Journal of Experimental Psychology:
Learning, Memory, and Cognition
2010, Vol. 36, No. 2, 552–557

0278-7393/10/$12.00 DOI: 10.1037/a0018277

Comprehension as a Basis for Metacognitive Judgments: Effects of Effort After Meaning on Recall and Metacognition

Franklin M. Zaromb
Washington University in St. Louis

Jeffrey D. Karpicke
Purdue University

Henry L. Roediger III
Washington University in St. Louis

We examined free recall and metacognitive judgments of ambiguous sentences studied with and without clues to facilitate their comprehension. Sentences were either studied without clues, with clues meaningfully embedded, or with clues following a 10-s interval delay. After presentation, subjects made judgments of comprehension (JCOMPs) or judgments of learning (JOLs). Puzzling over the meaning of sentences for several seconds prior to receiving the clue enhanced recall compared with studying sentences without clues or with embedded clues. This benefit of effort after meaning was not reflected in JCOMPs or JOLs. Rather, sentences considered relatively easy to understand received higher JOLs regardless of experimental condition. Although effort after meaning enhanced recall, subjects displayed no awareness of this benefit in their judgments. Our study adds to a growing literature showing students' ignorance of factors affecting their own learning, which have important implications for education. Making learning conditions more difficult, thus requiring students to engage more cognitive effort, often leads to enhanced retention.

Keywords: memory, learning, comprehension, metacognition, recall

Psychologists have long argued that comprehension is integral for the learning and retention of new information and skills (e.g., Ausubel, 1963; Bartlett, 1932; Katona, 1940). For example, Bartlett (1932) presumed a direct relationship between comprehension and acts of remembering by proposing that individuals engage in "effort after meaning" (p. 44). In encountering the world, people bring their past knowledge to bear on understanding their current situation. Bartlett also argued that when people remember the past, they attempt to make their recollections conform to prior background knowledge, cultural expectations, or the present context.

Several studies have corroborated the notion that a learner's attempts to understand new information directly influence subsequent retention. Auble and Franks (1978) introduced a task to examine the effects of effort after meaning on retention.[1] They had subjects listen to ambiguous sentences, such as "The home was small because the sun came out." Some subjects heard the sentences without any clues to disambiguate the sentences. Subjects in an embedded-clue condition heard the sentences with a clue included within the sentence ("The *igloo* was small because the sun came out"). Finally, subjects in a delayed-clue condition heard the ambiguous sentences and were required to puzzle over the meaning of the sentences for several seconds prior to hearing the disambiguating clues (*igloo*). Auble and Franks showed that making an effort after meaning in the delayed-clue condition enhanced later free recall relative to the other two conditions (also see Auble, Franks, & Soraci, 1979; Zaromb & Roediger, 2009). The memorial advantage of making an effort after meaning also occurs for sentences that are easier to comprehend in the absence of a clue, such as "The engine stopped because the liquid ran out." Although most subjects would not need to hear the clue word (*gasoline*) to comprehend this sentence, recall was still enhanced in a delayed-clue condition relative to an embedded-clue condition (Auble et al., 1979). These and related findings have suggested that processes involved in resolving the ambiguity of a stimulus facilitate learning (Auble et al., 1979; Carlin, Soraci, Dennis, Chechile, & Loiselle, 2001; Wills, Soraci, Chechile, & Taylor, 2000; Zaromb & Roediger, 2009).

In this research we used the effort after meaning paradigm to examine another way that comprehension may directly influence learning. Specifically, we asked whether people rely on their comprehension as a basis for metacognitive judgments and, if so, whether a comprehension-based heuristic is diagnostic of future recall. The two types of metacognitive judgments we investigated were retrospective judgments of comprehension (JCOMPs) and

Franklin M. Zaromb and Henry L. Roediger III, Department of Psychology, Washington University in St. Louis; Jeffrey D. Karpicke, Department of Psychological Sciences, Purdue University.

