1074-1083.
- SON, L. K. (2004). Spacing one's study: Evidence for a metacognitive control strategy. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 601-604.
- YOUNG, C. J. (2004). Contributions of metaknowledge to retrieval of natural categories in semantic memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 909-916.

NOTES

1. This effect was quite close to the threshold α level, at $p = .07$.
2. This technique is similar to that used in *profile analysis* (e.g., Morrison, 1967), in which the parallelism of two functions is evaluated separately from function coincidence or horizontality.
3. One degree of freedom was lost for this analysis because the correlation could not be estimated for the speeded prediction condition for 1 subject (null variance in predictions).
4. Gamma correlations, the typical measure of choice for relative accuracy in experiments of this sort, have been argued to be optimal for 2×2 designs (Nelson, 1984) and may also be inappropriately influenced by levels of criterion performance (Kimball & Metcalfe, 2003). The continuous nature of the performance variable in this experiment made the traditional correlation coefficient (r) better suited for this design.

(Manuscript received April 20, 2004;
revision accepted for publication December 27, 2004.)