Psychology and Aging
2006, Vol. 21, No. 2, 390–400

0882-7974/06/$12.00 DOI: 10.1037/0882-7974.21.2.390

Does Aging Influence People's Metacomprehension? Effects of Processing Ease on Judgments of Text Learning

John Dunlosky, Julie M. C. Baker, and Katherine A. Rawson
Kent State University

Christopher Hertzog
Georgia Institute of Technology

In 2 experiments, the authors investigated whether age-related differences exist in metacomprehension by evaluating predictions based on the ease-of-processing (EOP) hypothesis. According to this hypothesis, judgments of how well a text has been learned are based on how easily each text was processed; easier processing results in higher judgments. Participants read either sentence pairs or longer texts and judged their learning of each immediately afterward. Although an age-related difference in the use of processing ease in judgments was observed with sentence pairs, for longer texts older and younger adults' judgments were similarly related to processing ease. In both experiments, age equivalence was also evident in the accuracy of the judgments at predicting performance on the criterion test. The overall pattern of results suggests that judging text learning remains largely intact with aging.

Keywords: metacomprehension, ease of processing, fluency, judgments

Throughout the life span, people often need to comprehend and remember content from text materials covering innumerable topics, including, for example, labels on medicine bottles and food packages, medical instructions from doctors, newspaper articles, training manuals in occupational settings, and textbook materials in educational environments. Our interest here concerns *metacomprehension*, which involves how people judge their comprehension and learning of text materials.[1] Accurate metacomprehension itself can be valuable for fostering comprehension and memory of text materials (for a review, see Dunlosky, Hertzog, Kennedy, & Thiede, 2005). For instance, an individual who accurately judges that important text materials have not been learned well can selectively devote extra time to (or seek help with) just those materials.

This function of metacomprehension for effective learning prompts our main question: Does aging influence people's metacomprehension? Our review of the literature revealed that little is known about whether aging influences (a) how people judge the learning of text material and (b) the accuracy of these judgments. Our primary goal with the present research was to estimate the possible influences of aging on judgments of text learning, which was accomplished by empirically evaluating the ease-of-processing (EOP) hypothesis. A secondary goal was to evaluate whether age-related deficits occur in the accuracy of people's judgments of text learning, which bears on whether aging diminishes metacomprehension skills.

John Dunlosky, Julie M. C. Baker, and Katherine A. Rawson, Psychology Department, Kent State University; Christopher Hertzog, School of Psychology, Georgia Institute of Technology.

This research was supported by National Institute on Aging Grant R37 AG13148.

Correspondence concerning this article should be addressed to John Dunlosky, P. O. Box 5190, Psychology Department, Kent State University, Kent, OH 44242. E-mail: jdunlosk@kent.edu

The EOP Hypothesis

The EOP hypothesis was first proposed by Begg, Duft, Lalonde, Melnick, and Sanvito (1989) to account for how people judge their learning of simple materials, such as paired associates. According to this hypothesis, the ease with which people process materials during study influences their judgments of how well they will remember them. More specifically, people are expected to judge that they will better remember items that were easier to process during study than those that were more difficult to process.[2] To demonstrate EOP effects, Begg et al. (1989) had college students study individual words and make a judgment of learning (a confidence rating for the likelihood of later recall) for each word immediately after it had been studied. The list was composed of high-frequency and low-frequency words. As expected, the judg-

[1] Although it is important to distinguish between memory and comprehension of text materials (Kintsch, 1994), the bulk of the literature on metacomprehension has involved investigating how well people can predict their performance on a criterion test over text materials, criterion tests that often in part (if not entirely) tap memory for the texts. Several authors have argued that the distinction between memory and comprehension should also be integrated into theory and research on metacomprehension (Rawson, Dunlosky, & Thiede, 2000; Wiley, Griffin, & Thiede, 2005). In the present research, however, we investigated how older and younger adults predict their performance on a criterion test over text materials, which is also how other researchers examining cognitive aging have explored metacomprehension (Miles & Stine-Morrow, 2004; Olin & Zelinski, 1997). Thus, for historical continuity, we retain the term *metacomprehension* in general because our research focuses on metacognition for text materials, but we refer to the judgments made by participants as judgments of text learning.

[2] Processing ease is often not predictive of how well items will be remembered on a criterion test, so judgments based on processing ease may not be highly accurate at predicting test performance (e.g., Benjamin, Bjork, & Schwartz, 1998; Rawson & Dunlosky, 2002), which foreshadows results presented in both experiments of the present article.

ments were higher for high-frequency words than for low-frequency words, presumably because the former were easier to process at study. More recently, instead of merely inferring differences in processing ease from item manipulations (e.g., word frequency), investigators have begun to test the EOP hypothesis by measuring processing ease more directly (e.g., Benjamin et al., 1998; Hertzog, Dunlosky, Robinson, & Kidder, 2003; Koriat & Ma'ayan, 2005; Matvey, Dunlosky, & Guttentag, 2001). Consider the procedure used by Hertzog et al. (2003), which is most relevant to the methods adopted in the present experiments. In their experiments, participants studied paired associates (e.g., turtle–bean) and were instructed to generate an interactive image to associate the words within each pair (e.g., a turtle surfing on an oversized bean). While studying a given item, participants were instructed to press a key once they had generated an interactive image. Ease of processing was operationalized as the time from the onset of an item to a participant's keypress indicating that an image had been successfully generated. After studying each item, a judgment of learning was made. According to the EOP hypothesis, as the time to generate images decreases (indicating easier processing), participants' judgments of the likelihood of subsequent recall should increase. Consistent with this prediction, intraindividual correlations between the latency of imagery generation and the judgments were consistently negative.

Although ease of processing apparently influences judgments of learning for simple materials, current evidence is less definitive with respect to the role of processing ease in judgments of text learning. Inconsistent with the EOP hypothesis, Maki, Foley, Kajer, Thompson, and Willert (1990) reported that a manipulation that should influence the EOP texts (deleting letters from many words in some texts) had no influence on people's judgments of text learning. In contrast, Rawson and Dunlosky (2002) found that several manipulations (including deleting letters from words) did influence the judgments as expected (see also Miles & Stine-Morrow, 2004; Olin & Zelinski, 1997, described in detail later). More important, because these investigations did not include measures of processing ease, their outcomes are open to alternative interpretations. In particular, Koriat (1997) proposed that people construct judgments of learning in part based on an explicit analysis of item characteristics (e.g., concreteness). In the present case, the analytic use of text characteristics may also partly account for the aforementioned EOP effects. For instance, consider the finding that younger adults' judgments of text learning are higher for intact text than for text containing words with letters deleted (Rawson & Dunlosky, 2002). Because the latter would presumably be more difficult to process, this effect was interpreted as consistent with the EOP hypothesis. Nevertheless, the judgments may be a product of people's explicit belief that deleting letters from words reduces memory for those texts. Put differently, perhaps effects that had previously been viewed as evidence for the EOP hypothesis actually reflect people's explicit analysis of text characteristics and not processing ease per se.