These data were presented at the Psychonomics Meeting, Chicago, Illinois, November 2008. Support for this research was provided by James S. McDonnell Foundation Collaborative Activity Grant 220020041.

Correspondence concerning this article should be addressed to Franklin M. Zaromb, Department of Psychology, Box 1125, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130-4899. E-mail: fmzaromb@artsci.wustl.edu

[1] Although Auble and Franks (1978) preferred the term *effort toward comprehension*, we use the term *effort after meaning* throughout the article to provide historical continuity and to minimize confusion from switching between the terms.

prospective judgments of learning (JOLs). No studies to date have utilized metacognitive judgments in Auble and Franks's (1978) paradigm, so whether expending effort after meaning will influence JCOMPs and JOLs in relation to memory performance is unknown. Will subjects be aware that expending effort in the delayed-clue conditions enhances their comprehension (as reflected in JCOMPs) or their later recall (reflected in JOLs)?

Assuming effort after meaning does influence processes devoted to stimulus comprehension, JCOMPs during learning may index this extra cognitive work. For instance, prior research has shown that judgments of text comprehension are influenced by the relative ease with which text information is understood during study prior to making a judgment (Rawson & Dunlosky, 2002). In Auble and Franks's (1978) paradigm, sentences that are already embedded with clues should, therefore, receive higher JCOMPs than those studied under conditions in which individuals must puzzle over the meaning of the sentences before hearing the corresponding clues or when no clue is provided, because individuals in these latter two conditions expend greater effort processing the ambiguous verbal information.

Effort after meaning may also be associated with learners' awareness of achieving comprehension, such as having the experience of an "aha!" reaction (Auble et al., 1979). In Auble and Franks's (1978) paradigm, individuals might also develop insight when trying for an extended period of time to figure out the meaning of an ambiguous sentence, the presentation of the clue suddenly enables them to reorient their thinking to achieve comprehension (Luo, Niki, & Phillips, 2004). Thus, sentences that require effort after meaning should receive higher JCOMPs.

We also do not know whether subjects' effort after meaning will be reflected in their predictions of future recall, or JOLs. According to an *ease of processing* view, people believe that if a stimulus is easy to comprehend, it will be easy to remember. Previous research has demonstrated that JOLs may be determined by the relative ease with which to-be-learned items are processed during study prior to making a judgment (e.g., Begg, Duft, Lalonde, Melnick, & Sanvito, 1989), which suggests that sentences more readily understood (those in the embedded-clue condition) should receive higher JOLs than those studied in the delayed- and no-clue conditions.

A more refined view of JOLs is Koriat's (1997) cue-utilization framework, which distinguishes among different types of cues that inform metacognitive judgments. He proposed three classes of cues used for monitoring learning: *intrinsic*, *extrinsic*, and *mnemonic*. Intrinsic cues refer to features of materials that provide information about their a priori ease or difficulty of learning (e.g., word frequency, associative strength). Extrinsic cues refer to encoding conditions that are not intrinsic to the items, such as the use of interactive imagery to study paired associates or repeated study trials. Last, mnemonic cues refer to a subject's internal states, such as having the subjective experience of fluently processing an item. Importantly, the cue-utilization approach predicts that various types of heuristics link the three types of cues to JOLs and that these will vary depending on the experimental conditions (as discussed below). The underlying assumption is that JOLs are comparative in nature, so individuals focus on relative differences in ease of learning among study items while ignoring factors that can influence the overall level of recall.

Using Auble and Franks's (1978) paradigm, we have manipulated both intrinsic and extrinsic cues related to comprehension. Intrinsic cues for comprehension were manipulated by presenting a mixture of sentences that were easy and difficult to understand. This manipulation permitted us to ask whether learners are sensitive to intrinsic features of sentences that determine their comprehensibility. Extrinsic cues for comprehension were manipulated by varying the placement of the disambiguating clues, which varied effort after meaning. Specifically, we used three extrinsic cue conditions: a no-clue condition in which sentences were studied without disambiguating clues, an embedded-clue condition whereby clues were meaningfully embedded within the sentences, and a delayed-clue condition in which clues were heard 10 s following sentence presentation. All independent variables were manipulated within subjects.

