Individual differences, rereading, and self-explanation:
Concurrent processing and cue validity as constraints on metacomprehension accuracy

Thomas D. Griffin and Jennifer Wiley
University of Illinois, Chicago, Illinois

and

Keith W. Thiede
Boise State University, Boise, Idaho

The typical finding of metacomprehension studies is that accuracy in monitoring one's own level of understanding is quite poor. In the present experiments, monitoring accuracy was constrained by individual differences in both reading comprehension ability and working memory capacity (WMC), but rereading particularly benefited low-ability and low-WMC readers, effectively eliminating the relationship between monitoring accuracy and these reader characteristics. In addition, introducing a self-explanation reading strategy improved the accuracy of all the readers above mere rereading. The observed interaction between individual differences and rereading is interpreted in terms of concurrent-processing constraints involved in monitoring while text is processed, whereas the more general self-explanation effect is interpreted in terms of accessibility of valid, performance-predicting cues.

The typical finding from studies of metacomprehension is that monitoring accuracy is quite poor. Readers' judgments of their comprehension and their actual performance on tests correlate only at about .27 (for reviews, see Lin & Zabrucky, 1998; Maki, 1998; Weaver, Bryant, & Burns, 1995). These results show that readers have poor *relative accuracy*, meaning that they are unable to judge how their understanding varies from one text to the next. Without accurate metacognitive monitoring, readers will be ineffective in regulating their studying, strategies, and cognitive efforts that produce increased learning (e.g., Nelson & Narens, 1990; Thiede & Dunlosky, 1999).

In the present investigations, we explore two factors that may contribute to poor relative metacomprehension accuracy: (1) the requirement to carry out both text processing and monitoring processes concurrently and (2) limited accessibility of valid cues that predict comprehension. In the first experiment, we investigated whether individual differences assumed to relate to executing the concurrent tasks would predict monitoring accuracy but would become less relevant after rereading. The second experiment replicated Experiment 1 while also testing whether a self-explain task designed to shift readers to an inference level of text processing would improve monitoring accuracy for all readers, presumably by increasing access to inference-level monitoring cues.

Successful metacomprehension monitoring requires the execution of two separate tasks. Readers are typically given the goal of reading for comprehension, either explicitly or implicitly via the existence of comprehension tests. In fact, it makes little sense to ask readers to monitor their comprehension if they do not already have the goal of comprehension. However, monitoring of comprehension is a separate, additional task above and beyond reading for comprehension. Just because readers understand a text's meaning (i.e., comprehension) does not mean they have also correctly assessed how complete their understanding of the text actually is (i.e., metacomprehension).

The implications of metacomprehension's being a secondary process to text comprehension have been largely ignored, but they follow directly from what have been widely presumed to be critical features of the metacognitive system. The initial and current models of metacognition presume that different cognitive processes are occurring at two separate levels, referred to as procedural versus metaprocedural, cognitive versus metacognitive, or *object level* versus *metalevel* (e.g., Fischer & Mandl, 1984; Flavell, 1979; Nelson & Narens, 1990). Comprehension involves attention to and processing at the object level, where the object is the text, whereas metacomprehension involves attention to and processing at the metalevel, where the focus is upon one's own mental processes and

T. D. Griffin, tgriffin@uic.edu

representation of the text. In addition, the act of monitoring consists of an imbalanced "dominance relation" in which the metalevel is secondary to and must "be informed by the object-level" (Nelson & Narens, 1994, pp. 11–12). Our present hypotheses are derived largely from this basic assumption that monitoring is a secondary process additional to the primary process of comprehension.

Once one accepts this dual-processing assumption, several interesting implications unfold. When readers have only a single exposure to the text, as is most often the case in metacomprehension research, the success of the secondary process of metacomprehension may depend on whether it can be carried out concurrently with all of the comprehension processes that include even the lowest level text decoding. An explicit judgment of comprehension is presumably informed by, but is not the same as, attending to and encoding metalevel cues that could be later used in those explicit judgments. The actual judgments may still occur after reading, as they almost always do in metacomprehension research. However, the accuracy of those judgments will depend on whether or not the reader attended to the metalevel during reading and encoded the comprehension cues that are generated by and made most available during the reading process.

Without this concurrent attention, readers will have access to less metalevel information that provides the basis for inferring their comprehension levels whenever they do make their judgments. The less attention readers pay to the metalevel during reading, the more their judgments will be dominated by the metalevel information generated by the end of the text that happens to be active postreading. This metalevel information may be unrepresentative of their comprehension of the overall text. Except for rather short and simple texts, this postreading monitoring is likely to produce inaccurate judgments.

Readers are faced with the conundrum that monitoring may be most effective when done concurrently with reading, yet attending to the metalevel while processing the text may be quite difficult, due to cognitive constraints. One of the earliest theoretical treatments of metacomprehension foreshadowed the potential problem posed by concurrent monitoring and text processing by noting that we might expect "bottlenecks in processing capacity or some kind of interference between the metaprocedural and procedural levels" (Fischer & Mandl, 1984, p. 231). When accurate judgments require concurrent monitoring, both individual-differences and contextual factors related to executing the tasks in parallel should influence monitoring accuracy.

Accuracy could be influenced by individual-difference factors that either impact the ease with which readers can process text (e.g., reading ability) or are more generally related to concurrent-processing efficiency (e.g., working memory capacity [WMC]). Reading ability will impact the difficulty of the primary object-level task of text processing. Some readers struggle with one or more of the subprocesses involved in comprehension (e.g., Keenan, Betjemann, Wadsworth, DeFries, & Olson, 2006; Perfetti, 1985; Yuill & Oakhill, 1991). Thus, they should be limited in their ability to concurrently carry out the task of monitoring. A second individual difference, WMC, is theorized and op-erationalized in terms of attentional control and concurrent processing efficacy (e.g., Conway et al., 2005; Engle, Kane, & Tuholski, 1999). Readers who have limited WMC should have difficulty shifting their attention to the metalevel; thus, they may simply focus on the primary task of text processing. Reading ability and WMC relate via different pathways to concurrently executing text processing and monitoring; thus, they should independently influence judgment accuracy, when accuracy depends on concurrent monitoring.