To rule out this possibility, one needs to measure how fluently a given participant processes each text. By doing so, the effects of actual processing ease can be statistically estimated independently of any text manipulation. In both experiments reported here, we have adapted measures of processing ease from previous research (Hertzog et al., 2003; Karabenick, 1996). Our primary question is, Do age-related differences occur in the degree to which the ease of text processing influences people's judgments of text learning?

Aging and Metacomprehension

One answer to this question is informed by evidence from research that involves judgments of learning for paired-associate items. This research has consistently demonstrated that aging has little influence on how cues are incorporated into people's judgments of learning (for a review, see Hertzog & Hultsch, 2000; for a possible exception, see Souchay, Isingrini, & Espagnet, 2000). Most relevant, Robinson, Hertzog, and Dunlosky (in press) evaluated the EOP hypothesis for both younger and older adults using Hertzog et al.'s (2003) method, which involved having people generate interactive images during study and press a key once an image had been formed. Consistent with the EOP hypothesis, judgments of learning for both age groups were negatively related to the speed of generating images during study. Note, however, that conclusions from the paired-associate literature have not always generalized to the literature on judgments of text learning (e.g., Dunlosky, Rawson, & Middleton, 2005; Maki, 1998a; Matvey et al., 2001; Nelson & Dunlosky, 1991; Rawson & Dunlosky, 2002), indicating that caution should be used when drawing conclusions about one content domain based on findings from the other.

In summary, the question remains open as to whether the intact monitoring skills that older adults demonstrate in associative learning will generalize to text learning. One might expect that age-related differences will occur in the degree to which processing ease influences judgments of text learning. In particular, when participants judge their learning of a single paired associate, the ease of processing that pair would be readily accessible from working memory, and hence it would be available as a cue for both age groups. By contrast, text processing can be disrupted in multiple ways, such as when a reader encounters novel words, parses sentences with complex syntactic structures, or attempts to determine the referent of an ambiguous pronoun. Because judgments of text learning are usually made at the end of a passage break, further reading would usually intervene between any processing disruption and when the judgment is made after reading is complete. Hence, when the judgment is made, memories of the specific processing disruptions may be less accessible to older than younger adults. If so, the possible influence of ease of processing on judgments will be diminished for older adults compared with younger adults (cf. Lefevre & Lories, 2004).

Two investigations bear on this possibility by providing empirical tests of the EOP hypothesis for text processing by older and younger adults. First, Olin and Zelinski (1997) had older and younger adults read multiple paragraphs. After reading each text, participants made a variety of judgments for each paragraph, including (a) one about their certainty of correctly answering questions about the just-read paragraph and (b) another of processing ease, which was prompted with the question, "How easy was the above paragraph to understand?" After reading and judging the paragraphs, a recognition test was used to measure test performance across the paragraphs. Olin and Zelinski did not report correlations between the two judgments, but they did correlate each participant's judgments with test performance. The correlations for both age groups did not differ, which suggests that age

has a minimal influence on how the judgments were constructed and that processing ease was used similarly by both groups.[3] Nevertheless, such a comparison of correlations may have limited relevance to how judgments are made, because judgment–test correlations can be influenced by factors other than the differential use of ease of processing (e.g., Schwartz & Metcalfe, 1994; Thiede & Dunlosky, 1994; Weaver, 1990). For instance, even if two groups use completely different cues when making their judgments, the same level of accuracy may be obtained if the cues used by both have similar diagnosticity for predicting test performance.

Second, Miles and Stine-Morrow (2004) had younger and older adults read 36 short sentences (18 words each). After reading each one, participants predicted how much they would remember from the sentence. Most relevant for the present purposes, the sentences varied in propositional density: low density (4–5 propositions), medium density (8 propositions), and high density (9–10 propositions). Assuming that processing ease is inversely related to propositional density, the EOP hypothesis predicts that the judgments will be highest for low-density sentences and lowest for high-density sentences. Consistent with this prediction, both age groups' judgments were 6% higher for low-density sentences than for high-density sentences, and the judgments for the medium-density sentences fell in between the mean judgments for these two. Note, however, that Miles and Stine-Morrow's (2004) research was not specifically aimed at evaluating the EOP hypothesis, so processing ease was not measured, leaving their outcomes open to alternative interpretations, as described previously. Thus, although both studies (Miles & Stine-Morrow, 2004; Olin & Zelinski, 1997) provide preliminary evidence suggesting that aging does not influence the use of ease of processing as a cue for judgments of text learning, further evidence involving direct measurement of ease of processing during reading is required to more fully test the EOP hypothesis.

Evaluating the EOP Hypothesis for Older and Younger Adults

A primary aim of the present research was to estimate more directly the effect of processing ease on older adults' and younger adults' judgments of text learning. To do so, we manipulated text characteristics that presumably influence processing ease, which in turn were expected to influence the judgments. In Experiment 1, causal coherence was manipulated across two-sentence sets. In Experiment 2, coherence was manipulated by revising sentences to increase the difficulty of establishing connections between sentences within lengthy texts. Critically, in both experiments, we also obtained an online measure of processing ease while participants read the texts, which allowed us to directly estimate the influence of processing ease on the judgments.

Based on the present design, multiple predictions were evaluated. First, according to the EOP hypothesis, the judgments will be greater for texts that are more versus less coherent. Second, the measures of processing ease will be significantly correlated with the judgments in a manner indicating that easier processing is related to higher judgments. Third, this relationship will remain even when the correlations between processing ease and judgments are computed with coherence level partialed out. This outcome is critical in demonstrating that ease of processing is still influential when participants cannot be basing their judgments solely on the manipulated differences in text characteristics. If aging negatively influences the use of ease of processing in the construction of judgments of text learning, these effects will be smaller (or even nonexistent) for older adults compared with younger adults.

Experiment 1

We manipulated the causal coherence between the two sentences of each set (adapted from Myers, Shinjo, & Duffy, 1987). For example, causal coherence in the following pair is relatively high: "Nancy drank too much beer at a party yesterday. She suffered from a bad hangover this morning." Moderate- and low-coherence versions were created for each pair by writing alternative first sentences, in this case "Nancy had a good time at a party yesterday" (moderate causal relatedness) and "Nancy finished her work quite early on Friday" (low causal relatedness). As coherence decreases, readers need to engage in effortful processing to establish causal coherence (Fletcher, Hummel, & Marsolek, 1990; Myers et al., 1987). That is, as causal coherence decreases, so does processing ease. Thus, according to the EOP hypothesis, the judgments were expected to decrease monotonically with decreases in causal coherence.

