*Memory & Cognition*
*2003, 31 (6), 918-929*

# Delaying judgments of learning affects memory, not metamemory

DANIEL R. KIMBALL and JANET METCALFE
*Columbia University, New York, New York*

Judgments of learning (JOLs) for cue–target word pairs correlate particularly well with later target recall when made under conditions that permit delayed attempts to retrieve the targets—the *delayed-JOL effect*. Metamemory theories claim that memory monitoring improves under these conditions. However, another theory—the memory hypothesis—claims that the correlation increases because retrieved items receive a boost in recall from spaced study and are assigned high JOLs, whereas unretrieved items receive no spaced study and, therefore, no boost in recall and, consequently, are assigned low JOLs. When we eliminated differences in spaced study by reexposing word pairs following their JOLs, the delayed-JOL effect disappeared, supporting the memory hypothesis.

In the course of learning, people often have occasion to judge the extent to which they have learned information. For example, it is widely assumed that students make such judgments of learning (JOLs) to inform their study time allocation decisions (e.g., Metcalfe, 2002; Nelson, Dunlosky, Graf, & Narens, 1994; Nelson & Narens, 1990; Son & Metcalfe, 2000). Despite the seeming importance of JOLs in such contexts, early research showed that the judgments were only moderately correlated with subsequent recall (e.g., Bower & Winchester, 1970; Vesonder & Voss, 1985).

More recently, however, Nelson and Dunlosky (1991; Dunlosky & Nelson, 1992, 1994, 1997) have shown that JOLs can be highly correlated with future memory performance under certain conditions. In these studies, participants studied word pairs and were asked to rate how likely they were to recall the second word in a pair (the target) when provided with the first word (the cue) on a delayed test. The studies showed that JOLs made at a delay following initial study were more highly correlated with final recall (as measured by gamma[1]) than were JOLs made immediately after initial study, but only when the cue alone was provided at the time of the JOL, not when both the cue and the target were provided. Nelson and Dunlosky (1991) referred to this superiority in the gamma correlation for delayed cue-only JOLs as the *delayed-JOL effect*.

Nelson and Dunlosky (1991) viewed the higher gamma correlation for delayed cue-only JOLs as an indication that participants were better at assessing the state of their memory in that condition. On the basis of postexperiment interviews, they suggested that participants covertly attempted to retrieve information about the target before making their JOLs. They argued that, because final recall of a target was based on information in long-term memory alone, access to that information without interference from information about the target in short-term memory resulted in more accurate JOL decisions. Hence, delayed cue-only JOLs were highly correlated with final recall because, at the time of the JOL, information concerning the target was no longer available in short-term memory. By contrast, in other conditions—those involving immediate cue-only, immediate cue–target, and delayed cue–target JOLs—the undiagnostic information about the target was available in short-term memory at the time of the JOL because the target was present during the JOL, or it had just been presented for initial study, or both.

In a similar vein, several researchers have suggested that delayed retrieval attempts improve judgments of relative recallability by fostering transfer-appropriate monitoring that takes advantage of the consistency between the processes used in making the JOL and those used in final recall—that is, the processes used in delayed target retrieval (Begg, Duft, Lalonde, Melnick, & Sanvito, 1989; Dunlosky & Nelson, 1992; Glenberg, Sanocki, Epstein, & Morris, 1987). More recently, Koriat (1997) has suggested that, with delay, participants change the basis on which they make JOLs, relying less on an item's intrinsic preexperimental retrievability and relying more on the participant's own internal mnemonic cues, such as ease of processing and the relative accessibility of the target. These explanations share a common assumption that the delayed-JOL effect arises due to the effects of JOL timing and JOL format on the metacognitive processes involved in making a JOL.

Spellman and Bjork (1992) have offered an explanation of the delayed-JOL effect that differs markedly from these metacognitive explanations. They noted that the method of requesting JOLs in the delayed cue-only condition was

This research was reported in a paper presented at the 43rd Annual Meeting of the Psychonomic Society (November 2002, Kansas City). The research was supported by a grant from the James S. McDonnell Foundation to J. M. Special thanks go to Robert A. Bjork, Michael J. Kahana, and Barbara Spellman for comments on earlier drafts and to Tina Li and Sripriya Radhakrishna for assistance in collecting data. Correspondence concerning this article should be addressed to D. R. Kimball, Psychology Department, University of Texas at Arlington, P. O. Box 19528, Arlington, TX 76019-0528 (e-mail: dkimball@uta.edu).

itself likely to alter the accessibility of the targets in a way that would increase the correlation between JOLs and recall. They assumed, as had Nelson and Dunlosky (1991), that participants covertly attempted target retrieval when making a delayed cue-only JOL. However, they pointed to previous research on spacing effects showing that delayed retrieval practice improves final recall of successfully retrieved items (e.g., Allen, Mahler, & Estes, 1969) and that increasing the delay prior to retrieval practice increases the improvement in final recall that results from successful retrieval practice (Whitten & Bjork, 1977; see also Kelemen & Weaver, 1997). Spellman and Bjork argued that delayed cue-only JOLs are likely to be highly correlated with final recall because retrieved items (those given high JOLs) become even more retrievable because of the spacing of the retrieval practice, whereas unretrieved items (those given low JOLs) do not receive such spaced retrieval practice and, therefore, do not become more retrievable. Accordingly, Spellman and Bjork argued that the higher correlation between delayed cue-only JOLs and recall could arise because the method of obtaining recall predictions affects recall itself, the very phenomenon that is being predicted. Because this explanation claims that the delayed-JOL effect arises from the effect of spaced retrieval practice on subsequent recall, which is a memory effect, we refer to it as the *memory hypothesis.*

In the experiments presented here, we explored the extent to which the higher correlation that is normally observed between delayed cue-only JOLs and recall is attributable to the effects of spaced retrieval practice for high-JOL items and the absence of such effects for low-JOL items, as the memory hypothesis posits. We obtained JOLs by using delayed cue–target versus delayed cue-only stimuli (Experiments 1 and 2) and immediate cue-only versus delayed cue-only stimuli (Experiment 3).

Our key manipulation involved brief reexposure of a cue–target pair shortly following the JOL for that pair. The purpose of this post-JOL reexposure was to control for the effects on final recall of differences in study opportunities that occur at the time of a delayed cue-only JOL due to the retrieval of some items, but not others, at that time. In particular, by including a condition in which each target was not reexposed until after a delayed cue-only JOL had been made, we ensured that the JOLs in that condition were not informed by such target reexposure, but also that all the items in that condition experienced a spaced study opportunity at that time. Thus, this condition differs from the typical delayed cue-only condition, in which only some of the items—those that are retrieved—receive spaced study opportunities, and it also differs from the typical delayed cue–target condition, in which target reexposure occurs before the JOL and, thus, may influence both the JOLs and subsequent retrievability.

Applying principles derived from research on spacing effects (for reviews, see Crowder, 1976; Glenberg, 1979; Greene, 1989; Hintzman, 1974, 1976), the memory hypothesis makes specific predictions regarding the effects of post-JOL reexposure on both gamma and the pattern of final recall as a function of JOL level. To begin with, the memory hypothesis assumes that post-JOL reexposure of the cue–target pair serves as an additional study opportunity, just as it assumes that the retrieval or reexposure of the target at the time of the JOL serves as an additional study opportunity. Furthermore, the memory hypothesis assumes that, as is true for the study opportunity at the time of the JOL, the effect of a post-JOL study opportunity on the subsequent retrievability of an item varies as a function of the delay since the most recent study opportunity for that item.