If the cue-utilization framework is correct, easy sentences should receive higher JOLs than difficult sentences, but varying the conditions of effort after meaning (the extrinsic cue) should have little or no impact on JOLs. Thus, JOLs should be poor predictors of recall differences between the embedded- and delayed-clue conditions. Recall should be superior in the delayed-clue condition relative to the embedded-clue condition, but subjects should not predict this outcome. JOLs should be relatively good predictors of recall in the no-clue condition, in which comprehension should enhance retention of the easy sentences relative to the difficult, ambiguous sentences.

Method

Subjects

Forty-eight Washington University undergraduate students participated for either course credit or payment.

Materials

Stimuli were 36 sentences and their corresponding clues sampled from the stimuli of Auble et al. (1979) and Zaromb and Roediger (2009), and these are given in those articles. On the basis of ratings obtained in Auble et al.'s study and in our laboratory, 18 sentences were deemed difficult to comprehend (most subjects reported not understanding their meaning independent of the clues). Two examples are as follows: "The boy spilled his popcorn because the lock broke" (clue: *lion cage*), and "The notes were sour because the seam split" (clue: *bagpipe*). Eighteen sentences were easy to comprehend in that most subjects reported being able to understand their meaning independent of the clues. Two examples are as follows: "The colors appeared because the rain stopped" (clue: *rainbow*), and "The teacher stopped writing because the stick broke" (clue: *chalk*). The sentences and clues were voice-recorded on an IBM computer, and E-Prime experimental software was used for stimulus presentation and recording keyboard responses.

Design

The experiment used a 3 (Learning Condition: No-Clue, Embedded-Clue, Delayed-Clue) × 2 (Sentence Type: Easy vs. Difficult) × 2 (Judgment Type: JCOMP vs. JOL) mixed factorial design. The three learning conditions and sentence types were manipulated within subjects during the study phase as follows. Twelve sentences were presented without clues (no-clue condition). Twelve sentences had embedded clues (embedded-clue condition). Last, 12 sentences were followed by clues after a 10-s interval (delayed-clue condition). Half of the sentences in each learning condition were easy, and the other half were difficult to comprehend. The order of sentence presentation and learning condition was randomized with the constraint that across subjects, each sentence was assigned an equal number of times to each of the three conditions. Last, judgment type was manipulated between subjects such that half of the subjects made JCOMPs and half made JOLs for each sentence during the study phase.

Procedure

The experimenter initially instructed subjects that they would study and later be asked to recall a series of sentences and that additional clue words might be heard to facilitate sentence comprehension. The total study time for each sentence was 20 s, which included 5 s for sentence presentation and 2 s for clue presentation. Following auditory presentation of each sentence and the possible clue, half of the subjects were prompted by a computer display and allowed 10 s to make a JCOMP by indicating how well they understood the sentence by typing a number between 0 and 100, in which 0 means that the sentence was not understood at all, 100 indicates perfect comprehension, and values in between reflect intermediate levels of comprehension. The remaining 24 subjects were asked to make a JOL by indicating how likely they were to recall the sentence at a later point during the experimental session, again by typing a number between 0 and 100, in which 0 means that there is no chance they will recall the sentence, 100 means that they are perfectly confident they will recall it, and numbers in between reflect intermediate levels of confidence. Subjects were encouraged to use the entire scale as best as possible.

Following a 5-min number matching distracter task, subjects attempted to type as many and as much of the sentences as they could remember in any order into a blank Microsoft Word document for 10 min. The experimental session lasted about an hour.

Results

All results, unless otherwise stated, are significant at the .05 level. For all sets of individual comparisons, we controlled the Type I error rate at .05 using the False Discovery Rate procedure (Benjamini & Hochberg, 1995; Benjamini & Yekutieli, 2001).