To test this prediction, participants read eight different pairs at each level of causal coherence and were instructed to infer how the two sentences are related. Processing ease was then operationalized by two measures: (a) generation success (i.e., whether or not an inference was successfully generated, with successful generation reflecting greater processing ease), and (b) generation latency (i.e., the time between the presentation of the second sentence of a pair and a participant's keypress to indicate an inference had been successfully generated, with shorter times reflecting greater ease at inference generation). According to the EOP hypothesis, the judgments will be positively related to generation success and negatively related to inference latency.

Method

Participants

Thirty-five older ($M = 63.3$ years, $SD = 12.9$) and 35 younger ($M = 19.7$, $SD = 4.7$) adults participated in this study. Older adults were recruited from the community by newspaper ads and received $20 for participating. Younger adults were recruited from the department subject pool and received credit toward a course requirement in introductory psychology. The older adults were well educated ($M = 15.8$ years of education, $SD = 2.5$). Older adults reported taking more medications ($M = 2.6$, $SD = 1.9$) than younger adults ($M = 0.7$, $SD = 0.9$), $t(68) = 5.34$, $p < .05$. However, on a scale ranging from 1 (*excellent*) to 4 (*poor*), both older adults ($M = 1.5$, $SD = 0.6$) and younger adults ($M = 1.4$, $SD = 0.6$) rated themselves as being in quite good health.

[3] This interpretation is not the one offered by Olin and Zelinski (1997), who concluded that older adults rely more on ease of processing than do younger adults. Given the indirect nature of their analyses and the initially low levels of the single-order correlations, their outcomes support multiple interpretations. Most important, they did not report any evidence consistent with the hypothesis of age deficits in the use of ease of processing for making judgments of text learning.

Materials

Items consisted of causally related pairs of sentences taken from Myers et al. (1987). The same 24 target sentences used by Rawson and Dunlosky (2002) were used in this experiment. For each target sentence, there were three potential preceding sentences that varied in their causal relatedness to the target: low, moderate, and high. Each participant had eight pairs of sentences from each level of causal relatedness. The Appendix provides another example of the different levels of causal relatedness.

Procedure

Participants were instructed that they would be presented the first and second sentences for each pair (one at a time, the first immediately before the second) and were told that they would later be asked to recall the second sentence when shown the first sentence from each pair. They were also instructed to attempt to generate an inference that would make sense of the two sentences: "During the time while the second sentence is presented on the screen, press the spacebar as soon as you think that you understand how that sentence is related to the first sentence of the pair." They were informed that pressing the spacebar would not change the screen, but the computer would record that they had been successful in generating an inference. They were further informed that the sentences would appear on the screen for a set amount of time and to not worry if they were unable to generate an inference during the presentation of any given pair. Instead, we wanted them to indicate only when they had been successful.

Order of presentation of the sentence pairs was randomized for each participant. In an attempt to equate the likelihood that the two age groups would successfully generate an inference, presentation rate was set at 400 ms/word for younger adults and 700 ms/word for older adults. These rates were based on pilot data collected on each age group. We equated for the likelihood of inference generation to minimize effects of age-related slowing in processing speed on comprehension.

After each pair of sentences had been presented for study, participants were asked to judge how well they would be able to produce the second sentence when they were given the first sentence as a recall cue later in the experiment. Participants were prompted to make the judgment as follows: "You will be shown the first sentences from the pair you just studied in about 15 minutes. How confident are you that you will be able to recall the second sentence of the pair? Enter your judgment below, using any value from 0 to 100, where 0 = *definitely won't be able to recall* and 100 = *definitely will be able to recall.*"

After all sentence pairs had been studied and judged, participants were tested on their memory for the pairs. The computer presented each cue sentence, one at a time, and participants were instructed that they should type as much of the second sentence as they could remember, even if they could not remember the entire sentence. Recall was scored for propositional content. Participants' performance for each pair was based on the proportion of propositions correctly recalled.

Results

Judgment Magnitude, Measures of Processing Ease, and Test Performance

Judgment magnitude. For each participant, we computed the mean judgments. Means across individual values are presented in Table 1. A 2 (age: younger adult vs. older adult) × 3 (coherence level: high, moderate, and low) analysis of variance (ANOVA) revealed a main effect of age, $F(1, 68) = 5.02$, $MSE = 1173.3$, $p < .05$, with older adults making slightly higher judgments. Most important, a main effect occurred for coherence level, $F(2, 136) = 57.6$, $MSE = 67.1$, $p < .05$, and the Age × Coherence interaction was not significant, $F(2, 135) = 1.97$, $MSE = 67.1$, $p > .10$. Follow-up t tests for the main effect of coherence level revealed that the differences in judgment magnitude between each level of coherence were significant, ts$(69) > 4.7$, $p < .05$, indicating that as the coherence of the sentence pairs increased so did the judgments. These outcomes demonstrate that the judgments of both older and younger adults are sensitive to coherence level as predicted by the EOP hypothesis.

Table 1
Judgment Magnitude and Test Performance in Experiment 1

	Low coherence		Moderate coherence		High coherence	
Measure	*M*	*(SEM)*	*M*	*(SEM)*	*M*	*(SEM)*
Judgment magnitude						
Younger adult	31	(2.9)	43	(2.6)	48	(3.8)
Older adult	45	(3.9)	53	(3.4)	57	(3.5)
Test performance						
Younger adult	51	(3.9)	70	(2.8)	68	(2.8)
Older adult	28	(3.5)	41	(3.6)	46	(4.0)

Inference generation. Whether or not participants were able to generate an inference for a given pair provides one measure of processing ease. For each participant, we computed the proportion of pairs for which he or she indicated that an inference had been generated.

As evident from inspection of Table 2, both age groups were largely equivalent at generating inferences for the sentence pairs, $F(1, 68) = 0.01$, $MSE = .27$. Coherence level had a significant effect, $F(2, 136) = 33.8$, $MSE = 0.04$, $p < .05$, and the interaction was not significant, $F(2, 136) = 1.47$, $MSE = 0.04$. For coherence level, all three values were significantly different, ts$(69) > 3.4$, which establishes that, as sentence pairs became more coherent, the likelihood of generating an inference increased.

Generation latency. For pairs in which participants indicated they had successfully generated an inference, we computed each person's median generation latency (from onset of the target sentence to keypress response). Means across these medians are reported in Table 2. Although generation latencies tended to be longer for older than younger adults, the main effect of age was not significant, $F(1, 50) = 0.8$, $MSE = 4.0$. The main effect of coherence level as well as the interaction were also not significant (Fs < 2.2, MSEs < 0.6, ps $> .10$). Thus, the impact of coherence level on processing ease was mainly on whether people generated inferences and not on the speed of generation when they were successful.