In particular, if a target has just been reexposed or retrieved at the time of the JOL, then according to the memory hypothesis, reexposure of the target just after the JOL should have a minimal effect on retrievability because the reexposure acts as the second of two massed study opportunities. For that reason, the memory hypothesis predicts that post-JOL reexposure should have a minimal effect on final recall for most categories of items in our design, including (1) all items in the delayed cue–target condition, because targets are reexposed at the time of the JOL, (2) all items in the immediate cue-only condition, because targets are almost certainly retrieved at the time of the JOL and, besides, are initially presented just before that, and (3) items given high JOLs in the delayed cue-only condition, because they are successfully retrieved at the time of the JOL.

On the other hand, if a target has been neither reexposed nor retrieved at the time of the JOL, the memory hypothesis predicts that the post-JOL reexposure should boost retrievability to much the same degree as reexposure or retrieval at the time of the JOL would have boosted retrievability, had it occurred. According to the memory hypothesis, the only category of items in our design having targets that are neither reexposed nor retrieved at the time of the JOL are the items given low JOLs in the delayed cue-only condition; indeed, the hypothesis assumes that the failure to retrieve such targets is the reason the items are given low JOLs, consistent with participants' statements in postexperimental debriefings in Nelson and Dunlosky (1991) and in our study. Moreover, the post-JOL reexposure of those low-JOL items occurs at a substantial delay following initial study and, therefore, serves as a spaced study opportunity that should boost retrievability substantially, according to the memory hypothesis. Thus, the memory hypothesis predicts that the low-JOL items in the delayed cue-only condition are the only items in our design for which post-JOL reexposure should cause a substantial boost in retrievability and, therefore, in final recall. Indeed, by eliminating differences in spaced study opportunities for low-JOL items in the delayed cue-only versus delayed cue–target conditions, post-JOL reexposure should yield approximately equal recall of the low-JOL items in those two conditions, according to the memory hypothesis.

Regarding the effects of post-JOL reexposure on gamma, the memory hypothesis predicts that gamma in the delayed cue–target and immediate cue-only conditions should not vary substantially as a function of post-JOL reexposure, because the memory hypothesis predicts that post-JOL reexposure should not differentially affect recall of high-JOL versus low-JOL items in those conditions.[2] By contrast, in the delayed cue-only condition, the memory hypothesis predicts a substantial boost in recall for the low-JOL items, but not for the high-JOL items, as a consequence of post-JOL reexposure, so it predicts that gamma should drop substantially with post-JOL reexposure in that condition.[3] In fact, if the difference in gamma that makes up the delayed-JOL effect is attributable to differences in spaced study opportunities, as the memory hypothesis posits, and if post-JOL reexposure eliminates the differences in spaced study opportunities for low-JOL items in the delayed cue-only versus delayed cue–target conditions, then according to the memory hypothesis, post-JOL reexposure should reduce gamma in the delayed cue-only condition to approximately the same level as that in the delayed cue–target condition.

## EXPERIMENT 1

In both Experiments 1 and 2, we manipulated the format used to request delayed JOLs, using either the cue–target pair or the cue alone as a stimulus for the JOL. In Experiment 1, we briefly reexposed the cue–target pair shortly after the JOL for each pair in both JOL format conditions; in Experiment 2, we manipulated the occurrence of this post-JOL reexposure between subjects. In all other material respects, the methods used in the two experiments were identical.

### Method

**Participants**. The participants were 36 undergraduates enrolled in introductory psychology courses at Columbia University and Barnard College. They participated for course credit.

**Procedure**. After participating in a practice session as the experimenter read instructions, the participants studied 84 word pairs that appeared on an iMac monitor, one pair at a time for 4 sec each (500 msec of blank screen separated the items in all the phases). The participants were asked to learn the pairs for a later test in which they would be asked to type the second word when given the first.

In the judgment phase that followed study of all the pairs, the participants made delayed JOLs for each of the studied pairs. Each of the word pairs appeared one pair at a time, and for each pair, the participants made a JOL, using a 7-point scale to indicate their confidence that they would be able to recall the target when given the cue on the final test approximately 25 min later; a higher JOL indicated greater confidence. The participants were asked to make their judgments quickly; every 4 sec after the stimulus appeared, they received an oral "hurry" signal to enter their judgment immediately, a signal they were asked to avoid triggering. This signal was intended to prevent the participants from spending unlimited time intentionally rehearsing items, particularly in the cue–target condition.

Pairs were judged in sequences of six pairs; the JOL formats in each set of six pairs were either all cue-only or all cue–target. After every sixth pair was judged, the preceding six pairs reappeared for 1 sec each in random order as cue–target pairs, regardless of whether the pairs in that set had been presented in cue-only or cue–target format for the JOLs.[4] Note that, in contrast to Experiments 2 and 3, reexposure was not manipulated in Experiment 1; all the items were reexposed.

After the judgment phase, there followed a test phase in which the participants were asked to type the target word from each pair when given the cue word. They proceeded at their own pace.

**Materials and Apparatus**. The experiment was presented on iMac computers. Studied items consisted of 168 singular, four-letter, common nouns with high word frequency, familiarity, concreteness, and imageability. The computer randomly paired the words anew for each participant, yielding 84 word pairs that the computer randomly ordered for the study phase. Twelve of the pairs served as buffer pairs that were presented as the first and last 6 pairs in the study phase, as the first 12 pairs in the judgment phase, and as the first 12 pairs in the test phase. Results for buffer pairs were not analyzed.

In the judgment phase, the 72 nonbuffer pairs were assigned to 12 sets of 6 pairs, using block randomization to ensure that each set contained 1 pair from among the first, second, third, fourth, fifth, and sixth groups of 12 nonbuffer studied pairs. Half of the sets were assigned to each of the two JOL formats. Sets using the two formats were distributed across the judgment phase sequence, and this order was counterbalanced across participants. In the test phase, block randomization ensured that two items from each judgment phase set of six items appeared among the first, second, and last thirds of the test sequence.

### Results and Discussion

**Gamma correlations**. A dependent-samples $t$ test revealed that the mean gamma correlations in the delayed cue-only condition ($M = .52$, $SE = .07$) and the delayed cue–target condition ($M = .41$, $SE = .06$) did not differ reliably [$t(34) = 1.16$, $p > .25$]. (One participant was excluded from this analysis because gamma could not be calculated in the cue–target condition.) This lack of a reliable effect contrasts with the difference in gamma typically observed between these two conditions (e.g., Dunlosky & Nelson, 1992, 1997).

**Recall as a function of JOL level**. Figure 1 depicts mean recall percentages as a function of the participants' JOL selections. These curves are related to gamma in that they depict the relative recall percentages of items to which the participants assigned different JOLs, although recall percentages at each JOL level are averaged across participants. As can be seen from the figure, recall of items given low JOLs differed negligibly between the cue-only and the cue–target conditions. This contrasts with the typical pattern, in which recall of low-JOL items is higher in the cue–target condition than in the cue-only condition (see Dunlosky & Nelson, 1992, 1997). In fact, recall in the cue-only and cue–target conditions was highly similar across all JOL levels, and overall recall was virtually identical in the cue-only condition ($M = 45$, $SE = 5$) and the cue–target condition ($M = 44$, $SE = 4$; $t < 1$).