Sentence Recall

As shown in the top row of Table 1, the proportion of sentences correctly recalled for each subject was highest in the delayed-clue condition, followed by the no-clue and embedded-clue conditions, which did not differ. We conducted a 3 (Learning Condition: No-Clue vs. Embedded-Clue vs. Delayed-Clue) × 2 (Sentence Type: Easy vs. Difficult) analysis of variance (ANOVA), which revealed a significant effect of learning condition, $F(2, 92) = 56.60$, $MSE = 0.03$, $\eta_p^2 = .55$. Consistent with prior studies, recall was enhanced in the delayed-clue condition relative to the no-clue condition (0.47 vs. 0.23), $t(47) = 8.26$, $SEM = 0.03$, $d = 1.65$, and the embedded-clue condition (0.47 vs. 0.22), $t(47) = 7.78$, $SEM = 0.03$, $d = 1.61$, but with no difference between the latter two conditions (0.23 vs. 0.22, $t < 1$). Thus, retention was significantly affected by effort after meaning—an extrinsic factor (Auble & Franks, 1978; Auble et al., 1979; Zaromb & Roediger, 2009).

Although there was no main effect of sentence type ($F < 1$), there was a significant Learning Condition × Sentence Type interaction, $F(2, 92) = 4.60$, $MSE = 0.04$, $\eta_p^2 = .09$, because of enhanced recall in the no-clue condition for easy compared with difficult sentences (0.28 vs. 0.18), $t(47) = 2.93$, $SEM = 0.03$, $d = 0.57$. Individual pairwise comparisons between easy and difficult sentences were not significant in the embedded-clue condition (0.21 vs. 0.24, $t < 1$) or the delayed-clue condition (0.44 vs. 0.49), $t(47) = 1.20$, $SEM = 0.04$, *ns*, $d = 0.24$. Thus, the intrinsic factor of sentence difficulty influenced retention in the no-clue condition, but no such effect occurred in the embedded-clue and delayed-clue conditions, in which provision of clues made the sentences easier to understand.

JCOMPs

As shown in the second row of Table 1, mean JCOMPs were greater, on average, for sentences that were relatively easy to understand. They were also similarly high for sentences in the delayed-clue and embedded-clue conditions, and they were poorest in the no-clue condition. In examining the effects of extrinsic and intrinsic manipulations on JCOMPs, the ANOVA revealed a significant effect of learning condition, $F(2, 46) = 51.80$, $MSE = 0.03$, $\eta_p^2 = .69$, with higher JCOMPs in the delayed-clue and

Table 1
Proportion of Easy and Hard to Comprehend Sentences Correctly Recalled, Mean Judgments of Comprehension (JCOMPs), and Mean Judgments of Learning (JOLs) in the No-Clue, Embedded-Clue, and Delayed-Clue Presentation Conditions

	No-clue		Embedded-clue		Delayed-clue	
Measure	Easy	Hard	Easy	Hard	Easy	Hard
Free recall	0.28 (0.06)	0.18 (0.04)	0.21 (0.06)	0.24 (0.06)	0.44 (0.06)	0.49 (0.06)
JCOMPs	0.79 (0.12)	0.38 (0.14)	0.91 (0.07)	0.82 (0.09)	0.93 (0.12)	0.83 (0.07)
JOLs	0.50 (0.10)	0.35 (0.11)	0.51 (0.11)	0.45 (0.12)	0.55 (0.11)	0.48 (0.11)

Note. Values in parentheses are 95% confidence intervals calculated according to the method of Loftus and Masson (1994). JCOMPs and JOLs were converted from whole number values to proportions.

embedded-clue conditions relative to the no-clue condition: delayed-clue condition (0.86 vs. 0.58), $t(23) = 7.92$, $SEM = 0.04$, $d = 1.64$; embedded-clue condition (0.88 vs. 0.58), $t(23) = 9.40$, $SEM = 0.03$, $d = 1.89$. However, there was no difference between the former conditions (0.86 vs. 0.88, $t < 1$). The similar results in the delayed-clue and embedded-clue conditions suggest that, unlike recall performance, JCOMPs were not sensitive to variations of the extrinsic manipulation of effort after meaning.