Test performance. For each participant, we computed the percentage of correctly recalled propositions from the target sentences separately for the three coherence levels. The mean values across participants are presented in Table 1. A 2 (age) × 3 (coherence level) ANOVA revealed that recall was greater for younger than older adults, $F(1, 68) = 36.2$, $MSE = .09$, and that recall differed across coherence levels, $F(2, 136) = 37.3$, $MSE = .02$. The interaction was not significant, $F(2, 136) = 1.3$, $MSE = .02$. Concerning the main effect of coherence level, recall was significantly less for low-coherence pairs than for high-coherence pairs, $t(69) = 7.0$, or for moderate coherence pairs, $t(69) = 6.9$. Recall for the latter pairs did not significantly differ, $t(69) = 0.9$.

Table 2
Measures of Processing Ease, Experiment 1

Measure	Low coherence		Moderate coherence		High coherence	
	M	*(SEM)*	*M*	*(SEM)*	*M*	*(SEM)*
Inference generation						
Younger adult	.34	(.04)	.55	(.05)	.67	(.06)
Older adult	.39	(.06)	.55	(.07)	.60	(.06)
Generation latency (s)						
Younger adult	2.4	(.12)	2.3	(.06)	2.1	(.08)
Older adult	2.5	(.39)	2.6	(.35)	3.1	(.37)

Note. Inference generation is the proportion of sentence pairs for which participants indicated successfully generating an inference during study. Generation latency is the time between the presentation of the target sentence of a pair and participant's response that an inference had been generated.

Relations Among Measures of Processing Ease and Judgments

We now examine the relationship between the judgments of text learning and processing ease for inference generation and generation latency. For both measures, the issue concerns whether they will correlate with the judgments as predicted by the EOP hypothesis and whether the correlation between processing ease and judgments will remain different from zero with coherence level partialed out.

Inference generation and judgments. For each participant, we computed three correlations across sentence pairs to evaluate the EOP hypothesis: (a) between coherence level (1 = low, 2 = moderate, 3 = high) and judgments, which should be positive given the outcomes presented in Table 1; (b) between whether an inference was generated and the judgments, which was predicted to be positive; and (c) between whether an inference was generated and the judgments, with coherence level partialed out. The Goodman-Kruskal gamma was used for all correlations, as recommended by Nelson (1984; for partial gamma, see Quade, 1974). The means across individual correlations are reported in Table 3.

Several outcomes are noteworthy. First, as evident in Table 1, both age groups' judgments were positively related to the coherence level of the sentence pairs. Second, the judgments for both age groups were significantly related to generation success, although this relationship was considerably weaker for older than younger adults. Third, and most important, when partialed on coherence level, the correlation between generation success and judgments remained significant for younger adults but not for older adults. In summary, whereas both age groups' judgments were entrained to coherence level of the sentence pairs, processing ease as measured by the success of inference generation was less influential for older than younger adults within levels of coherence.

Generation latency and judgments. To assess the relationship between generation latencies and judgments, we computed the three correlations for only those sentence pairs in which participants indicated they had successfully generated an inference (Table 3). In contrast to generation success, participants' judgments were not significantly related to the latency of generating inferences on this task. Evidently, this particular aspect of processing ease in this context contributed minimally to judgments of text learning.

Accuracy of the Judgments and the Diagnosticity of Processing Ease

Another issue that is focal to the metacomprehension literature is the degree to which people's judgments accurately predict test performance. Because little is known about whether aging influences the accuracy of judgments for text learning, we explore this issue here along with the diagnosticity of inference generation for predicting test performance. *Cue diagnosticity* refers to how well the cue of processing ease predicts test performance and hence can contribute to the overall magnitude of judgment accuracy.

Relative judgment accuracy. As in previous research, we operationalized the relative accuracy of judgments as the intraindividual gamma correlation between each participant's judgments and subsequent test performance across the sentence pairs. This correlation represents the degree to which a participant accurately predicts different levels of test performance for each sentence pair relative to others. The means across individual correlations are reported in Table 4. Both groups demonstrated above-zero accuracy ($ts > 5.0$). An age-related trend exists in judgment accuracy, but this difference was not significant, $t(66) = 0.53$. Note that the trend favors older adults and hence is inconsistent with an age-related deficit in judgment accuracy.

Cue diagnosticity. Given that inference generation was related to judgments for both age groups (although this relationship was diminished for older adults), a question arises as to whether this particular cue contributes to the above-chance accuracy of the judgments at predicting test performance (see Table 4). That is, even if people use this cue for judging how well they have learned

Table 3
Relations Among Measures of Processing Ease, Judgments, and Coherence Level in Experiment 1

Measure	Young		Old				
	M	*(SEM)*	*M*	*(SEM)*	*t*	*df*	*p*
Inference generation							
Judgment, coherence	.39[a]	(.04)	.32[a]	(.07)	0.84	67	.40
Judgment, generation	.56[a]	(.06)	.29[a]	(.11)	2.16	59	.03
Partial[c]	.46[a]	(.06)	.12	(.13)	2.52	57	.02
Generation latency							
Judgment, coherence	.29[a]	(.07)	.16[b]	(.09)	1.07	59	.29
Judgment, latency	−.18	(.07)	.04	(.09)	1.93	59	.06
Partial[c]	−.10	(.08)	−.09	(.11)	0.08	59	.94

Note. Cell entries are means of intra-individual correlations between the two measures listed in each row (with corresponding standard errors in parentheses). Coherence = 1 for low, 2 for moderate, and 3 for high coherence. Generation means whether or not an inference had been generated for a sentence pair.
[a] Different from 0 ($p < .05$). [b] $.05 < p < .10$. [c] Correlation between judgments and the corresponding measure of processing ease, with coherence level partialed out.

Table 4
Relations Among Judgments and Test Performance (Accuracy) and Processing Ease and Test Performance (Cue Diagnosticity)

	Accuracy[a]		Cue diagnosticity[b]		Partial[c]	
Age group	*M*	*(SEM)*	*M*	*(SEM)*	*M*	*(SEM)*
Experiment 1[d]						
Younger adult	.24[e]	(.03)	.27[e]	(.06)	.19[e]	(.04)
Older adult	.29[e]	(.05)	.18	(.11)	.23[e]	(.06)
Experiment 2						
Younger adult	.16	(.08)	−.11	(.08)	.11	(.10)
Older adult	.23[e]	(.09)	−.14	(.08)	.22[e]	(.09)

[a] Mean across-intraindividual correlations between judgments and test performance. [b] Mean across-intraindividual correlations between measures of processing ease and test performance. [c] Correlation between judgments and test performance, with the corresponding measure of processing ease partialed out. [d] Inference generation was used as the measure of processing ease for these analyses. See text for details. [e] Different from 0, $p < .05$.

a text, doing so will not benefit judgment accuracy if the cue is not predictive—or diagnostic—of criterion test performance.