In previous research, the hypothesized relationship between WMC and monitoring accuracy has not been examined, and only limited support has been found for a relationship between reading ability and monitoring accuracy. Some studies have shown evidence consistent with a positive relationship (Bol & Hacker, 2001; Bol, Hacker, O'Shea, & Dwight, 2005; Garner, 1987; Glover, 1989; Hacker, Bol, Horgan, & Rakow, 2000; Maki & Berry, 1984; Maki, Shields, Wheeler, & Zacchilli, 2005; Schneider, 1985), but others have failed to show any relationship (Glenberg & Epstein, 1985; Maki, Jonas, & Kallod, 1994; Maki & Swett, 1987; Pressley, Snyder, Levin, Murray, & Ghatala, 1987). Unfortunately, this body of studies contains such enormous variability in the factors that could moderate ability effects or the capacity to observe them (e.g., text characteristics, test characteristics, or the way reading ability is measured) that they provide little insight into whether, when, or why reading ability might affect monitoring accuracy.

If monitoring accuracy depends on attending to the metalevel during reading and lower ability readers have difficulty attending to the metalevel during effortful text processing, the monitoring accuracy of these readers should benefit from contexts that reduce the required text processing. During an immediate second reading, many of the subprocesses involved in reading do not need to be reexecuted (e.g., Millis, Simon, & tenBrock, 1998; Perfetti, 1985). This reduction in the need for low-level text processing means that readers who otherwise struggle to monitor during text processing could focus more attention at the metalevel. Thus, their monitoring accuracy could improve during rereading.

Sizable increases in monitoring accuracy have been observed as a result of having readers read each text twice before making their judgments of understanding (Dunlosky & Rawson, 2005; Rawson, Dunlosky, & Thiede, 2000; but see Maki, Holder, & McGuire, 2001). Importantly, previous research has not viewed the rereading effect in terms of reducing the need for concurrent processing, nor have the potential interactions between rereading and cognitive constraints been examined. However, prior results are compatible with the current perspective, and Dunlosky and Rawson (2005) have noted that "rereading may afford more resources for comprehension monitoring" (p. 51).

Experiment 1 employed naturalistic texts and tests that would be likely to create a context in which the text-processing demands would be high and the accuracy of metacomprehension judgments would depend on concurrent monitoring. The goal of Experiment 1 was to examine how rereading can affect monitoring accuracy and how this effect is moderated by individual differences in the ability to execute the components of text processing that reading requires.

EXPERIMENT 1

Method

Participants and Design

As partial fulfillment of a course requirement, 75 introductory psychology students were randomly assigned to either a read-once or a reread condition. Three participants were dropped due to a lack of variance in their judgments. Both of the present experiments employed the most common measure of metacomprehension monitoring—*relative* accuracy, or *resolution* (Nelson, 1984)—which measures whether readers can judge their understanding of each text relative to other texts, rather than against some absolute referent, such as complete understanding.

Materials and Measures

Texts. The texts were five explanatory texts that described complex causal phenomena from the natural or social sciences (i.e., antibiotic use causing allergies, biological evolution, volcano formation and eruption, racial differences on IQ tests, and ice ages). The texts varied from 650–900 words in length and had Flesch–Kincaid grade levels of 11–12 and Flesch reading ease scores in the *difficult* range of 31–49 (see the Appendix). The antibiotics text served as a practice trial, and the others served as the target texts used in the analyses.

Comprehension tests. The comprehension tests were inference verification tests (IVTs), designed to assess situation-level representations that could not be answered using only surface memory for the text (Royer, Carlo, Dufresne, & Mestre, 1996). The test for each text consisted of 16–20 statements that readers judged as "true or false, based upon what [they] read in the text." The statements were not simple paraphrases or negations but, rather, assertions that were either consistent or inconsistent with inferences that were logically implied by one or more assertions in the text. The set of items were designed to provide relatively complete coverage of all the central ideas and their implicit logical relations (see the Appendix for examples).

Comprehension ability. For several reasons, readers' ability to process and comprehend text was assessed using a composite factor score of readers' performance across the IVTs for the four target texts. A principal-factors analysis was performed (no rotation necessary), yielding a single factor with loading values that represented the degree to which each individual test shared common variance with each other test. These loading values were used as regression weights to compute a factor score for each reader. Principal-factors analysis incorporates only shared variance among the tests; thus, factor scores can be meaningfully interpreted as representing a latent construct (e.g., comprehension ability) that exerts a common influence across tests, independently of influences such as interest or topic knowledge that might influence any individual test.

A more traditional ability measure, such as the Nelson–Denny (N–D), was not used for several reasons. The texts and tests in standardized measures such as the N–D are focused on surface memory for details, rather than on situation-level comprehension of intertextual relations that are the focus of our texts and the present research goals (see Magliano, Millis, Ozurub, & McNamara, 2007). Use of such a set of texts and tests not only would fail to capture important variance in the processing of complex, attention-demanding texts, but also could create problems of fatigue and interactions arising from presenting two sets of texts with fundamentally different structures and expected comprehension levels (see Wiley, Griffin, & Thiede, 2005).

Although performance on the target tests was used in the computation of both comprehension skill and metacomprehension, each measure is independent, due to the use of different and orthogonal variance related to test performance. Comprehension ability scores are dependent only on the weighted average of the absolute levels of performance across the tests. Which test a reader performed best and worst on, relative to each other test, has no statistical bearing upon this ability score. Conversely, relative monitoring accuracy scores are not affected by absolute performance and depend only on these relative differences in performance that do not affect the ability measure. In fact, the most often touted advantage of relative accuracy scores is their statistical independence from ability or other factors represented by the aggregate absolute level of performance that constitutes our ability measure (Nelson, 1984, 1996).

To validate the assumption that the IVT scores share a common influence from a general comprehension ability factor, 50 pilot participants completed both the IVT tests for the target texts and the N–D Reading Comprehension Test (Form H). The single computed factor (accounting for 48% of the variance) for the IVT scores correlated with the N–D scores at $r = .64$ ($p < .05$). The assumption that the IVT factor scores represent a general latent construct (such as comprehension ability) was further supported by the fact that its correlation with N–D was greater than that of any IVT score for the individual texts ($rs = .27–.54$).