By contrast, the intrinsic factor of sentence difficulty did significantly influence JCOMPs, as easy sentences received higher JCOMPs than difficult sentences overall (0.87 vs. 0.68), $F(1, 23) = 118.56$, $MSE = 0.01$, $\eta_p^2 = .84$. This outcome occurred in each learning condition: no-clue condition, $t(23) = 8.62$, $SEM = 0.05$, $d = 1.80$; embedded-clue condition, $t(23) = 3.65$, $SEM = 0.02$, $d = 0.55$; and delayed-clue condition, $t(23) = 3.96$, $SEM = 0.03$, $d = 0.74$. The ANOVA further revealed a significant Learning Condition × Sentence Type interaction, $F(2, 46) = 25.16$, $MSE = 0.01$, $\eta_p^2 = .52$, which was due to the greater difference between easy and difficult sentences in the no-clue (0.40) relative to the embedded-clue (0.09) and delayed-clue (0.10) conditions. Although provision of the clues made the sentences easier to understand and reduced the gap between the sentence types, JCOMPs were still influenced by the intrinsic factor of sentence difficulty. In sum, whereas the extrinsic manipulation of effort after meaning significantly affected recall but not JCOMPs, the intrinsic factor of sentence difficulty influenced JCOMPs but not recall.

JOLs

As shown in the third row of Table 1, JOLs were greater, on average, for sentences that were relatively easy to understand, and the pattern of data was similar to that for JCOMPs. JOLs were similarly high for sentences in the delayed-clue and embedded-clue conditions, and they were poorest in the no-clue condition. In examining the effects of extrinsic and intrinsic manipulations on JOLs, the ANOVA confirmed a significant effect of learning condition, $F(2, 46) = 14.62$, $MSE = 0.01$, $\eta_p^2 = .39$, because of higher JOLs in the delayed-clue and embedded-clue conditions relative to the no-clue condition: delayed-clue condition (0.52 vs. 0.42), $t(23) = 4.75$, $SEM = 0.02$, $d = 0.64$; embedded-clue condition (0.48 vs. 0.42), $t(23) = 3.60$, $SEM = 0.02$, $d = 0.39$. However, no significant difference emerged between the former two conditions (0.52 vs. 0.48), $t(23) = 2.29$, $SEM = 0.02$, $d = 0.25$, *ns*. Again, the similar results in the delayed-clue and embedded-clue conditions indicate that JOLs were not affected by effort after meaning.

By contrast, the intrinsic factor of sentence difficulty did significantly influence JOLs, with easy sentences receiving higher JOLs than difficult sentences overall (0.52 vs. 0.43), $F(1, 23) = 34.18$, $MSE = 0.01$, $\eta_p^2 = .60$. Again, this outcome occurred in each learning condition: no-clue condition, $t(23) = 7.12$, $SEM = 0.02$, $d = 0.94$; embedded-clue condition, $t(23) = 2.75$, $SEM = 0.02$, $d = 0.35$; and delayed-clue condition, $t(23) = 3.59$, $SEM = 0.02$, $d = 0.41$. The ANOVA further revealed a significant Learning Condition × Sentence Type interaction, $F(2, 46) = 7.35$, $MSE < 0.01$, $\eta_p^2 = .24$, which was due to the greater difference between easy and difficult sentences in the no-clue (0.15) relative to the embedded-clue (0.06) and delayed-clue (0.07) conditions. Again, as with JCOMPs, the provision of clues made the sentences easier, which in turn raised JOLs and eliminated the difference between hard and easy sentences. Thus, whereas the extrinsic factor of effort after meaning significantly affected recall but not JOLs, the intrinsic variable of sentence difficulty influenced JOLs but not recall performance.