This pattern suggested that the brief post-JOL cue–target reexposure provided a crucial study opportunity that was otherwise unavailable for the low-JOL items in the delayed cue-only condition and that eliminating the difference in study opportunities was sufficient to eliminate the difference in gamma correlations. To determine whether the brief post-JOL cue–target reexposure was indeed responsible for eliminating the difference in gamma, we manipulated the occurrence of such reexposure between subjects in Experiment 2.

## EXPERIMENT 2

By factorially manipulating JOL format and post-JOL target reexposure for delayed-JOL items in this experi-

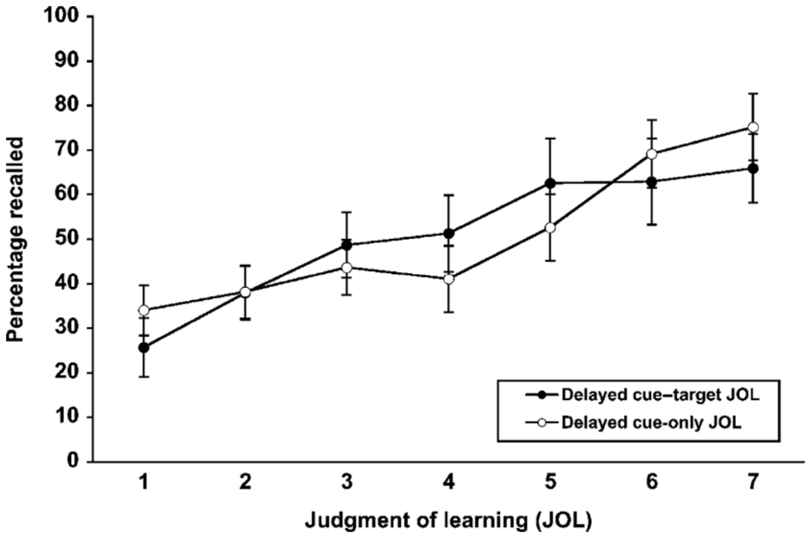


**Figure 1. Mean percentage of correct final recall (with standard errors) as a function of judgment of learning (JOL) format and JOL level in Experiment 1.**

ment, we were able to compare gamma correlations and recall patterns in four conditions that differed as to the occurrence and timing of target reexposure. In these conditions, targets were reexposed (1) both before and after the JOL (the delayed cue–target reexposed condition), (2) after but not before the JOL (the delayed cue-only reexposed condition), (3) before but not after the JOL (the delayed cue-target unreexposed condition), or (4) neither before nor after the JOL (the delayed cue-only unreexposed condition).

The memory hypothesis predicted that gamma correlations and recall patterns would be similar in the first three of these conditions, because all the items in these conditions would benefit from spaced study. The memory hypothesis also predicted that those patterns would differ from the pattern in the delayed cue-only unreexposed condition, in which only retrieved items (those given high JOLs) would receive a boost in recallability from spaced study, but unretrieved items (those given low JOLs) would not receive spaced study and, thus, would receive no boost in recallability.

## Method

The method was identical to that used in Experiment 1, except in three respects: (1) The participants were 72 undergraduates enrolled in introductory psychology courses at Columbia University and Barnard College, participating for course credit, (2) a randomly assigned 36 participants received the brief post-JOL reexposure of cue–target pairs as in Experiment 1, whereas the other 36 participants did not, and (3) we shortened the time that the participants spent restudying items other than those included in our analyses (see note 4).

## Results

**Gamma correlations**. The mean gamma correlations in each condition are shown in Figure 2. A mixed analysis of variance (ANOVA) revealed a reliable reexposure × JOL format interaction [$F(1,70) = 7.48, MS_e = 0.16, p = .0079$]. The pattern in the conditions with post-JOL reexposure replicated the pattern from Experiment 1: Gammas in the delayed cue-only condition ($M = .39, SE = .09$) and the delayed cue–target condition ($M = .46, SE = .05$) were in the moderately positive range and did not differ reliably ($F < 1$). However, in the unreexposed condition, gamma was reliably higher for the delayed cue-only condition ($M = .72$, $SE = .06$) than for the delayed cue–target condition [$M = .43$, $SE = .07$; $F(1,70) = 9.93$, $MS_e = 0.16, p = .0024$], consistent with the typical results reported by Dunlosky and Nelson (1992, 1997). In addition, reexposure had no reliable effect on gamma for cue–target JOLs ($F < 1$), but gamma was reliably higher for cue-only JOLs without reexposure than for those with reexposure [$F(1,70) = 8.76$, $MS_e = 0.22, p = .0042$].

**Recall as a function of JOL level**. Figure 3 plots the curves for final recall as a function of the participants' JOL selections. We conducted a mixed ANOVA that included groups of JOL levels as a factor, grouping the lowest three JOL levels together and the highest four JOL levels together and using a conservative alpha level of .01. This analysis excluded 6 participants who had been included in the gamma analysis (4 in the reexposed condition, 2 in the unreexposed condition), because they not use JOL levels from both of these sets in both the cue-only and the cue–target conditions. The ANOVA revealed a reliable three-way interaction of JOL format, JOL level, and reexposure [$F(1,64) = 10.60, MS_e = 300.06, p = .0018$], such that there was a reliable simple effect of JOL format for low-JOL items without reexposure, with higher recall in the cue–target condition than in the cue-only condition [$F(1,64) = 78.56, MS_e = 257.23, p < .0001$], but JOL format did not reliably affect recall of high-JOL items without reexposure ($F < 1$), low-JOL items with reexposure ($F < 1$), or high-JOL items with reexposure [$F(1,64) =$

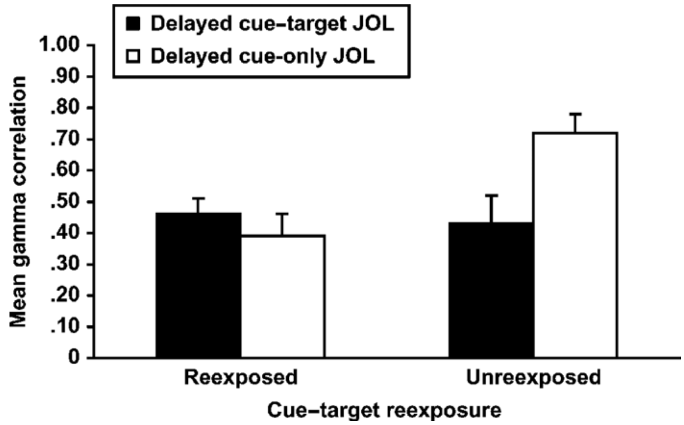


**Figure 2. Mean gamma correlations between judgments of learning (JOLs) and final recall (with standard errors) as a function of JOL format and cue–target reexposure in Experiment 2.**

2.70, $MS_e = 273.39, p > .10$]. In addition, reexposure increased recall of the low-JOL items in the cue-only condition [$F(1,64) = 30.48$, $MS_e = 490.51$, $p < .0001$] but did not affect recall of the high-JOL items in that condition ($F < 1$) or recall of either the low-JOL or high-JOL items in the cue–target condition ($F$s $< 1$).[5]

**Overall recall**. A mixed ANOVA indicated that the mean recall percentage in the cue-only condition was reliably higher when the target was reexposed ($M = 51$, $SE = 4$) than when it was not reexposed [$M = 33$, $SE = 4$; $F(1,70) = 9.47$, $MS_e = 580.14$, $p = .0030$], but in the cue–target condition, recall did not differ reliably between the reexposed condition ($M = 56$, $SE = 4$) and the unreexposed condition ($M = 56$, $SE = 4$; $F < 1$), yielding a reliable reexposure × JOL format interaction [$F(1,70) = 26.27$, $MS_e = 106.18$, $p < .0001$].