To estimate the diagnosticity of inference generation, we computed intraindividual correlations between inference generation and test performance. Mean values across participants are presented in Table 4. Inference generation was significantly diagnostic of test performance for younger adults but not for older adults, although the difference between these correlations was not statistically significant, $t(59) = 0.48$. When judgment accuracy was computed with inference generation partialed out (see Table 4), judgment accuracy was still greater than zero for both age groups. This outcome suggests that the use of processing ease as a cue contributed minimally to the predictive accuracy of people's judgments of text learning.

Discussion

Two major conclusions from Experiment 1 are that (a) age-related differences were not evident in judgment accuracy, suggesting that aging leaves this metacomprehension skill intact, and (b) processing ease is related to people's judgments as predicted by the EOP hypothesis, but these EOP effects were diminished for older adults. A question arises as to whether these outcomes will generalize to judgments of longer, more naturalistic text materials. Older adults' judgments may be even less sensitive to processing ease for materials that are longer than the short sentence pairs used in Experiment 1. That is, one reason for expecting age differences in judgments of text learning was that older adults would be more likely to forget processing disruptions that occur while reading when they later made their judgments, which could be even more problematic when people read and judge longer texts. Alternatively, the obvious differences in causal coherence across the two-sentence text sets used in Experiment 1 may have captured older adults' attention (cf. Hertzog, Kidder, Powell-Moman, & Dunlosky, 2002), overshadowing the more subtle cue of ease of inference generation. If so, older adults' judgments may be more sensitive to the measured processing ease with longer texts in which the coherence manipulation may be less salient. Accordingly, in Experiment 2, we again estimate EOP effects and judgment accuracy but instead use longer text, which also better reflects the kinds of material most often used in the metacomprehension literature.

Experiment 2

In Experiment 2, participants read eight expository paragraphs on different topics (e.g., scientific discovery, art forgery) that, on average, contained 370 words/paragraph. To provide evidence relevant to the EOP hypothesis, each paragraph had two versions: one that was revised to be coherent and easily read and a second that was revised to increase the difficulty of establishing connections between sentences, which was expected to disrupt reading (for details on text construction, see Rawson & Dunlosky, 2002, Experiment 3). Such revisions were designed to decrease the ease of constructing a coherent representation of the text while reading. Each participant read eight different paragraphs: four coherent ones and four less coherent ones. According to the EOP hypothesis, participants' judgments will be greater for the more coherent than for the less coherent text.

We also obtained a measure of online processing ease, which was modified from Karabenick (1996). Each sentence of a text was presented individually for self-paced reading, and the computer recorded how often each participant indicated, by making a keypress, having difficulties for each sentence. Processing ease was operationalized as the number of keypresses per sentence that occurred while a participant was reading a given text. According to the EOP hypothesis, judgments of text learning should be positively related to coherence level and negatively related to the number of keypresses, which indicate online processing difficulty. The latter relationship between judgments and number of keypresses was also expected to remain after partialing out coherence level. Finally, these methods will also support an analysis of our secondary issue: whether age-related differences occur in the accuracy of judgments for predicting test performance across longer texts.

Method

Participants

Twenty-eight older ($M = 64.1$ years, $SD = 14.2$) and 33 younger ($M = 19.4$ years, $SD = 2.9$) adults were recruited in the same manner and received the same compensation as in Experiment 1. The older adults were well educated ($M = 14.2$ years of education, $SD = 2.4$). Older adults reported taking more medications ($M = 1.7$, $SD = 1.9$) than younger adults ($M = 0.3$, $SD = 0.6$), $t(58) = 3.95$, $p < .05$. However, on a scale ranging from 1 (*excellent*) to 4 (*poor*), both older adults ($M = 1.6$, $SD = 0.5$) and younger adults ($M = 1.6$, $SD = 0.7$) rated themselves as being in quite good health.

Materials

Eight paragraphs (from Rawson & Dunlosky, 2002, Experiment 3), each discussing a different topic, were used. Two versions of each paragraph were constructed: a more coherent version (coherent) and a less coherent (incoherent) version. Each participant received all eight topics, but four paragraphs were coherent and four were incoherent. Assignment of text to coherence condition was randomized across participants, as was the order of text presentation. Participants were tested on each paragraph with eight

multiple-choice questions. Half of the questions for each paragraph could be answered directly from the text, and half could be answered based on inferences from the text. See the Appendix for an example of each version from one text, along with corresponding test questions.

Procedure

Participants were instructed that they would be reading paragraphs one sentence at a time and would be asked to answer questions about them later. In addition, any time they were having difficulty understanding the material, participants were to press the *Enter* key on the keyboard. The computer recorded how often they pressed the key for each sentence of a text.

Participants controlled the sentence presentation rate by pressing a key on the screen when they had finished reading each sentence. While reading, they pressed the *Enter* key if they had any problems with the sentence on the screen. Participants were not allowed to return to previously read sentences. After reading a given text, participants completed the judgments by responding to the question, "How certain are you that you will be able to correctly answer questions about this text in approximately 10 minutes?", on a scale ranging from 0 (*definitely unable*) to 100 (*definitely able*). After making their judgment, they read the next text. After all paragraphs had been read, participants were tested over each.

Results

Judgment Magnitude, Measure of Processing Ease, and Test Performance

Judgment magnitude. For each participant, we computed the mean judgment separately for the two levels of coherence. Mean values across participants are presented in Table 5. A 2 (age: younger adult vs. older adult) × 3 (coherence level) ANOVA was conducted. The main effect of coherence level was significant, $F(1, 59) = 71.6$, $MSE = 51.5$, $p < .05$, indicating that judgments were greater for coherent than for incoherent paragraphs. Although the main effect of age approached significance, $F(1, 59) = 2.6$, $MSE = 715.9$, $p = .11$, both age groups demonstrated the critical effect of coherence level on judgments, an observation that is supported by the nonsignificant interaction, $F(1, 59) = 0.01$, $MSE = 51.5$. These outcomes demonstrate that both older and younger adults' judgments are sensitive to coherence level as predicted by the EOP hypothesis.

Number of disruptions. Processing ease was operationalized as the number of times a participant pressed a key to indicate that his or her understanding was disrupted. For each participant, we computed the mean number of keypresses for each sentence per text. The means for coherent and incoherent paragraphs were .09 ($SE = .02$) and .69 ($SE = .30$) for younger adults, respectively, and .28 ($SE = .17$) and .62 ($SE = .16$) for older adults, respectively. Neither the main effect of age, $F(1, 59) = 0.06$, $MSE = 1.5$, nor the interaction, $F(1, 59) = 0.6$, $MSE = 0.9$, was significant. The main effect of coherence level, $F(1, 59) = 7.67$, $MSE = 0.9$, $p < .05$, was significant, indicating that participants reported having more difficulty processing the incoherent texts than the coherent texts, which supports the construct validity of the keypress measure of processing ease.