Procedure

All materials were presented and responses recorded via computer. The participants were given a general summary of the tasks and their order. Following this verbal overview, all the participants read a brief instruction page which stated, "In this experiment you will be asked to read several texts so that you understand them." The read-once instruction also stated, "Read each text carefully one time, as though studying for an exam." The reread instruction stated, "Read the text quickly the first time just to get the gist or basic idea of the passage. Then, read it the second time more carefully, as if studying for an exam." Texts were presented one paragraph at a time, and the readers advanced to the next paragraph at their own pace. After the last paragraph, the rereading participants were automatically rerouted to the beginning of the text. After reading, the participants were asked, "How well do you think you understand the text you just read?" using a scale of 1 (*not at all*) to 7 (*very well*); then they completed the comprehension test. They repeated this read, judge, test sequence for each text.

Results

Descriptives, Judgments, and Test Performance

Mean judgment magnitude did not significantly differ between conditions [$t(70) = 1.52$, n.s.; see Table 1 for mean judgments and test performance]. Even though average test performance varied from 60% to 85% correct among the four target texts, the patterns of mean performance among the texts were the same for the read-once and reread conditions, and the conditions did not differ in mean performance on any of the tests [all $ts(70) < 0.75$, n.s.].

Metacomprehension Monitoring Accuracy

Relative monitoring accuracy was computed using Pearson correlations of each participant's judgments with his or her corresponding test performances.[1] Monitoring accuracy scores were regressed onto reading condition, reading ability, and the computed interaction term. Table 2 reports the resulting beta weights and t values of the regression. There was a main effect of rereading, so that monitoring accuracy for the reread condition ($M = .46$, $SEM = .06$) was significantly better than that for the read-once condition ($M = .22$, $SEM = .10$). There was no main effect of comprehension ability. However, there was a significant comprehension ability × rereading interaction (see Figure 1). Comprehension ability accounted for a sizable amount of the variance in monitoring accuracy when the participants read the texts only once ($R^2 = .25$) but accounted for no variance when they read the texts twice ($R^2 = .00$). Rereading produced marked accuracy improvements for low-ability readers, but not for higher ability readers.

Table 1
Mean Judgment Magnitude and Proportion Correct for Each Target Text
by Reading Condition (With Standard Errors of the Means)

| | | | Proportion Correct on IVT | | | | | | | |
| | Judgment | | Race and IQ | | Volcanoes | | Evolution | | Ice Ages | |
Condition	M	SEM	M	SEM	M	SEM	M	SEM	M	SEM
Experiment 1										
Read once	4.17	.18	.85	.01	.69	.02	.64	.02	.60	.02
Reread	4.54	.16	.83	.01	.69	.02	.66	.02	.58	.02
Experiment 2										
Read once	4.60	.20	.81	.02	.69	.02	.67	.02	.61	.02
Reread	4.45	.17	.79	.02	.69	.02	.68	.02	.57	.02
Self-explain	4.62	.16	.85	.02	.71	.02	.67	.02	.60	.02

Note—No between-condition differences were significant at $p < .05$. IVT, inference verification test.

Discussion

The observed interaction between comprehension ability and rereading on monitoring accuracy is consistent with the concurrent-processing assumption. Lower ability readers struggled to successfully monitor their performance during a single reading but were able to monitor equally well as high-ability readers when they read the texts twice. The lack of any rereading effect on test performance itself is not surprising, given a similar lack of rereading effects in some previous research (e.g., Dunlosky & Rawson, 2005; Dunlosky, Rawson, & Hacker, 2002). In addition, the present tests were designed to measure inference-level comprehension. Rereading may afford an opportunity to process the text at a more conceptual level. However, without special instruction or training, such as that introduced by McNamara (2004), typical college students may not take advantage of that opportunity and may continue to process the text at a surface level that fosters memorizing details, rather than constructing intertextual inferences.

A limitation of Experiment 1 is the fact that comprehension skill is an indirect indicator of text-processing difficulty. Furthermore, both comprehension ability and monitoring are components of reading skills more generally and likely share a number of common influences. Besides replication, two major goals of Experiment 2 were to examine whether WMC has a similar but independent relationship with monitoring accuracy and rereading, and to introduce a reading strategy manipulation that could improve accuracy for readers of all levels.

Table 2
Stepwise Regression of Metacomprehension Accuracy
on Reading Condition, Comprehension Ability,
and the Condition × Ability Interaction

| | Variables | | | |
Condition	B	SE_B	β	t Value
Step 1 ($R^2 = .08$)				
Rereading	.28	.10	.31	2.86*
Comprehension ability	.10	.06	.18	1.55
Step 2 ($R^2 = .20$)				
Rereading × ability	.42	.13	.35	3.22*

Note—Statistically significant ($p < .05$) betas are those retained in the final model. *$p < .05$.

EXPERIMENT 2

The effect of comprehension ability and its interaction with rereading in Experiment 1 are presumed to result from the amount of effort and attention required for the text-processing task itself in the single-read condition. As comprehension ability increases, the attention required for text processing decreases, and more attention can be directed to monitoring metalevel cues. WMC should impact monitoring accuracy and interact with rereading in a similar pattern as comprehension ability, but for slightly different reasons. WMC is presumed to impact the efficiency and costs of concurrent-processing situations.[2] Thus, in the face of high text-processing demands, higher WMC should allow readers to more efficiently devote some of their attention to monitoring metalevel cues about their comprehension. Thus, the predicted result is that monitoring of low-WMC readers should suffer in single-read conditions because of their inability to control their attention and less efficient execution of whatever concurrent processing is required.

An additional goal of Experiment 2 was to introduce a new manipulation that, unlike rereading, would improve monitoring accuracy for all readers, particularly for explanatory texts where comprehension involves constructing a *situation model* in which the explicit and implied causal and logical interrelations among text ideas and with prior knowledge are represented (Kintsch, 1994). The text comprehension literature holds important lessons for metacomprehension research—namely, that certain texts and tests foster and allow for assessment of surface-memory versus situation-model-level comprehension and that the quality of one level of representation need not correspond to or be a valid predictor of the quality at other levels (Kintsch, 1998; Wiley et al., 2005).