## Discussion

The results were consistent with the predictions of the memory hypothesis. In the standard conditions, those not involving post-JOL cue–target reexposure, the only difference in final recall was that recall of the low-JOL items was lower in the delayed cue-only condition than in the delayed cue–target condition. Post-JOL reexposure affected recall only of the low-JOL items in the delayed cue-only condition, which increased to a level comparable to that of low-JOL items in the delayed cue–target condition. This finding is consistent with the claim of the memory hy-

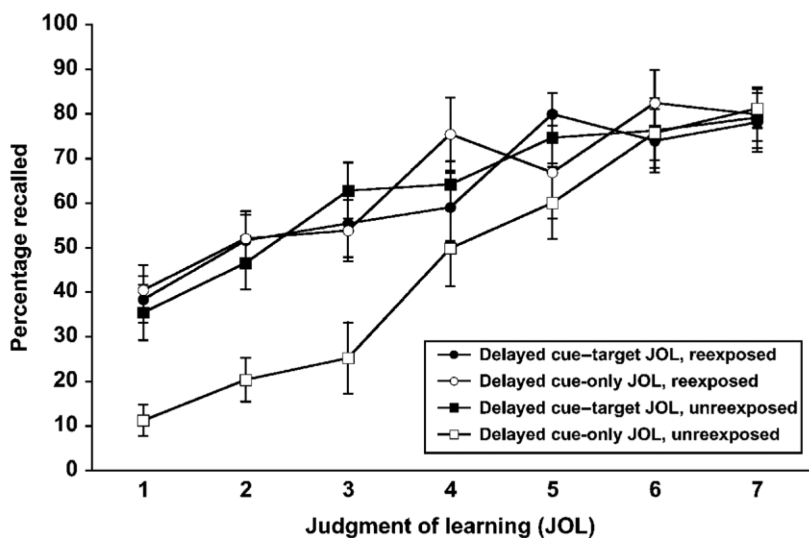


**Figure 3. Mean percentage of correct final recall (with standard errors) as a function of judgment of learning (JOL) format, cue–target reexposure, and JOL level in Experiment 2.**

pothesis that, in the standard conditions not involving post-JOL reexposure, the low-JOL items in the delayed cue-only condition are the only items that are neither covertly retrieved nor reexposed at the time of the JOL and, therefore, are the only items to experience a spaced study opportunity through post-JOL reexposure and the only items to receive an attendant boost in retrievability.

Redressing the inequality in spaced study opportunities through post-JOL reexposure not only erased the single difference between the conditions in the pattern of recall, but also erased the superiority in gamma for delayed cue-only JOLs over delayed cue–target JOLs. The gamma results thus supported the claim of the memory hypothesis that, in the standard conditions not involving post-JOL reexposure, differences in spaced study opportunities at the time of the JOL are responsible for the higher gamma correlation in the delayed cue-only condition.

In addition, the similarity in gamma and recall patterns regardless of whether target reexposure preceded or followed a delayed JOL poses a challenge for metamemory explanations of the delayed-JOL effect. We address this issue in the General Discussion section.

## EXPERIMENT 3

Given that Experiments 1 and 2 suggested that the superiority in the gamma correlation for delayed cue-only JOLs over delayed cue–target JOLs is attributable to differences in spaced study opportunities at the time of the delayed JOL, in Experiment 3 we sought to determine whether a similar explanation applies to the superiority in gamma for delayed cue-only JOLs over immediate cue-only JOLs. To address this question, we again manipulated the occurrence of post-JOL cue–target reexposure, this time within subjects.

Manipulation of the timing of the JOL provided an opportunity to test the claim of the memory hypothesis that the timing of the differential study opportunities plays a critical role in the delayed-JOL effect. The memory hypothesis predicts that spacing effects should be observed only in the delayed cue-only condition and only for items in that condition that are retrieved or reexposed at the time of a delayed JOL, immediately thereafter, or at both times. By contrast, in the immediate cue-only condition, the memory hypothesis posits that the absence of delay between initial study and JOL and between JOL and post-JOL reexposure yields no spacing of study opportunities and, therefore, no spacing effects. Accordingly, with one important exception, the memory hypothesis predicts that, regardless of post-JOL reexposure, high- and low-JOL items in the delayed cue-only condition should be recalled at higher rates than high- and low-JOL items, respectively, in the immediate cue-only condition. The exception is that the hypothesis predicts that low-JOL items in the delayed cue-only condition that are not reexposed following the JOL should not receive spaced study and, therefore, should be recalled at a rate comparable to that of low-JOL items in the immediate cue-only condition.

Predictions regarding gamma in the delayed cue-only condition are identical to those made in Experiment 2: The memory hypothesis predicts that post-JOL cue–target reexposure should decrease gamma substantially by eliminating differences in spaced study opportunities for high-JOL versus low-JOL items. In the immediate cue-only condition, the memory hypothesis predicts that post-JOL reexposure should not affect gamma, because it should not differentially affect the availability of spaced study opportunities for high-JOL versus low-JOL items, unlike in the delayed cue-only condition.

### Method

The method was designed to match that used by Nelson and Dunlosky (1991) in material respects, and as a result, it differed substantially from the method used in Experiments 1 and 2.

**Participants**. The participants were 40 undergraduates enrolled in introductory psychology courses at Columbia University and Barnard College, participating for course credit.

**Design**. The experiment used a 2 (JOL timing: immediate cue-only JOL vs. delayed cue-only JOL) × 2 (target reexposure: cue–target vs. cue-only post-JOL reexposure) within-subjects factorial design.

**Procedure**. The participants first proceeded through a practice session as the experimenter read instructions. In the experiment proper, there were two phases, a study-and-JOL phase and a test phase. In the study-and-JOL phase, the participants were asked to learn word pairs for a later test in which they would be asked to type the second word when given the first word. The participants studied each of 84 word pairs during the 4 sec it appeared on an iMac screen (500 msec of blank screen separated the items in both the phases). The first 12 word pairs were buffer items that were not tested. The other 72 word pairs were divided into two sets of 36 pairs, and the participants performed the same set of tasks for one set and then for the other. In each set, the participants made immediate cue-only JOLs for 18 of the pairs immediately following study of the pair. For the other 18 pairs in the set, the participants made delayed cue-only JOLs in a block that followed presentation of the 36th pair in that set.

The participants were asked to make their judgments quickly; every 4 sec after an item appeared, they received an oral "hurry" signal to enter their judgment immediately, a signal they were asked to avoid triggering. Following entry of an immediate or delayed JOL for an item, either the cue–target pair or the cue alone was reexposed for 2 sec; this procedure implemented our manipulation of target reexposure. In each of the four 18-pair subsets, both the cue and the target were reexposed for 9 pairs, and the cue alone was reexposed for the other 9 pairs.

In the test phase, the participants were asked to type in the target word from each pair when given the cue word. They proceeded at their own pace.