Correlations

To directly examine whether JCOMPs and JOLs were influenced by similar factors and whether they were equally accurate at predicting recall, we needed to average data across subjects for each of the 36 study sentences.[2] As shown in Table 2, in the no-clue condition, when sentences varied in terms of their intrinsic comprehensibility, JCOMPs were strongly correlated with JOLs ($r = .70$, $p < .01$), and both types of judgments were positively correlated with recall performance (JCOMPs: $r = .49$, $p < .01$; JOLs: $r = .54$, $p < .01$). By contrast, metacognitive judgments and recall were uncorrelated with one another in the embedded-clue condition, perhaps because sentences were generally understood, and intrinsic variations in sentence difficulty may have been less salient. However, there were mixed results in the delayed-clue condition, with JCOMPs being positively correlated with JOLs ($r = .49$, $p < .01$), but neither judgment was predictive of recall performance. Perhaps by studying the sentences alone for 10 s prior to hearing the disambiguating clues in the delayed-clue condition, subjects based their metacognitive judgments on intrinsic sentence difficulty and discounted the positive effect of the delayed clues on retention.

Discussion

Our experiment demonstrated that requiring subjects to puzzle over the meaning of sentences for several seconds prior to comprehension enhanced recall relative to conditions in which the sentences were presented with an embedded-clue (making them readily understood) or a no-clue condition (in which the difficult sentences were generally not understood). Moreover, in the embedded-clue and delayed-clue conditions, sentences that were considered relatively easy to understand were no better recalled than sentences considered ambiguous. These data are consistent with prior studies and corroborate the notion of a memorial advantage for what Bartlett (1932) termed *effort after meaning* (Auble & Franks, 1978; Auble et al., 1979; Zaromb & Roediger, 2009).

However, these effects on recall were not predicted by metacognitive JCOMPs and JOLs. Although the delayed-clue condition produced the highest level of recall, subjects were not aware of this benefit; JOLs and JCOMPs in the delayed-clue condition were similar to those in the embedded-clue condition. Moreover, sentences considered relatively easy to understand received higher judgments than sentences perceived to be ambiguous and difficult to understand, but no difference existed in recall between the two

[2] It was not possible to compare within-subject correlations between JCOMPs and JOLs because the judgments were assessed in different groups of subjects. Also, insufficient data were collected in each condition (because of no sentences being recalled by some subjects in some conditions) to report within-subject correlations between metacognitive judgments and recall performance.

Table 2
Correlations Across the 36 Sentences Between Mean Judgments of Comprehension (JCOMPs), Judgments of Learning (JOLs), and Recall Performance (Rec.) in the No-Clue, Embedded-Clue, and Delayed-Clue Presentation Conditions

	No-clue			Embedded-clue			Delayed-clue		
Measure	JCOMP	JOL	Rec.	JCOMP	JOL	Rec.	JCOMP	JOL	Rec.
JCOMP	—	.70*	.49*	—	.21	.07	—	.49*	−.06
JOL	.70*	—	.54*	.21	—	−.19	.49*	—	.25
Recall	.49*	.54*	—	.07	−.19	—	−.06	.25	—

* Significant at the $p < .01$ level.

types of sentences when the sentences had been presented with clues. By contrast, when the sentences were studied alone in the no-clue condition, JCOMPs and JOLs were highly correlated and served as moderately good predictors of recall. Taken together, these findings indicate that JOLs were influenced by the intrinsic manipulation of difficulty of the individual sentences but not by extrinsic variations in the learning conditions of sentence and clue presentation (effort after meaning).

These findings are surprising for several reasons. First, prior findings have shown that JCOMPs are influenced by the relative ease with which study materials are understood (e.g., Rawson & Dunlosky, 2002), so subjects should have reported the highest JCOMPs in the embedded-clue condition. This did not occur. On the other hand, if effort after meaning is associated with the unique subjective experience or heightened awareness of achieving comprehension, subjects should have reported the highest JCOMPs in the delayed-clue condition. Instead, JCOMPs were similar in both the delayed-clue and embedded-clue conditions. The current findings also violate predictions derived from the ease of processing account of JOLs (Begg et al., 1989). This view predicts that sentences processed more fluently (those in the embedded-clue condition) should achieve greater comprehension and thereby receive higher JOLs than those studied in the delayed-clue-condition JOLs. However, the predictions were similar in the two conditions.