These level-of-representation issues not only are important for the design of appropriate texts and tests; they also directly imply that effective monitoring of cues related to mere recall of surface representations may have limited accuracy in predicting performance that is dependent on the quality of one's situation model. This implication is compatible with the more general and increasingly popular *cue utilization* view, which holds that "the accuracy of metacognitive judgments . . . should depend on the validity of the cues on which [they] rest" (Koriat,
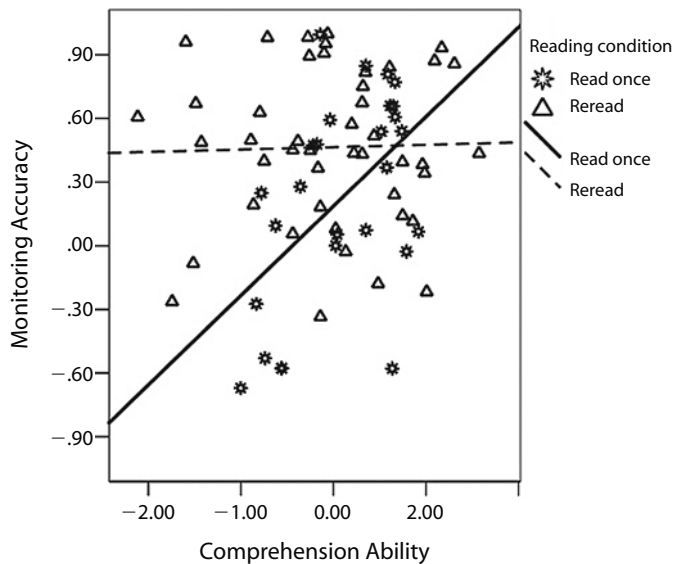
Figure 1. Monitoring accuracy as a function of reading condition and comprehension ability (Experiment 1).

2007, p. 295). Also, a number of metacognitive researchers have suggested that individuals generally base their judgments on superficial cues that are not valid predictors of the assessed test performance (e.g., Benjamin, Bjork, & Schwartz, 1998; Dunlosky et al., 2002; Koriat, 1997). Furthermore, both high- and low-ability readers report basing their metacomprehension judgments on surface-level recall cues (Thiede & Anderson, 2004).

By combining these common perspectives in the metacognition and text comprehension literatures, we can infer that even when readers pay ample attention to the meta-level during reading, judgment accuracy will be limited if test performance requires situation-level comprehension and readers monitor only surface-level cues. This could explain why high-ability readers failed to benefit from the increased opportunity for monitoring during rereading in Experiment 1. With this in mind, we introduced a reading task of *self-explanation*, which has previously been shown to focus readers on their situation-level text representations (Chi, 2000; Wiley & Voss, 1999) and, thus, should increase the salience of cues predictive of inference test performance. We presumed that self-explaining would increase access to valid cues, as opposed to rereading, which we assumed impacted attention to the metalevel more generally. Thus, we expected self-explanation to improve accuracy for all the readers above that with mere rereading.

Method

Participants and Design

Ninety-four introductory psychology students were randomly assigned to a read-once, reread, or self-explanation condition. Two participants were dropped due to a lack of variance in their judgments.

Materials and Procedure

In the first session, the participants completed two measures of WMC, reading span (Rspan) and operation span (Ospan). In accord with the procedures described in Conway et al. (2005), the participants evaluated the correctness of sets of 2–5 sentences (Rspan) or simple arithmetic problems (Ospan), while concurrently encoding and retaining for later recall the single letters (Rspan) or words (Ospan) that appeared next to each sentence or problem. Scores were computed on the basis of the proportion of letters or words correctly recalled for each set. The average score across Rspan and Opsan was computed to generate a composite that represented a general attentional control measure independent of the reading or math abilities that could influence each individual score. Mean WMC scores were equivalent for the read-once ($M = .63$, $SD = .14$), reread ($M = .63$, $SD = .15$), and self-explain ($M = .65$, $SD = .14$) conditions [$F(2,89) < 1$, n.s.].

In the second session, all materials and procedures were identical to those in Experiment 1, with the exception of a written instruction read in the self-explanation condition, based on the work of Chi (2000; Chi, de Leeuw, Chiu, & LaVancher, 1994). The self-explanation instruction added the following to the reread instruction:

> As you read the text the second time, you should try to explain to yourself the meaning and relevance of each sentence or paragraph to the overall purpose of the text. Ask yourself questions like: What new information does this paragraph add? How does it relate to previous paragraphs? Does it provide important insights into the major theme of the text? Does the paragraph raise new questions in your mind? So, try your best to think about these issues and ask yourself these kinds of questions about the text as you read it for the second time.

The instruction also provided a 50-word example text and hypothetical self-explanation comments for each sentence.

As a manipulation check, time on task was analyzed for the participants in the self-explanation and rereading conditions. On average, the explainers' time per text ($M = 365.66$ sec, $SD = 88.25$) was 14.35% longer than the rereaders' time ($M = 319.78$ sec, $SD = 89.62$) [$t(59) = 2.02$, $p < .05$]. This suggests that the readers were engaging in some additional processing in the self-explanation condition.

Results

Judgments and Test Performance

Mean judgment magnitude did not reliably differ among conditions [$F(2,89) = 0.30$, $MS_e = 0.93$; see Table 1]. Mean test performance varied in the same manner as in Experi-

ment 1; again, the pattern of performance across texts was the same for each condition, and performance did not reliably differ between conditions on any of the tests [$Fs(2,89) <$ 1.20, $MS_e < 0.02$], except for RaceIQ [$F(2,89) = 2.23$, $MS_e = 0.01$, $p < .10$]. The marginal effect of condition on the RaceIQ text resulted from the self-explain group's performing slightly better than the other groups.

Metacomprehension Monitoring Accuracy

Comprehension ability and WMC scores were positively correlated at Pearson's $r = .32$, and neither showed a reliable bivariate correlation with relative monitoring accuracy (Pearson's $rs = .10$ and $.12$, respectively). In accord with Aiken and West (1991, p. 71), the three reading conditions were dummy coded into two variables that, when entered together in a regression, represented the effect of rereading, as compared with reading once, and the effect of self-explanation, as compared with rereading. Also, each dummy variable was multiplied by both comprehension ability and WMC scores to compute corresponding interaction terms. Relative accuracy scores were regressed onto the predictors in three steps, as detailed in Table 3, along with the betas and corresponding t tests for each predictor. The greater accuracy for rereading ($M = .39$, $SEM = .07$) over reading once ($M = .21$, $SEM = .09$) was not significant. However, self-explanation reliably improved accuracy ($M = .63$, $SEM = .07$) over rereading. Neither comprehension ability nor WMC had an overall relationship with accuracy.