Test performance. For each participant, we computed the percentage of correct answers to test questions. The mean values across participants are reported in Table 5. A 2 (age) × 2 (coherence level) ANOVA revealed a main effect of age favoring older over younger adults, $F(1, 59) = 11.3$, $MSE = 0.03$, and a main effect of coherence level favoring coherent over incoherent texts, $F(1, 59) = 6.28$, $MSE = 0.01$. The interaction was not significant, $F(1, 59) = 0.20$.

Table 5
Judgment Magnitude and Test Performance in Experiment 2

	Incoherent		Coherent	
Measure	*M*	*(SEM)*	*M*	*(SEM)*
Judgment magnitude				
Younger adult	34	(3.6)	45	(3.9)
Older adult	27	(2.8)	38	(3.6)
Test performance				
Younger adult	36	(2.6)	42	(2.4)
Older adult	48	(3.2)	52	(2.6)

Relations Among Processing Ease and Judgments

As in Experiment 1, we computed three correlations across texts to evaluate the EOP hypothesis: (a) between coherence level (0 = *incoherent,* 1 = *coherent*) and judgments, which should be positive given the outcomes presented in Table 5; (b) between the mean number of keypresses per sentence and the judgments, which was predicted to be negative; and (c) between the number of keypresses and the judgments, with coherence level partialed out. The means across individual correlations are reported in Table 6.

Several outcomes are germane to aging and metacomprehension. First, the correlations indicate that both older and younger adults' judgments were positively related to the coherence level of the texts (see Table 6). Second, the judgments for both age groups were significantly related to the measure of processing ease, and the magnitude of this relationship was not significantly different for older and younger adults. Third, when partialed on coherence level, the relationship between processing ease and judgments was significantly reduced for younger adults, $t(27) = 4.58$, and for older adults, $t(24) = 2.44$. More important, the partial correlations remained greater than zero and did not significantly differ for the two age groups. These results provide strong support for the EOP hypothesis.

Accuracy of the Judgments and the Diagnosticity of Processing Ease

Relative accuracy. We operationalized accuracy as in Experiment 1. Predictive accuracy (see Table 4) did not differ between age groups, $t(58) = 0.58$, and accuracy was not significantly different from zero for younger adults, $t(32) = 1.92$, $p = .06$, but it was significantly greater than zero for older adults, $t(26) = 2.73$. Although both values are low (as in previous research; e.g., see Maki, 1998b), the important outcome here is the absence of age deficits.

Cue diagnosticity. We also assessed the cue diagnosticity of processing ease by computing intraindividual correlations between number of keypresses and test performance. Mean values across

Table 6
Relations Among Processing Ease, Judgments, and Coherence Level in Experiment 2

	Young		Old				
Measure	*M*	*(SEM)*	*M*	*(SEM)*	*t*	*df*	*p*
Judgment, coherence	.62[a]	(.06)	.66[a]	(.09)	0.37	58	.72
Judgment, processing ease	−.55[a]	(.08)	−.66[a]	(.07)	1.02	52	.31
Partial[b]	−.31[a]	(.12)	−.53[a]	(.09)	1.48	51	.15

Note. Coherence = 0 for incoherent texts and 1 for coherent text. Processing ease is the mean number of keypresses per sentence (to indicate a difficulty in understanding while reading).
[a] Different from 0, $p < .05$. [b] Correlation between judgments and the measure of processing ease, with coherence level partialed out.

participants (see Table 4) were not significantly different from zero for either age group. Not surprisingly, the correlation between judgments and test performance (accuracy) was not influenced by partialing this correlation on number of keypresses.

General Discussion

Aging and Metacomprehension

Outcomes from the present experiments not only further establish the metacognitive abilities of older adults, but, in doing so, they also provide new and critical support for the EOP hypothesis for judgments of text learning. Specifically, in each experiment, a text characteristic was manipulated to influence processing ease across texts. These coherence manipulations had the expected effects on younger adults' judgments (replicating outcomes reported by Rawson & Dunlosky, 2002), and these effects were also evident for older adults. Thus, aging does not appear to decrease the sensitivity of judgments of learning to changes in the manipulated coherence of text materials.

Although these effects support the EOP hypothesis for both age groups, an alternative interpretation arises from recent theory about how learning judgments are made (Koriat, 1997). As detailed early in this article, judgments may be based on an explicit analysis of the text characteristics (e.g., "those two sentences are not at all related to one another so I probably won't remember them") in a manner that could partly account for the main effects illustrated in Tables 1 and 5. This rival of the EOP hypothesis was ruled out by analyses based on the measures of processing ease. For younger adults, the likelihood of inference generation (Experiment 1) and the number of reported disruptions (Experiment 2) were significantly correlated with judgments even after the corresponding coherence manipulations were statistically controlled. For older adults, EOP effects were less evident for the shorter text sets used in Experiment 1, but the number of disruptions that older adults reported when reading longer texts was significantly correlated with judgments after the coherence manipulation was statistically controlled. Of course, the remaining EOP–judgment relationships may be produced by some unknown factor that is related to text characteristics (within each level of coherence). Given the homogeneity of texts within each level, however, the influence of such an unspecified factor would presumably be small and may even be overshadowed by the salient variability in processing ease.

Processing ease was not significantly influential on judgment in only one situation: for older adults in Experiment 1. Their gamma correlation between inference generation and judgments was relatively low (.29), and it did not remain significant after partialing on the coherence manipulation (.12). The corresponding correlations for younger adults were significantly larger. This age-related difference may have arisen if intraindividual variability in inference generation was less for older than younger adults, because diminished variability in a cue can reduce its influence on judgments of text learning (Baker & Dunlosky, in press). In contrast to this possibility, the variability in inference generation for older and younger adults was essentially the same, as indicated by the age equivalence in inference generation across the pairs (see Table 2). Another plausible explanation is that using simple sentence pairs that differ dramatically in relatedness (see Appendix) is responsible for this age difference. In particular, older adults may rely heavily on an analysis of the highly salient cue of causal coherence for these short sentence pairs (cf. Hertzog et al., 2002), which may shift their attention away from more subtle differences in processing ease that occurred across the pairs. According to this explanation, older and younger adults are equally able to incorporate various cues into judgments of text learning, but older adults may more heavily weigh a single cue (e.g., causal coherence between sentence pairs) when it readily captures attention. We leave investigation of this possibility to future research.

Despite this single age difference, the overall pattern of results for younger and older adults was clearly more similar than different. Older and younger adults' judgments were consistently related to the manipulations of coherence, and when they read and judged lengthy and more naturalistic texts in Experiment 2, the judgments from both age groups were significantly related to processing ease. In both experiments, the accuracy of the judgments at predicting test performance did not differ across age groups, and the trend slightly favored older adults. Thus, consistent with the literature on associative learning (Hertzog & Hultsch, 2000), aging appears to leave monitoring of learning of text materials largely intact. These results also lead to another empirical question: Did using processing ease as a basis for their judgments contribute to judgment accuracy? We return to this issue after briefly reviewing previous research on aging and judgment accuracy.