**Materials**. The studied items were the 168 words from Experiments 1 and 2. Of these, 24 words were randomly selected and paired and were used as buffer items for all the participants. The remaining 144 words were randomly paired for each participant anew, and the resulting 72 pairs were randomly ordered for study. In the study-and-JOL phase, pairs were assigned to conditions, using block randomization to ensure that the four conditions were each assigned 3 pairs in each of six successive blocks of 12 pairs. The order for delayed JOLs preserved the order of these six blocks but rerandomized the order of the 6 delayed-JOL pairs within each of the six blocks. The test order consisted of the first 36 studied items in rerandomized order, followed by the last 36 studied items in rerandomized order.

### Results

**Gamma correlations**. Figure 4 shows the mean gamma correlations in each condition. We conducted a within-subjects ANOVA on the gamma correlations between

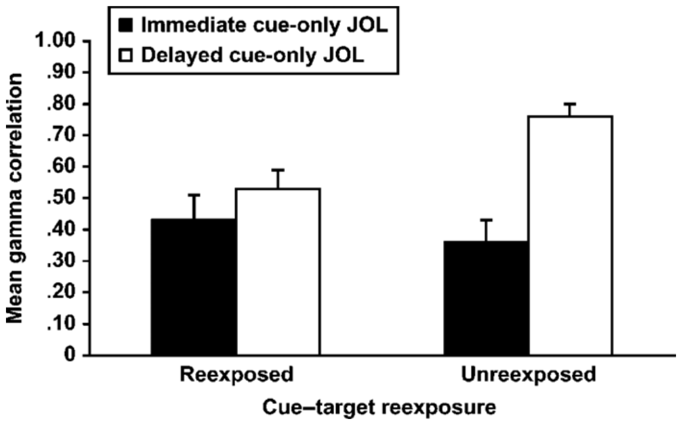


**Figure 4. Mean gamma correlations between judgments of learning (JOLs) and final recall (with standard errors) as a function of JOL timing and cue–target reexposure in Experiment 3.**

JOLs and final recall, using data from the 36 participants for whom gamma could be calculated in all conditions. The ANOVA revealed a reliable target reexposure × JOL timing interaction [$F(1,35) = 6.31$, $MS_e = 0.13$, $p = .0168$]. When the target was not reexposed, there was a typical delayed-JOL effect, with a higher gamma for delayed cue-only JOLs ($M = .76$, $SE = .04$) than for immediate cue-only JOLs [$M = .36$, $SE = .06$; $F(1,35) = 24.29$, $MS_e = 0.12$, $p < .0001$]. By contrast, when the target was reexposed, gammas for delayed cue-only JOLs ($M = .53$, $SE = .07$) and immediate cue-only JOLs ($M = .43$, $SE = .08$) did not differ reliably ($F < 1$). In addition, for delayed cue-only JOLs, gamma was higher when the target was unreexposed than when it was reexposed [$F(1,35) = 8.95$, $MS_e = 0.10$, $p = .0051$], replicating the pattern observed for comparable conditions in Experiment 2. Reexposure had no reliable effect on gamma for immediate cue-only JOLs ($F < 1$).

**Recall as a function of JOL level**. Figure 5 depicts mean correct recall percentages in each condition as a function of JOL level. We conducted a within-subjects ANOVA, using a conservative alpha level of .01, that included groups of JOL levels as a factor, again grouping the lowest three JOL levels together and the highest four JOL levels together. The analysis excluded 7 of the 36 participants who were included in the gamma analysis, because who did not use JOL levels from both of these JOL groups in all four conditions; this exclusion had negligible effects on the means for each JOL group in each condition. This analysis showed that the pattern for delayed cue-only JOLs replicated that for comparable conditions in Experiment 2, in that target reexposure reliably increased

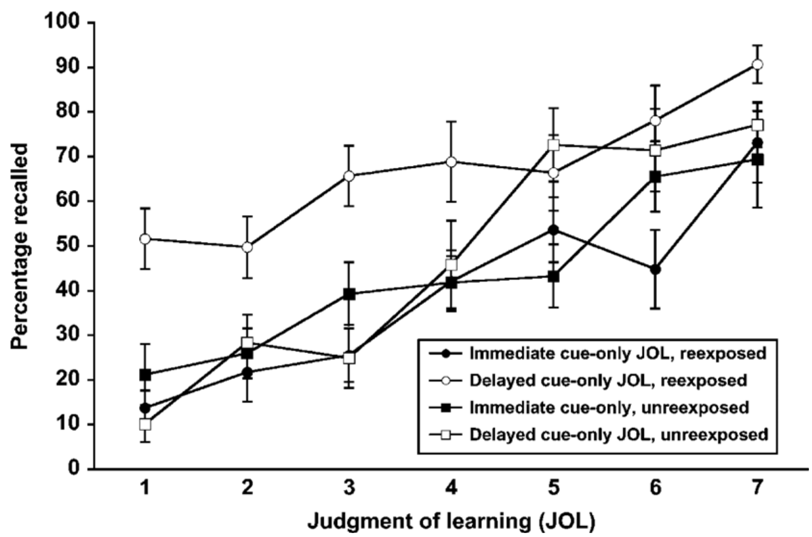


**Figure 5. Mean percentage of correct final recall (with standard errors) as a function of judgment of learning (JOL) timing, cue–target reexposure, and JOL level in Experiment 3.**

recall of low-JOL items [$F(1,28) = 39.24$, $MS_e = 293.15$, $p < .0001$], but not recall of high-JOL items [$F(1,28) = 4.89$, $MS_e = 362.49$, $p = .0354$], yielding a reliable simple interaction of JOL level and target reexposure for delayed cue-only JOLs [$F(1,28) = 8.36$, $MS_e = 253.85$, $p = .0073$].[6]

The pattern of ANOVA results for conditions with unreexposed targets was consistent with previous research (Dunlosky & Nelson, 1992; Weaver & Kelemen, 1997): As compared with recall in the immediate-JOL condition, recall in the delayed-JOL condition was reliably higher for high-JOL items [$F(1,28) = 17.97$, $MS_e = 329.79$, $p = .0002$] and numerically but unreliably lower for low-JOL items [$F(1,28) = 4.69$, $MS_e = 370.07$, $p = .0389$], yielding a reliable simple interaction of JOL level and JOL timing for items with unreexposed targets [$F(1,28) = 19.44$, $MS_e = 362.17$, $p = .0001$]. By contrast, in conditions with reexposed targets, recall in the delayed-JOL condition was reliably higher than that in the immediate-JOL condition for both high-JOL items [$F(1,28) = 27.87$, $MS_e = 396.50$, $p < .0001$] and low-JOL items [$F(1,28) = 33.05$, $MS_e = 218.40$, $p < .0001$], and JOL level and timing did not interact ($F < 1$). Consequently, there was a reliable three-way interaction of target reexposure, JOL timing, and JOL level [$F(1,28) = 10.23$, $MS_e = 237.10$, $p = .0034$].

**Overall recall**. A within-subjects ANOVA of overall recall percentages revealed a reliable target reexposure × JOL timing interaction [$F(1,35) = 55.02$, $MS_e = 94.80$, $p < .0001$]. Recall in the delayed-JOL condition following target reexposure ($M = 61$, $SE = 4$) was reliably higher than recall in each of the other three conditions ($p$s $< .0001$), among which recall did not differ reliably ($p$s $> .10$)—that is, delayed JOLs without target reexposure ($M = 40$, $SE = 3$), immediate JOLs with target reexposure ($M = 37$, $SE = 3$), and immediate JOLs without target reexposure ($M = 41$, $SE = 4$).