As in Experiment 1, comprehension ability interacted significantly with rereading (see Figure 2). Lower ability readers had particularly poor accuracy when they read only once ($R^2 = .16$), but their accuracy was higher when they reread (whereas that of higher ability readers remained unchanged), to the point where ability no longer predicted accuracy in the reread condition ($R^2 = .01$). In contrast to rereading, the positive effect of self-explaining did not in-

Table 3
Stepwise Regression of Metacomprehension Accuracy on Condition, Comprehension Ability, Working Memory Capacity (WMC), and the Condition × Ability Interactions

Condition	Variables			
	B	SE_B	β	t Value
Step 1 ($R^2 = .12$)				
Rereading vs. once	.19	.11	.19	1.67**
Explain vs. rereading	.30	.08	.34	3.63*
Comprehension ability	.02	.06	.06	0.56
WMC	.24	.32	.09	0.90
Step 2 ($R^2 = .17$)				
Rereading × ability	.28	.14	.19	2.00*
Explain × ability	.09	.12	.03	0.32
Step 3 ($R^2 = .23$)				
Rereading × WMC	1.58	.60	.25	2.62*
Explain × WMC	.13	.73	.01	0.08

Note—Statistically significant ($p < .05$) betas are those retained in the final model.   *$p < .05$.   **$p < .10$.

teract with ability, and there was no relationship between accuracy and ability in the self-explain condition ($R^2 = .01$). Likewise, WMC interacted with rereading, but not with self-explanation (see Figure 3). Lower WMC readers had poorer monitoring accuracy when reading once ($R^2 = .21$) but benefited from rereading to the extent that WMC no longer affected accuracy ($R^2 = .01$). The benefits of self-explaining above rereading were unrelated to WMC ($R^2 = .00$).

Discussion

Experiment 2 replicated the comprehension ability × rereading effect and demonstrated that WMC also interacts with rereading. Although the pattern of results for comprehension ability and WMC were nearly identical, their effects were independent, as evidenced by their retention in the final stepwise model. In fact, if comprehension ability is entirely removed from the analyses, the
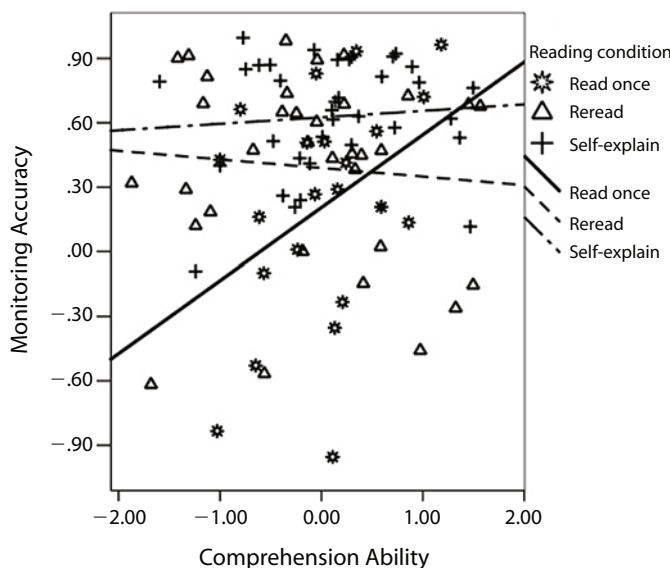

Figure 2. Monitoring accuracy as a function of reading condition and comprehension ability (Experiment 2).
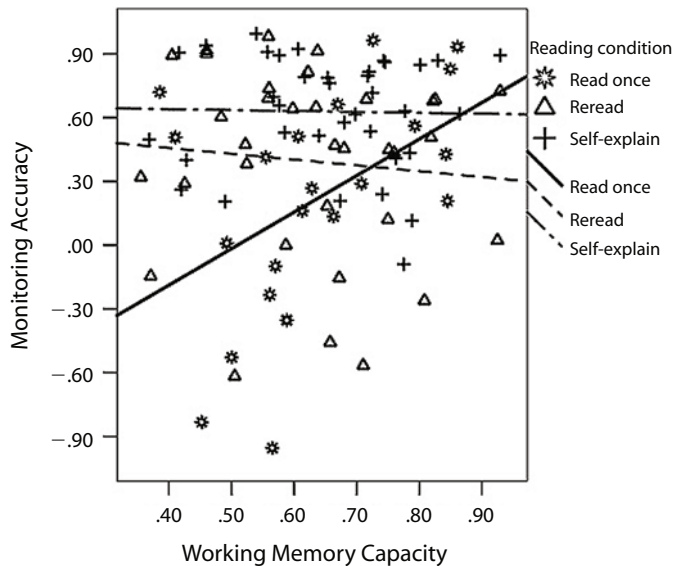
Figure 3. Monitoring accuracy as a function of reading condition and working memory capacity.

WMC $\times$ rereading β increases only from .25 to .28. The independence of the WMC effect on accuracy is consistent with our assumption that both comprehension skill and WMC affect monitoring via their independent relations to concurrent monitoring during text processing. Comprehension skill affects the attention and effort required for text processing, whereas WMC impacts the efficiency of switching attention between the two tasks and, thus, between the object level and the metalevel.

In addition, giving readers an instruction to self-explain during rereading led to a significant improvement in monitoring accuracy for all the readers, unlike rereading by itself, which showed no benefits for high-ability readers. This qualitative difference between the effects suggests a qualitative difference in the constraints impacted by rereading and self-explaining. The rereading effect and its interaction with both comprehension ability and WMC can be accounted for by assuming that rereading reduces the need for low-level text processing that otherwise makes it difficult for some readers to attend to available metalevel cues about their comprehension. The more general effect of self-explanation and the lack of an interaction with individual-difference measures suggest that this effect is not tied to processing constraints that impede monitoring efforts.