Aging and Judgment Accuracy

Current conclusions from the three studies on aging and the accuracy of judgments of text learning are mixed. Two studies

reported age equivalence in accuracy. Lin, Zabrucky, and Moore (2002) had older and younger adults read texts of varying difficulty levels. Across levels of difficulty, accuracy ranged from −.09 to .47, and, most important, age-related effects were not found across these conditions. Olin and Zelinski (1997) also reported age equivalence in accuracy; the correlation between judgments and test performance was about .10 for both age groups. In both of these studies, participants read and judged relatively lengthy paragraphs. In contrast, Miles and Stine-Morrow (2004) had participants read sentences of varying propositional density. They reported small age-related differences in correlations favoring younger adults by about .08 to .18 (values estimated from their Figure 5).

Results from the present research inform this small but growing literature in two ways. First, in both experiments reported here, age equivalence in predictive accuracy was evident, and older adults' accuracy was consistently above zero. Thus, except for the small age deficits reported by Miles and Stine-Morrow (2004), it appears that aging leaves the accuracy of judgments of text learning largely spared. Nevertheless, the lack of age deficits in judgment accuracy should not be viewed too optimistically. The irony lies in the fact that, although age equivalence in judgment accuracy is prevalent, the absolute magnitude of accuracy for both older and younger adults in these studies has consistently been poor. Thus, there may not be an age deficit, but everyone's performance is poor. Also, other conditions that have not yet been investigated may demonstrate age-related deficits in accuracy. For instance, some conditions produce more accurate predictions for younger adults (e.g., Dunlosky et al., 2005; Thiede, Dunlosky, Griffin, & Wiley, 2005), and perhaps age deficits will emerge as overall accuracy increases. Moreover, the bulk of the metacomprehension literature has mainly focused on judgments of text learning and has not systematically explored people's ability to judge their comprehension of recently studied text (see footnote 1). Whether age deficits will occur when adults judge their comprehension is an empirical question. Investigating older and younger adults' abilities here would prove valuable because accurately judging one's comprehension is arguably as important as accurately judging one's learning of text materials.

Second, the effects of processing ease presumably result from the application of a heuristic that easier processing results in higher performance. On the basis of this heuristic approach to human judgment, the accuracy of the judgments is, in part, driven by the diagnosticity of the cues used to make the judgments. That is, high accuracy is not ensured but instead results from the degree to which cues empirically correlate with test performance. As shown in Table 4, processing ease was not significantly correlated with test performance in three of four cases. Given such low diagnosticity, it is not surprising that using processing ease as a basis for judgments of text learning contributed minimally to predictive accuracy. What is perhaps most remarkable here is that, in all cases except one, participants' judgments were influenced by processing ease even though this cue was not diagnostic of test performance.

On the basis of this evidence, we propose that the most influential cues will be those that discriminate among to-be-judged items on any dimension that people believe is relevant to criterion performance (cf. Baker & Dunlosky, in press; Dunlosky & Matvey, 2001; Koriat & Ma'ayan, 2005) regardless of the actual diagnosticity of those cues. One implication, however, is that age effects on predictive accuracy will be manifest under some conditions, such as when a diagnostic cue discriminates between items for younger adults but not for older adults. Note, however, that such effects would not be indicative of age-related deficits in judgment accuracy per se but instead would reflect age-related differences in the available inputs for judgments. Such heuristic-based hypotheses of judgment accuracy can be empirically evaluated using the methods described here, which involve measuring relevant cues to assess their influence on judgments and their diagnosticity at predicting test performance (cf. Hertzog et al., 2003; Koriat & Ma'ayan, 2005; Nelson, Narens, & Dunlosky, 2004).

Summary

Outcomes from two experiments supported the EOP hypothesis. Although an age-related difference was evident in the use of processing ease for simple two-sentence pairs, processing ease had a major impact on both age groups' judgments when longer paragraphs were read. In both experiments, age equivalence was also found in the accuracy of the judgments. Overall, these results support the conclusion that aging leaves people's ability to judge their learning of text material largely intact.

References

Baker, J. M. C., & Dunlosky, J. (in press). Does momentary accessibility influence metacomprehension judgments? The influence of study-judgment lags on accessibility effects. *Psychonomic Bulletin & Review.*

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.

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

Dunlosky, J., Hertzog, C., Kennedy, M., & Thiede, K. (2005). The self-monitoring approach for effective learning. *Cognitive Technology, 10,* 4–11.

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., Rawson, K. A., & Middleton, E. (2005). What constrains the accuracy of metacomprehension judgments? Testing the transfer-appropriate-monitoring and accessibility hypotheses. *Journal of Memory and Language, 52,* 551–565.

Fletcher, C. R., Hummel, J. E., & Marsolek, C. J. (1990). Causality and the allocation of attention during comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 233–240.

Hertzog, C., Dunlosky, J., Robinson, E., & Kidder, D. (2003). Encoding fluency is a cued used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 22–34.

Hertzog, C., & Hultsch, D. F. (2000). Metacognition in adulthood and old age. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 417–466). Mahwah, NJ: Erlbaum.

Hertzog, C., Kidder, D. P., Powell-Moman, A., & Dunlosky, J. (2002). Aging and monitoring associative learning: Is monitoring accuracy spared or impaired? *Psychology and Aging, 17,* 209–225.

Karabenick, S. A. (1996). Social influences on metacognition: Effects of

colearner questioning on comprehension monitoring. *Journal of Educational Psychology, 88,* 689–703.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49,* 294–303.

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., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language, 52,* 478–492.

Lefevre, N., & Lories, G. (2004). Text cohesion and metacomprehension: Immediate and delayed judgments. *Memory & Cognition, 32,* 1238–1254.

Lin, L., Zabrucky, K. M., & Moore, D. (2002). Effects of text difficulty and adults' age on relative calibration of comprehension. *American Journal of Psychology, 115,* 187–198.

Maki, R. H. (1998a). Predicting performance on text: Delayed versus immediate predictions and tests. *Memory & Cognition, 26,* 959–964.

Maki, R. H. (1998b). Test predictions over text material. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 117–145). Hillsdale, NJ: Erlbaum.

Maki, R. H., Foley, J. M., Kajer, W. K., Thompson, R. C., & Willert, M. G. (1990). Increased processing enhances calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 609–616.

Matvey, G., Dunlosky, J., & Guttentag, R. (2001). Fluency of retrieval at study affects judgments of learning (JOLs): An analytic or nonanalytic basis for JOLs? *Memory & Cognition, 29,* 222–232.