## Discussion

The results provided additional evidence that the delayed-JOL effect is attributable to differences in spaced study opportunities at the time of delayed cue-only JOLs, as the memory hypothesis posits. Using a method similar to that in Nelson and Dunlosky (1991), we replicated the pattern of results from Experiment 2 for the delayed cue-only condition: Recall of items given low JOLs increased reliably with post-JOL cue–target reexposure, but recall of items given high JOLs did not, as predicted by the memory hypothesis. As in Experiment 2, this pattern is consistent with the claim of the memory hypothesis that high-JOL items are retrieved at the time of the JOL and, therefore, receive study opportunities at that time, whereas the low-JOL items are not retrieved at that time and receive an additional spaced study opportunity only upon post-JOL cue–target reexposure.

The results further showed that the timing of the additional study opportunities plays a critical role in the delayed-JOL effect by yielding spacing effects only for items that receive study opportunities after a delay. Leaving aside the delayed cue-only unreexposed condition for a moment, the other three conditions exhibit a pattern that reflects typical spacing effects across all JOL levels. As can be seen in Figure 5, cue–target reexposure had no effect in the immediate-JOL conditions across all JOL levels, but as compared with those conditions, recall in the reexposed delayed-JOL condition was substantially higher across all JOL levels. This pattern reflects typical spacing effects, in that all the items in the immediate-JOL conditions received massed study opportunities (initial study and retrieval at the time of the immediate JOL, with or without post-JOL reexposure) and all the items in the reexposed delayed-JOL condition received spaced study opportunities (initial study and reexposure following the delayed JOL, with or without retrieval at the time of the JOL).[7]

By comparison, in the delayed cue-only unreexposed condition, the high-JOL items exhibited recall comparable to that of high-JOL items in the delayed cue-only reexposed condition, whereas low-JOL items exhibited recall comparable to that of low-JOL items in the immediate cue-only condition. This pattern is consistent with the high-JOL items' being retrieved at the time of the JOL and, thus, receiving a spaced study opportunity that the unretrieved low-JOL items did not. Thus, the evidence supported the claims of the memory hypothesis that, in the delayed cue-only unreexposed condition, differences in covert retrieval success for high-JOL items versus low-JOL items create differences in study opportunities and that the timing of those differential study opportunities affects final recall through differences in spacing effects.

By erasing differences in spaced study opportunities, post-JOL cue–target reexposure also eliminated the superiority in gamma for delayed cue-only JOLs over immediate cue-only JOLs. The elimination of this difference in gamma points to the sufficiency of the differences in spaced study opportunities as an explanation for the delayed-JOL effect in the standard conditions not involving post-JOL reexposure.

## ADDITIONAL ANALYSIS
### Distributions of JOL Assignments

In agreement with previous research (Dunlosky & Nelson, 1994, 1997; Weaver & Kelemen, 1997), our participants used the more extreme portions of the JOL rating scale when assigning JOLs in the delayed cue-only condition than in the delayed cue–target condition and the immediate cue-only condition. Figure 6 shows the percentage of items that the participants assigned to particular JOL levels in the various conditions in Experiments 1, 2, and 3. This polarization in JOL assignments is consistent with the participants' according great weight to the current accessibility of target-related information in making delayed cue-only JOLs (see Benjamin, Bjork, & Schwartz, 1998) and a high degree of differentiation in such accessibility across items in the delayed cue-only condition, due to differences in the success rates for covert retrieval.

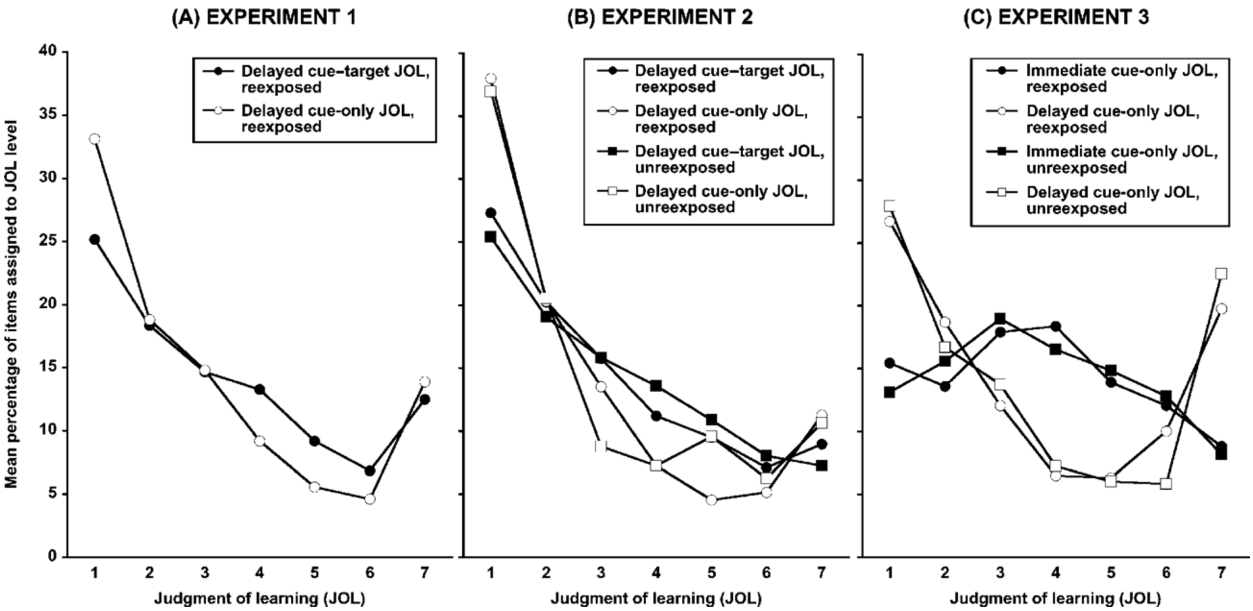


**Figure 6. Judgment of learning (JOL) distributions as a function of cue–target reexposure, JOL format, and/or JOL timing in Experiments 1, 2, and 3.**

The issue of interest here is whether such polarization of JOL assignments makes a unique contribution to the delayed-JOL effect, distinct from the contribution of differences in spaced study opportunities. Because the delayed-JOL effect disappeared when post-JOL cue–target reexposure controlled for differential study opportunities, there is no portion of the effect that remains to be explained by JOL polarization. Accordingly, our results dovetail with results reported by Weaver and Kelemen (1997), who similarly found that the contribution of this polarization to the delayed-JOL effect is negligible.

Moreover, the polarization of JOL distributions is entirely consistent with the memory hypothesis. According to that hypothesis, judgments are likely to be more polarized in the delayed cue-only condition than in the other conditions simply because the distribution of associative strengths is likely to be more polarized among the pairs in that condition following attempts to retrieve targets. Only in that condition do some items (the retrieved high-JOL items) receive a boost in retrievability from spaced study, but other items (the unretrieved low-JOL items) do not, according to the memory hypothesis. By contrast, all of the items in the delayed cue–target condition receive a boost in retrievability from spaced study, and none of the items in the immediate cue-only condition receive such a boost, so the distribution of strengths in each of those conditions is likely to be less polarized than that in the delayed cue-only condition, according to the memory hypothesis. Thus, the differences in the polarization of JOL assignments among the conditions can be explained by differences in the polarization of associative strengths among the conditions, and therefore, they need not be attributed to differences in metacognitive acuity among the conditions.