Neither the rereading nor the self-explanation effects can be attributed to effects on test performance itself, because the groups did not differ in test performance. The lack of a rereading effect on comprehension itself is consistent with prior findings (e.g., Dunlosky & Rawson, 2005; Dunlosky et al., 2002). The self-explanation instruction used here was not designed to impact test performance, because it did not involve the practice, feedback, or training in constructing quality explanations that has been central to interventions designed to increase comprehension itself (e.g., Chi, 2000; McNamara, 2004). Instead, our

brief self-explanation instruction was intended to prompt readers merely to attempt intertextual explanations, with the assumption that this would increase readers' access to valid metalevel cues, so that their judgments of comprehension would be based on valid cues related to the quality of their situation model.

It is not necessary for readers to actually construct a *good* explanation; they just need to attempt to create any explanation in order to increase the salience and accessibility of inference-predicting cues. Indeed, the effect of self-explanation on monitoring accuracy, but not on test performance, suggests that it was the attempt to explain, and not the successful production of a high-quality explanation, that led to the accuracy effects. This further accounts for the fact that ability and WMC differences did not interact with the self-explanation condition. Higher ability readers might construct superior explanations, but all the readers could attempt to explain, which is presumably what gave them access to valid cues. Thus, the improvement in monitoring performance, the lack of improvement in test performance, and the lack of an explanation $\times$ ability interaction can all be explained by the cue access account of the self-explanation effect.

GENERAL DISCUSSION

The results of the present experiments reveal that variance in the accuracy of monitoring judgments is related to contextual factors (i.e., single vs. multiple readings) and reader characteristics (i.e., individual differences in comprehension ability and WMC) that are plausibly related to being able to monitor metalevel cues during the primary task of text processing. The present assumptions account for the findings, particularly the independent interactions between rereading and both comprehension ability and

WMC, the unmoderated effect of self-explaining, and the fact that comprehension and metacomprehension accuracy diverged in their relationships to other variables.

The core assumptions underlying the present work are the following: that metacognitive monitoring is a distinct process secondary to the processes being monitored (e.g., comprehension); that some comprehension-predicting cues are less accessible after reading than during reading; that readers' success at engaging in concurrent monitoring will depend on available attentional resources; and that even with concurrent monitoring, the types of cues readers typically attend to will have only limited relevance to situation-model-level comprehension. The present findings do not directly test these assumptions but, rather, take them as a starting place from which all our hypotheses and interpretations were derived. Regarding the first of these assumptions, it is inherent in the basic formulations of the metacognitive construct (Fischer & Mandl, 1984; Flavell, 1979; Nelson & Narens, 1990), and it is unclear how any construct that could reasonably be labeled metacognitive monitoring would not be a secondary process to whatever is being monitoring at the object level. The other core assumptions are less self-evident but seem to be in step with more general assumptions in the memory, attention, and metacomprehension literatures. The novel contribution of the present work lies in combining these various assumptions and identifying their implications for how potential individual differences and contextual factors ought to moderate monitoring accuracy.

Our account of the present findings implies that the important factor underlying the rereading × ability/WMC interactions is the varying extent to which readers were able to encode metalevel cues during reading, rather than merely waiting until their postreading judgment to pay any attention to metalevel information. The evidence that accuracy was impacted by the ability to attend to the metalevel while reading is indirect. The core assumptions that we have laid out suggest the following: that such concurrent monitoring would likely influence monitoring accuracy; that comprehension ability, WMC, and rereading would be expected to influence one's ability to execute such concurrent processing; and that each of the observed findings are what would be expected if concurrent monitoring was having an influence.

However, we did not directly observe *when* readers were monitoring or whether any readers actually did monitor while reading, rather than entirely at the time of judgment. Thus, it is important to consider alternatives that might account for the findings. It is possible that all the readers engaged only in postreading monitoring. Perhaps rereading merely gives readers a larger sample of experiences to reflect back on during retrospective judgment; but such an effect should be equal for readers of all ability levels, and it was not. Alternatively, it may be that readers need to process the text at a conceptual level before their reading experiences become predictive of comprehension, and such processing is least likely to occur for lower ability readers during a single reading. However, this account seems to allow WMC to influence accuracy only via its influence on comprehension itself; thus, it fails to explain the independent WMC effects. Other possibilities may exist, but their plausibility will be tested by their ability to account for the totality of the find-

ings, particularly the independent interactions between rereading and both comprehension ability and WMC, plus the main effect of self-explanation that implies that its source of influence is qualitatively different from whatever is responsible for the rereading effect. A challenge for future research is to come up with more direct but unobtrusive methods for assessing when monitoring is actually occurring.

Self-Explaining and Monitoring of the Situation Model

Self-explaining and rereading appear to improve accuracy in different ways. If rereading increases attention to metalevel cues available during reading, perhaps self-explanation increases readers' access to certain types of cues that are most indicative of situation-model-level comprehension. This cue access interpretation of the self-explanation effect is quite compatible with the increasingly popular *cue utilization* view that emphasizes cue validity more generally (e.g., Koriat, 1997; Maki, 1998; Rawson & Dunlosky, 2002; Weaver et al., 1995). It is also compatible with more recent emphases on situation-level cues for predicting performance requiring conceptual understanding (e.g., Dunlosky & Rawson, 2005; Wiley et al., 2005). One of the largest and most robust improvements in accuracy has been produced by the use of prejudgment generation tasks (e.g., summary or keyword listing) that require readers to access their representations (accuracy levels around .6–.7; Thiede, Anderson, & Therriault, 2003; Thiede, Dunlosky, Griffin, & Wiley, 2005). The fact that summary and keyword effects critically depend on a delay between reading and generation has been interpreted as resulting from the decay of surface representations that forces readers to access their situation models, giving them more valid cues for predicting inference-level comprehension (Thiede et al., 2005). However, those findings could also be explained by a simple transfer-appropriate-monitoring (TAM) hypothesis, due to the greater processing similarity (i.e., long-term memory retrieval) between delayed generation and delayed testing. In contrast, the present findings favor a cue accessibility interpretation over the TAM hypothesis, because self-explanation during reading is quite unlike the process of inference verification after reading. Just as Dunlosky, Rawson, and Middleton (2005) recently argued regarding their effects of term-specific judgments, it is not some match between the judgment and performance processes themselves (as presumed by TAM), but rather that the cues produced by the judgment process allow for valid inferences about performance.