On the other hand, Koriat's (1997) cue-utilization framework does account well for these data. If varying the relative comprehensibility of the sentences represents an intrinsic cue and manipulating the conditions of clue presentation (effort after meaning) represents an extrinsic cue, the cue-utilization approach predicts that effort after meaning should have little or no impact on JOLs. Consistent with this prediction, JOLs were similar in the delayed-clue and embedded-clue conditions even though recall differed substantially between these conditions. Second, just as intrinsic cues were expected to significantly influence metacognitive judgments, JOLs in the no-clue condition in which sentences were poorly understood were lower than those reported in the embedded-clue and delayed-clue conditions. More importantly, JOLs were higher for sentences considered relatively easy to understand across all learning conditions.

As successful as the cue-utilization framework is in explaining these data, it is not without limitations (Dunlosky & Matvey, 2001). First, despite our correlational data showing a strong relation between JOLs and JCOMPs, it remains unclear what specific features determine the ease or difficulty of sentence comprehension; it is not clear, either, how sentence difficulty relates to other normative measures (e.g., ease of learning judgments, reading fluency, word frequency). Second, by manipulating intrinsic and extrinsic cues within subjects, we afforded individuals the opportunity to base their JOLs on comparisons among different sentence types, learning conditions, or both; nonetheless, only the intrinsic factor affected JOLs.

Prior research has shown that learners discount extrinsic cues relative to intrinsic cues even when they are both distributed within subjects (e.g., Koriat, 1997; Rabinowitz, Ackerman, Craik, & Hinchley, 1982); however, it is still unclear under what conditions and why intrinsic cues outweigh extrinsic cues (see Dunlosky & Matvey, 2001). One possibility is that the intrinsic cue of sentence difficulty is so salient that it "overshadows" the extrinsic cue of effort after meaning (e.g., Price & Yates, 1993). Another possible explanation is that individuals utilize the intrinsic cue of sentence difficulty as an analytic heuristic, inferring that memory performance should be better for easier to understand sentences. Alternatively, individuals may treat the intrinsic cue of sentence difficulty as a nonanalytic heuristic, relying on their extensive prior experience of processing normal verbal discourse over their limited experience of learning ambiguous discourse with delayed clues. Future researchers may help to distinguish between the alternative explanations by, for instance, using multiple study and test trials to determine whether individuals shift from reliance on an analytic (theory-based) heuristic to a nonanalytic (experience-based) heuristic as they gain experience learning and attempting to remember the different types of sentence materials under conditions that vary effort after meaning (Koriat, 1997).

The present findings also provide a unique demonstration of how two different types of metacognitive judgments—retrospective JCOMPs and prospective JOLs—are similar in terms of their sensitivity to the same cues. These two types of judgments have primarily been studied with different paradigms and measures. JOLs have been the focus of metamemory studies in which subjects typically study and later attempt to remember simple verbal materials, such as word pair associates (e.g., Bjork, 1994; Nelson & Narens, 1990; for a recent review, see Dunlosky & Metcalfe, 2009). JCOMPs have been the focus of metacomprehension studies that have investigated how people assess their comprehension of texts (e.g., Dunlosky & Lipko, 2007; Glenberg & Epstein, 1985; Maki & Berry, 1984). One enigma is that findings in the metamemory literature seldom generalize to studies of metacomprehension (Dunlosky & Lipko, 2007; Dunlosky & Metcalfe, 2009).

By contrast, our findings demonstrate important similarities between JCOMPs and JOLs in Auble and Franks's (1978) paradigm where they can be directly compared. First, both types of judgments were sensitive to the relative comprehensibility of the study sentences but were insensitive to variations in the learning conditions. Second, both judgments were moderate predictors of recall in the no-clue condition in which comprehension enhanced retention of the easy sentences relative to the difficult sentences. Moreover, JCOMPs and JOLs were poor predictors of recall in the embedded-clue and delayed-clue conditions, in which effort after meaning greatly influenced retention but did not influence metacognitive judgments. The only apparent difference between the two types of judgments was that JCOMPs were greater in magnitude than JOLs. This outcome makes sense; subjects might understand a sentence without necessarily believing that they can remember it later. Stated another way, subjects may use higher criteria (but similar processes) when making predictions of recall than postdictions of comprehension.