## GENERAL DISCUSSION

Our results strongly support the memory hypothesis as an explanation of the delayed-JOL effect. The results indicate that the delayed-JOL effect stems from differences in spaced study opportunities for high-JOL and low-JOL items in that condition, caused by differences in the success of covert retrieval for those items at the time of the delayed JOL. Our evidence therefore indicates that the delayed-JOL effect is a memory phenomenon rather than a metamemory phenomenon.

### Ramifications for Metamemory Explanations of the Delayed-JOL Effect

In Experiment 2, gamma and recall patterns were similar regardless of whether target reexposure preceded or followed a delayed JOL. This similarity in patterns poses a challenge for metamemory explanations of the delayed-JOL effect. Those explanations posit that the delayed-JOL effect arises because JOL timing (immediate vs. delayed) and JOL format (cue-only vs. cue–target) influence the metacognitive processes used in making JOLs. However, those metacognitive processes can be informed only by pre-JOL target reexposure, not by post-JOL target reexposure. Therefore, any influences of target reexposure on such metacognitive processes cannot explain the effects of post-JOL target reexposure on gamma and recall patterns or why those effects are similar to the effects of pre-JOL target reexposure.[8]

Of course, proponents of metamemory explanations could posit some additional mechanism to explain the effects of post-JOL target reexposure on gamma and recall. Any such additional mechanism is likely to be very much like the spacing mechanism posited by the memory hypothesis, but in any event, our data place important constraints on the nature of such an additional mechanism. In particular, such proponents cannot simply claim that post-JOL target reexposure would disrupt the relationship between JOLs and recall in some abstract and unspecified way. Instead, they must propose a mechanism that explains the pattern of results that we actually observed. Accordingly, the mechanism must explain not just the effects of post-JOL target reexposure in the delayed cue-only condition—the decrease in gamma, the increase in recall of low-JOL items, and the absence of any change in recall of high-JOL items—but also the absence of any effects of such reexposure on gamma and recall in the immediate cue-only and delayed cue–target conditions. Barring the proposal of an additional mechanism that completely explains the observed pattern of results following post-JOL reexposure, the spacing mechanism posited by the memory hypothesis must be regarded as superior because it does provide such a complete explanation.

Moreover, because proponents of metamemory explanations must posit an additional mechanism to explain the effects of post-JOL reexposure in our study, they face problems with parsimony. To avoid these problems, such proponents face the additional challenge of explaining why it is necessary to posit two different mechanisms to explain the highly similar patterns that we observed for gamma and recall in the conditions in which target reexposure preceded or followed delayed JOLs. In particular, they would need to provide evidence that each of those mechanisms explains some portion of the data that the other mechanism cannot. By contrast, the memory hypothesis does not face these problems with parsimony, because recallability of a cue–target pair can be affected by spaced target reexposure regardless of whether it precedes or follows a delayed JOL, so the spacing mechanism posited by the memory hypothesis has no difficulty in explaining the similarity in gamma and recall patterns that we observed.

## Memory Monitoring and JOLs

It is not our claim that memory monitoring plays *no* role in JOLs. Obviously, the moderately positive gamma values for delayed cue–target and immediate cue-only JOLs indicate that participants are able to monitor the state of their memories with some degree of accuracy. The metacognitive nature of the processes involved is reflected in our participants' reports that they considered several types of information in making their JOLs, consistent with other research and with the cue utilization framework suggested by Koriat (1993, 1997). They reported considering such things as degree of association between the cue and the target (e.g., Rabinowitz, Ackerman, Craik, & Hinchley, 1982), ability to form an interactive image between the cue and the target (e.g., Dunlosky & Nelson, 1994), cue familiarity (e.g., Metcalfe, Schwartz, & Joaquim, 1993), and ease and speed with which information came to mind at recall (e.g., Kelley & Lindsay, 1993; Mazzoni & Nelson, 1995).

In the delayed cue-only condition, our participants also reported considering the success or failure of a covert attempt to retrieve the target, with retrieved items being given higher JOLs than were unretrieved items. The memory hypothesis assumes that participants use just such a strategy in assigning JOLs. However, that assumption does not imply that retrieval success or failure is the only factor they consider in assigning JOLs. For example, there could well be a further classification among the retrieved high-JOL items based on consideration of the other factors noted above; the memory hypothesis is silent on that issue. Accordingly, the memory hypothesis is not inconsistent with our generally upward-sloping recall functions for high-JOL items or with such findings as the moderately positive gamma that Nelson and Dunlosky (1992) reported for items that participants had succeeded in overtly retrieving prior to making delayed cue-only JOLs.

Importantly, however, the memory hypothesis claims that, even though people use metacognitive processes to make JOLs and even though those processes yield moderately accurate recall predictions, those processes are not themselves also responsible for the increase in gamma for delayed cue-only JOLs. Our evidence supports that claim.

## Overall Recall

In Experiment 3, we found that mean overall recall did not differ between the delayed cue-only and the immediate cue-only conditions. The memory hypothesis is not inconsistent with this finding, because it makes no prediction regarding overall final recall. In particular, it does not predict that the *overall* probability of recall will necessarily be greater in the delayed cue-only condition than in the other conditions (cf. Nelson & Dunlosky, 1992). Rather, it predicts that a delay between initial study and attempted retrieval at the time of the JOL will increase the *conditional* probability that items retrieved successfully at that point will also be successfully recalled later. As Kelemen and Weaver (1997) have noted, the effect of such successful retrieval on overall final recall depends not only on that conditional probability, but also on the proportion of such items that are, in fact, successfully retrieved at the time of the JOL (see also Whitten & Bjork, 1977). Of course, that proportion is likely to vary with the interval between initial study and the JOL-related retrieval; increasing the delay might increase the conditional probability of successful final recall given successful covert retrieval (as the memory hypothesis predicts), but overall final recall might remain the same or even decline, because the proportion of items that are successfully retrieved would decline as the delay increased. This reasoning applies in reverse to shorter delays: When Kelemen and Weaver used a relatively short interval between initial study and JOL, they reported higher mean final recall in the delayed cue-only condition than in the immediate cue-only condition.

## Implications for Education

Our results suggest that students may be equally capable of judging which information they are likely to remember or forget across a retention interval, regardless of whether their judgment follows an immediate recall test, a delayed recall test, or a delayed restudy episode. This suggestion is subject to an important assumption, however: At the beginning of the retention interval, the information must have had a similar study history—that is, it must have been restudied or retrieved at similar points in time, including any restudy or retrieval that occurs at the time of the judgment. Nonequivalent study histories can make it appear as though the stimulus for the judgment, or its timing, has an effect on metacognitive accuracy.

As a practical matter, our results raise a question as to whether there is any measurable metacognitive benefit to making JOLs based on delayed retrieval attempts. At a minimum, our results suggest that the target information should be reexposed following a delayed JOL, to take advantage of the cognitive benefit from such a spaced study opportunity. This cognitive benefit is evident in the higher overall recall in the delayed cue-only condition with such reexposure than without. Such a strategy would afford a student all the cognitive benefits of spacing, with apparently no substantial metacognitive costs.

## Implications for Metacognitive Research

The most obvious and direct implication of our results for metacognitive research is that the delayed-JOL effect should not be regarded as a metacognitive effect. More generally, our results underscore the importance of considering and controlling for the effects on memory of variables used to study metamemory. Distinguishing between memory effects and metamemory effects can be troublesome, especially with a phenomenon such as the delayed-JOL effect, in which the underlying processes used to make JOLs are indeed metacognitive but the difference to be explained is attributable to processes that are not metacognitive. Failure to consider memory effects may lead, as in this case, to a Heisenberg-like effect: A manipulation thought to affect metamemory processes instead affects the memories that those processes are used to assess.