Use of different cues or just more of the same? Again, we did not directly observe the types of cues the readers were monitoring; thus, it is conceivable that the self-explanation task merely increased the quantity of monitoring but that all the groups accessed and monitored qualitatively similar types of cues. The two possibilities under this alternative approach would be either that all the readers used surface cues or that all the readers used situation-level cues. The notion that surface-level cues were used by all the readers is suspect because it seems implausible that surface cues would be so highly predictive of performance on an inference test. Self-explainers reached

an uncommonly high level of accuracy (.63) in predicting their inference performance. The idea that surface cues would predict inference performance so well is at odds with common views about the relative independence of representation quality between levels (e.g., Kintsch, 1998) and the need for a match between the representation levels of the cues and the assessment items (e.g., Koriat, 1997), especially considering that the test materials were specifically created to keep the surface and the inference levels distinct. Alternatively, the notion that everyone, even the readers in the read-once condition, relied on situation-level cues contradicts findings and common views that readers typically rely on superficial, surface-level monitoring cues to judge their level of comprehension (for a review, see Maki, 1998). For example, Thiede and Anderson (2004) found that both high- and low-skilled readers report relying heavily on surface-level memory cues.

More important, the present empirical findings show a qualitative difference between rereading and self-explanation effects, with only the latter increasing accuracy for readers of all ability levels. This qualitative difference suggests a distinction in the underlying mechanisms responsible for the effects. Thus, using a more-of-the-same account for both the rereading and the self-explain effects raises problems. Although we find the more-of-the-same account theoretically less plausible than a cue access account, we are currently conducting a series of studies using a test expectancy paradigm that directly tests whether the benefits of increased monitoring are contingent upon its being directed toward the level of representation that matches the one being assessed.

Potential Moderators of Ability and Rereading Effects

It is empirically true that the benefits of rereading were not observed for higher ability readers and that an overall rereading effect would not have been observed without the inclusion of low-ability readers in this sample. Thus, in the present context, reader ability and rereading interacted and moderated each other. However, these results should not be interpreted to suggest that rereading will always benefit low-ability readers or that failures to observe rereading effects may always be due simply to the absence of low-ability readers. Instead, a combination of factors was likely necessary for the present results. Both the relatively long and the conceptually complex explanatory texts, plus the inference-level tests employed in these studies, may be important components, because they likely increase the need for concurrent monitoring to produce accurate judgments. We contend that such texts and tests are needed to research metacomprehension, and not just metamemory (Wiley et al., 2005), but they may also be important in creating a context in which concurrent monitoring is important and factors such as reader ability and rereading have room to play a role in successful monitoring.

It is difficult to determine the reasons for prior null results and the failed replications for ability and rereading effects in the literature, given the myriad of potential factors at play. As was noted previously, the materials and ability measures have all varied greatly across studies. In addition, reviews of the literature have revealed problems with many prior studies including unreliable measures (such as single-item tests), which necessarily limits the theoretical implications that can be drawn (for reviews, see Weaver, 1990; Wiley et al., 2005). However, future research that manipulates methodological features such as the nature of the texts and tests that are used would help elucidate the boundaries of ability and rereading effects and would further evaluate the concurrent-processing assumption and the breadth of circumstances under which it holds.

Conclusions

Both rereading and self-explanation improved monitoring accuracy, but in somewhat different ways, suggesting influences on different constraining factors. Our account for these effects presumes that they may be particular to situations in which relatively long texts explain complex causal (or logical) relations about which readers must draw inferences and build a situation model. In such contexts, rereading may afford benefits to some readers who, due to various processing constraints, are struggling to concurrently process the texts and monitor the metalevel cues regarding the quality of their mental representations. In another vein, self-explaining may make monitoring generally more effective by increasing the access and salience of cues that are actually predictive of performance that requires inference- and situation-model-level understanding. The results indicated that both a rereading manipulation that reduced concurrent-processing demands and a self-explanation manipulation that increased access to inference-level comprehension cues improved monitoring accuracy. Notably, the self-explanation instruction produced a very high level of accuracy (.63) comparable to the best levels of metacomprehension monitoring that have been found in the literature thus far (Thiede et al., 2005). Taken together, the results suggest that metacomprehension accuracy is constrained both by the ability of the reader to engage in monitoring during reading and by access to valid cues. Future consideration of how accuracy may depend on individual differences in readers' abilities, or readers' task misconstrual and reliance on suboptimal cues, should be helpful toward understanding why accuracy is notoriously poor and how, when, and for whom accuracy can be improved.

REFERENCES

AIKEN, L. S., & WEST, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

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

BOL, A. S., & HACKER, D. J. (2001). A comparison of the effects of practice tests and traditional review on performance and calibration. *Journal of Experimental Education*, **69**, 133-151.

BOL, A. S., HACKER D. J., O'SHEA, P., & DWIGHT, A. (2005). The influence of overt practice, achievement level, and explanatory style on calibration accuracy and performance. *Journal of Experimental Education*, **73**, 269-290.

CHI, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 5, pp. 161-238). Mahwah, NJ: Erlbaum.

CHI, T. M. H., DE LEEUW, N., CHIU, M., & LAVANCHER, C. (1994). Eliciting self-explanation improves understanding. *Cognitive Science*, **18**, 439-477.

CONWAY, A. R. A., KANE, M. J., BUNTING, M. F., HAMBRICK, D. Z., WILHELM, O., & ENGLE, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, **12**, 769-786.

DUNLOSKY, J., & RAWSON, K. A. (2005). Why does rereading improve metacomprehension accuracy? Evaluating the levels-of-disruption hypothesis for the rereading effect. *Discourse Processes*, **40**, 37-55.

DUNLOSKY, J., RAWSON, K. A., & HACKER, D. J. (2002). Metacomprehension of science text: Investigating the levels-of-disruption hypothesis. In J. Otero, J. A. Leon, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 255-279). Mahwah, NJ: Erlbaum.

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

ENGLE, R. W., KANE, M. J., & TUHOLSKI, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102-134). Cambridge: Cambridge University Press.