In sum, although puzzling over the meaning of ambiguous verbal materials for a short period prior to achieving comprehension can facilitate learning, students are not aware of this benefit. Instead, JCOMPs and JOLs are primarily affected by the inherent comprehensibility of the materials alone. These findings add to the body of literature showing metacognitive illusions from which students suffer (Karpicke, 2009; Koriat & Bjork, 2005; Metcalfe & Finn, 2008). From an educational standpoint, these findings add further evidence that making learning conditions challenging for students—in this case, forcing them to engage in an effort to discover meaning of difficult material—aids their retention even though they are unaware of this fact during learning. Bjork (1994) has pointed to the need to introduce desirable difficulties into learning opportunities, and our results show that requiring effort after meaning is one such desirable difficulty. For example, posing difficult questions during lectures or texts, and having students ponder them before providing an answer, should aid learning and retention.

References

Auble, P. M., & Franks, J. J. (1978). The effects of effort toward comprehension on recall. *Memory & Cognition, 6,* 20–25.

Auble, P. M., Franks, J. J., & Soraci, S. A. (1979). Effort toward comprehension: Elaboration or "aha!"? *Memory & Cognition, 7,* 426–434.

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning: An introduction to school learning.* New York, NY: Grune & Stratton.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge, England: Cambridge University Press.

Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language, 28,* 610–632.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B, 57,* 289–300.

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics, 29,* 1165–1188.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–206). Cambridge, MA: MIT Press.

Carlin, M. T., Soraci, S. A., Dennis, N. A., Chechile, N. A., & Loiselle, R. C. (2001). Enhancing free-recall rates of individuals with mental retardation. *American Journal on Mental Retardation, 106,* 314–326.

Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science, 16,* 228–232.

Dunlosky, J., & Matvey, G. (2001). Empirical analysis of the intrinsic–extrinsic distinction of judgments of learning (JOLs): Effects of relatedness and serial position on JOLs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1180–1191.

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition.* Los Angeles, CA: Sage.

Glenberg, A. M., & Epstein, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 702–718.

Karpicke, J. D. (2009). Metacognitive control and strategy selection: Deciding to practice retrieval during learning. *Journal of Experimental Psychology: General, 138,* 469–486.

Katona, G. (1940). *Organizing and memorizing: Studies in the psychology of learning and teaching.* New York, NY: Hafner Publishing.

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.

Koriat, A., & Bjork, R. A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 187–194.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subjects designs. *Psychonomic Bulletin & Review, 1,* 476–490.

Luo, J., Niki, K., & Phillips, S. (2004). Neural correlates of the "Aha! reaction." *NeuroReport, 15,* 2013–2017.

Maki, R. H., & Berry, S. L. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 663–679.

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin Review, 15,* 174–179.

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. 125–173). New York, NY: Academic Press.

Price, P. C., & Yates, J. F. (1993). Judgmental overshadowing: Further evidence of cue interaction in contingency judgment. *Memory & Cognition, 21,* 561–572.

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.

Rawson, K. A., & Dunlosky, J. (2002). Are performance predictions for text based on ease of processing? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 69–80.

Wills, T. W., Soraci, S. A., Chechile, R. A., & Taylor, H. A. (2000). "Aha" effects in the generation of pictures. *Memory & Cognition, 28,* 939–948.

Zaromb, F. M., & Roediger, H. L. (2009). The effects of "effort after meaning" on recall: Differences in within- and between-subjects designs. *Memory & Cognition, 37,* 447–463.

Received August 30, 2009
Revision received October 19, 2009
Accepted October 21, 2009 ■