## REFERENCES

Allen, G. W., Mahler, W. A., & Estes, W. K. (1969). Effects of recall tests on long-term retention of paired associates. *Journal of Verbal Learning & Verbal Behavior*, **8**, 463-470.

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., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamemory index. *Journal of Experimental Psychology: General*, **127**, 55-68.

Bower, G., & Winchester, P. (1970). *Metamemory judgments.* Unpublished manuscript, Stanford University, Palo Alto, CA.

Crowder, R. G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Erlbaum.

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 & Language*, **33**, 545-565.

Dunlosky, J., & Nelson, T. O. (1997). Similarity between the cue for judgments of learning (JOL) and the cue for test is not the primary determinant of JOL accuracy. *Journal of Memory & Language*, **36**, 34-49.

Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition*, **7**, 95-112.

Glenberg, A. M., Sanocki, T., Epstein, W., & Morris, C. (1987). Enhancing calibration of comprehension. *Journal of Experimental Psychology: General*, **116**, 119-136.

Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 371-377.

Hintzman, D. L. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola symposium* (pp. 77-99). Hillsdale, NJ: Erlbaum.

Hintzman, D. L. (1976). Repetition and memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 11, pp. 47-91). New York: Academic Press.

Kelemen, W. L., & Weaver, C. A., III (1997). Enhanced metamemory at delays: Why do judgments of learning improve over time? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 1394-1409.

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 & Language*, **32**, 1-24.

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

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, **126**, 349-370.

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

Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning? *Journal of Experimental Psychology: General*, **131**, 349-363.

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 851-861.

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., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect." *Psychological Science*, **2**, 267-270.

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

Nelson, T. O., Dunlosky, J., Graf, A., & Narens, L. (1994). Utilization of metacognitive judgments in the allocation of study during multitrial learning. *Psychological Science*, **5**, 207-213.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 26, pp. 125-173). New York: Academic Press.

Rabinowitz, J. C., Ackerman, B. P., Craik, F. I. M., & Hinchley, J. L. (1982). Aging and metamemory: The roles of relatedness and imagery. *Journal of Gerontology*, **37**, 688-695.

Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 204-221.

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.

Vesonder, G. T., & Voss, J. F. (1985). On the ability to predict one's own responses while learning. *Journal of Memory & Language*, **24**, 363-376.

Weaver, C. A., III, & Kelemen, W. L. (1997). Judgments of learning at delays: Shifts in response patterns or increased metamemory accuracy? *Psychological Science*, **8**, 318-321.

Whitten, W. B., & Bjork, R. A. (1977). Learning from tests: Effects of spacing. *Journal of Verbal Learning & Verbal Behavior*, **16**, 465-478.

## NOTES

1. As is now standard in metacognitive research following the suggestion of Nelson (1984), Nelson and Dunlosky (1991) used the Goodman–Kruskal gamma correlation coefficient ($G$) to measure the relative accuracy of the JOLs—that is, the extent to which participants can judge correctly which of two items is more likely to be recalled on the criterion test. The constituent elements for calculating gamma are the number of concordances, in which the higher rated of two items is recalled and the lower rated item is not, and the number of discordances, in which the lower rated of two items is recalled and the higher rated item is not. Specifically, gamma is equal to the ratio of (1) the number of concordances minus the number of discordances, over (2) the number of concordances plus the number of discordances. Gamma has a range of $-1.0$ to $+1.0$, with 0 representing chance performance and $+1.0$ representing perfect accuracy. Gamma also has a probabilistic interpretation: Given that two items are assigned different JOLs and only one of the two is recalled, the probability that the recalled item was given the higher JOL is determined by the equation $P = .5 + .5G$.

2. It might seem odd to consider a gamma correlation between JOLs and final recall when there is an intervening event (post-JOL reexposure) that participants could not be expected to take into account in making JOL assignments. However, we do not claim that the participants should somehow have predicted any effects of post-JOL reexposure on recall. We are instead interested in the predictions made by the memory hypothesis and by metamemory explanations, rather than by the participants, regarding the effects of post-JOL reexposure on gamma and the pattern of final recall and in the consistency of those predictions with the results.

3. On average across items and participants, as more low-JOL items are recalled without any similar increase in recall of high-JOL items, there would be fewer concordances (in which a higher rated item is recalled and a lower rated item is not) and more discordances (in which a lower rated item is recalled and a higher rated item is not). Under these conditions, the numerator of gamma would decrease, inasmuch as it is equal to the number of concordances minus the number of discordances. The denominator of gamma, which is equal to the number of concordances plus the number of discordances, may increase, decrease, or remain unchanged as a result of reexposure, depending on whether the absolute increase in the number of discordances is greater than, less than, or equal to the absolute decrease in the number of concordances, respectively. However, regardless of the change in the denominator, it can be shown algebraically that gamma will decline so long as the number of concordances decreases and the number of discordances increases.

4. Prior to such reexposure, the six pairs were presented in a single array in the same format as at study (cue–target or cue-only) for purposes of selecting items to be presented in a separate, later restudy phase. The restudy phase is not relevant for present purposes; the items that were restudied in that phase were not included in the analyses reported here, and such restudy had no reliable effect on the items that were analyzed.

5. Two additional analyses yielded similar outcomes. (1) A mixed ANOVA using the slopes of the recall functions as the dependent measure, which included all 72 participants and had the advantage that the slopes for individual participants were independent of the portion of the JOL scale they used, revealed that the slope in the delayed cue-only condition was steeper than those in the other conditions, yielding a reliable interaction of JOL format and reexposure [$F(1,70) = 8.07$, $MS_e = 0.01$, $p = .0059$]. (2) Separate mixed ANOVAs for each JOL level, using a more conservative alpha level of .01, indicated that recall did not differ as a function of reexposure or JOL format for the highest four JOL levels, whereas for the three lowest JOL levels, recall in the unreexposed cue-only condition was reliably lower than that in the other three conditions.

6. As in Experiment 2, we performed two additional analyses on the Figure 5 data, and the analyses yielded similar outcomes. (1) A within-subjects ANOVA on the slopes of the recall functions revealed that the slope in the delayed cue-only condition was steeper than those in the other conditions, yielding a reliable target reexposure × JOL timing interaction [$F(1,35) = 6.67$, $MS_e = 0.01$, $p = .0141$]. (2) Dependent-samples $t$ tests of the effect of target reexposure on recall at individual JOL levels for immediate JOLs and delayed JOLs separately yielded results similar to those for the two ANOVAs, except that in the delayed-JOL condition, target reexposure reliably increased recall of items given a JOL of 7.

7. These results also indicate that the recall patterns in the delayed cue–target conditions in Experiment 2 did indeed reflect spacing effects, in that all the items in those conditions received spaced study opportunities (initial study and reexposure at the time of the delayed JOL, with or without post-JOL reexposure), resulting in patterns of recall that were similar to the pattern in the delayed cue-only reexposed condition in both Experiments 2 and 3 (cf. Figures 3 and 5).

8. For this reason, any effects of spacing on such metacognitive processes cannot explain the similarity in the patterns in these conditions, in contrast to the effects of spacing on recallability, which can explain that similarity.

(Manuscript received January 23, 2003;
revision accepted for publication April 10, 2003.)