Miles, J. R., & Stine-Morrow, E. A. L. (2004). Adult age differences in self-regulated learning from reading sentences. *Psychology and Aging, 19,* 626–636.

Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory & Language, 26,* 453–465.

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., Narens, L., & Dunlosky, J. (2004). A revised methodology for research on metamemory: Pre-judgment recall and monitoring (PRAM). *Psychological Methods, 9,* 53–69.

Olin, J. T., & Zelinski, E. M. (1997). Age differences in calibration of comprehension. *Educational Gerontology, 23,* 67–77.

Quade, D. (1974). Nonparametric partial correlation. In H. M. Blalock (Ed.), *Measurement in the social sciences* (pp. 369–398). London: Macmillan.

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.

Rawson, K. A., Dunlosky, J., & Thiede, K. W. (2000). The rereading effect: Metacomprehension accuracy improves across reading trials. *Memory & Cognition, 28,* 1004–1010.

Robinson, A. E., Hertzog, C., & Dunlosky, J. (in press). Aging, encoding fluency, and metacognitive monitoring. *Aging, Neuropsychology, & Cognition.*

Schwartz, B. L., & Metcalfe, J. (1994). Methodological problems and pitfalls in the study of human metacognition. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 93–113). Cambridge, MA: MIT Press.

Souchay, C., Isingrini, M., & Espagnet, L. (2000). Aging, episodic memory feeling-of-knowing, and frontal functioning. *Neuropsychology, 14,* 299–309.

Thiede, K. W., & Dunlosky, J. (1994). Delaying students' metacognitive monitoring improves their accuracy at predicting their recognition performance. *Journal of Educational Psychology, 86,* 290–302.

Thiede, K. W., Dunlosky, J., Griffin, T., & Wiley, J. (2005). Understanding the delayed-keyword effect on metacomprehension accuracy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1267–1280.

Weaver, C. A. (1990). Constraining factors in calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 214–222.

Wiley, J., Griffin, T. D., & Thiede, K. W. (2005). Putting the comprehension in metacomprehension. *Journal of General Psychology, 132,* 408–428.

(*Appendix follows*)

Appendix

Example Materials for Experiment 1

Sentence type	Sentence
High causal cue	Jake wanted to help the Third World people.
Moderate causal cue	Jake was very concerned about world hunger.
Low causal cue	Jake majored in engineering in Vermont.
Target	He joined the Peace Corps after graduation.

Example Materials for Experiment 2

"Joan of Arc": Coherent Version

Joan of Arc, reported to be a supernatural visitant sent by God to save France in the midst of the Hundred Years' War, inspired the French but clouded the minds and froze the energies of the English. The English were robbed of their assurance by a sense of awe, and even fear, of Joan of Arc. Joan of Arc, who was a village girl from the Vosges, had appeared as the ambassador and plenipotentiary (one given full power to act) of God. Upon Joan of Arc's invocation the spirit of victory changed sides, and Joan led the French in an offensive that never rested until the English invaders were driven out of France. Joan called for an immediate onslaught upon the English besiegers, and herself led the storming parties against them. Joan was wounded by an arrow, but plucked it out and returned to the charge. Joan was hurled half-stunned into the ditch as she mounted the scaling ladders, but as she lay prostrate on the ground, she commanded new efforts. She cried, "Forward, fellow countrymen, God has delivered them into our hands!" Joan saved Orleans, as one by one the English forts fell and English garrisons were slain, and the siege was broken. However, despite Joan's victories and her services to the Court of Charles VII, King of France, the attitude of both the Court and the Church toward Joan eventually began changing. The Court realized that Joan served France rather than one particular political interest, and the Church realized that she served God rather than the Church. Thus, the Court and the Church, which were the powerful particularist interests that had hitherto supported Joan, were estranged. Joan was captured by the Burgundians, a rival French faction of Orleans, and sold to the rejoicing English for a moderate sum. For a whole year Joan's fate hung in the balance, while careless, ungrateful Charles lifted not a finger to save her. There is no record of Charles offering any ransom; however, there is record of the comment of an English soldier who witnessed her death at the stake. The English soldier said, "We are lost, as we have burned a saint." The soldier's comment proved true: Joan of Arc perished on May 29, 1431, and thereafter the tides of war flowed remorselessly against England.

"Joan of Arc": Incoherent Version

In the midst of the Hundred Years' War to save France was reported to be a supernatural visitant sent by God, Joan of Arc, inspiring the French but clouding the minds and freezing the energies of the English. By a sense of awe, and even fear, of Joan of Arc were the English robbed of their assurance. As the ambassador and plenipotentiary (one given full power to act) of God, a village girl from the Vosges who was Joan of Arc had appeared. The spirit of victory upon Joan of Arc's invocation changed sides, and until the English invaders were driven out of France Joan led in an offensive that never rested the French. Storming parties against them herself leading, an immediate onslaught upon the English besiegers Joan called for. By an arrow was wounded Joan, but out she plucked it and to the charge returned. Half-stunned into the ditch Joan was hurled as the scaling ladders she mounted, but she commanded new efforts as on the ground prostrate she lay. "Forward, fellow countrymen," she cried, "God has delivered them into our hands!" As one by one fell the English forts and were slain English garrisons Joan saved Orleans, and broken was the siege. The attitude of both the Court and the Church toward Joan eventually began changing, despite Joan's victories and her services to the Court of Charles VII, King of France, however. That Joan served France rather than one particular political interest realized the Court, and that rather than the Church she served God the Church realized. The powerful particularlist interests that had hitherto supported Joan that were the Court and the Church were thus estranged. By the Burgundians, a rival French faction of Orleans was Joan captured, and to the rejoicing English was for a moderate sum sold. While not a finger to save her lifted careless, ungrateful Charles, in the balance for a whole year Joan's fate hung. Of the comment of an English soldier who her death at the stake witnessed there is record; however of Charles offering any ransom there is no record. "We are lost, as we have burned a saint," said the English soldier. Joan of Arc perished on May 29, 1431, and the tides of war remorselessly against England flowed thereafter, the soldier's comment proving true.

Example of Inference Question

Inferring from the passage, which of the following would have been the most likely outcome if Joan had not been executed by the English?

- The state would have recognized her as a national hero.
- The Church would have provided for her in a convent.
- She would have been exiled and excommunicated. <correct answer>
- She would have supported the French Burgundians.
- She would only have indirectly assisted the French army by offering tactical advice.

Example of Text-Based Question

According to the text, what happened after Joan's death in 1431?

- England began to lose the war. <correct answer>
- The French were driven out of England.
- The Court and the Church were estranged.
- The siege was broken and Orleans was saved.
- Joan was reported to be a supernatural visitant sent by God.

Received July 27, 2005
Revision received December 9, 2005
Accepted December 28, 2005 ■