FISCHER, P. M., & MANDL, H. (1984). Learner, text variables, and control of text comprehension and recall. In H. Mandl, N. L. Stein, & T. Trabasso (Eds.), *Learning and comprehension of text* (pp. 213-254). Hillsdale, NJ: Erlbaum.

FLAVELL, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, **34**, 906-911.

GARNER, R. (1987). *Metacognition and reading comprehension*. Norwood, NJ: Ablex.

GLENBERG, M. A., & EPSTEIN, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 702-718.

GLOVER, J. A. (1989). Improving readers' estimates of learning from text: The role of inserted questions. *Reading Research Quarterly*, **28**, 68-75.

HACKER, D. J., BOL, L., HORGAN, D. D., & RAKOW, E. A. (2000). Text prediction and performance in a classroom context. *Journal of Educational Psychology*, **92**, 160-170.

KEENAN, J. M., BETJEMANN, R. S., WADSWORTH, S. J., DEFRIES, J. C., & OLSON, R. K. (2006). Genetic and environmental influences on reading and listening comprehension *Journal of Research in Reading*, **29**, 75-91.

KINTSCH, W. (1994). Learning from text. *American Psychologist*, **49**, 294-303.

KINTSCH, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.

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. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *Cambridge handbook of consciousness* (pp. 289-325). New York: Cambridge University Press.

LIN, L., & ZABRUCKY, K. M. (1998). Calibration of comprehension: Research and implications for education and instruction. *Contemporary Educational Psychology*, **23**, 345-391.

MAGLIANO, J. P., MILLIS, K., OZURUB, Y., & MCNAMARA, D. S. (2007). A multidimensional framework to evaluate reading assessment tools. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theory, interventions, and technologies* (pp. 107-136). Mahwah, NJ: Erlbaum.

MAKI, R. H. (1998). 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., & BERRY, S. L. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **4**, 663-679.

MAKI, R. H., HOLDER, E. W., & MCGUIRE, M. J. (2001, November). *Metacomprehension of text: A test of the optimum effort hypothesis*. Paper presented at the 42nd Annual Meeting of the Psychonomic Society, Orlando, FL.

MAKI, R. H., JONAS, D., & KALLOD, M. (1994). The relationship between comprehension and metacomprehension ability. *Psychonomic Bulletin & Review*, **1**, 126-129.

MAKI, R, H., SHIELDS, M., WHEELER, A., & ZACCHILLI, T. (2005). Individual differences in absolute and relative metacomprehension accuracy. *Journal of Educational Psychology*, **97**, 723-731.

MAKI, R. H., & SWETT, S. (1987). Metamemory for narrative text. *Memory & Cognition*, **15**, 72-83.

MCNAMARA, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes*, **38**, 1-30.

MILLIS, K. K., SIMON, S., & TENBROEK, N. S. (1998). Resource allocation during the rereading of science texts. *Memory & Cognition*, **26**, 232-246.

NELSON, T. O. (1984). A comparison of current measure of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, **95**, 109-133.

NELSON, T. O. (1996). Gamma is a measure of the accuracy of predicting performance on one item relative to another item, not of the absolute performance on an individual item. *Applied Cognitive Psychology*, **15**, 395-411.

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: Academic Press.

NELSON, T. O., & NARENS, L. (1994). Why investigate metacognition? In J. Metcalf & A. P. Shimanura (Eds.), *Metacognition* (pp. 1-26). Cambridge, MA: MIT Press.

PERFETTI, C. A. (1985). *Reading ability*. New York: Oxford University Press.

PRESSLEY, M., SNYDER, B. A., LEVIN, J. R., MURRAY, H. G., & GHATALA, E. S. (1987). Perceived readiness for examination performance (PREP) produced by initial reading of text and text containing adjunct questions. *Reading Research Quarterly*, **22**, 219-236.

RAWSON, K. A., & DUNLOSKY, J. (2002). Are performance predictions for text based on ease of processing? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **8**, 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.

ROYER, J. M., CARLO, M., DUFRESNE, T., & MESTRE, J. (1996). The assessment of levels of domain expertise while reading. *Cognition & Instruction*, **14**, 373-408.

SCHNEIDER, W. (1985). Developmental trends in the metamemory–memory behavior relationship: An integrative review. In D. L. Forrest-Pressley, G. E. Mackinnon, & T. G. Walker (Eds.), *Metacognition, cognition, and human performance* (pp. 57-110). Orlando, FL: Academic Press.

THIEDE, K. W., & ANDERSON, M. C. M. (2004, April). *Students' insights into comprehension monitoring*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego.

THIEDE, K. W., ANDERSON, M. C. M., & THERRIAULT, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, **95**, 66-73.

THIEDE, K. W., & DUNLOSKY, J. (1999). Toward a general model of self-

regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 1024-1037.

THIEDE, K. W., DUNLOSKY, J., GRIFFIN, T. D., & WILEY, J. (2005). Understanding the delayed keyword effect on metacomprehension accuracy. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 1267-1280.

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

WEAVER, C. A., BRYANT, D. S., & BURNS, K. D. (1995). Comprehension monitoring: Extensions of the Kintsch and van Dijk model. In C. A. Weaver, S. Mannes, & C. Fletcher (Eds.), *Discourse comprehension: Essays in honor of Walter Kintsch* (pp. 177-193). Hillsdale, NJ: Erlbaum.

WILEY, J., GRIFFIN, T. D., & THIEDE, K. W. (2005). Putting the comprehension in metacomprehension. *Journal of General Psychology*, **132**, 408-428.

WILEY, J., & VOSS, J. F. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology*, **91**, 301-311.

YUILL, N., & OAKHILL. J. (1991). *Children's problems in text comprehension: An experimental investigation*. Cambridge: Cambridge University Press.

NOTES

1. Although results were similar with either Pearson or gamma correlations, Pearsons are reported because the inherent insensitivity of gamma to the magnitude of relative differences produced segmented distributions, fewer unique scores, and more extreme scores, making it inappropriate for regression analyses and undermining most of the benefits for using judgment and performance measures with multiple possible values.

2. Several competing perspectives regarding the specific mechanisms underlying WMC span tasks share similar implications for the impact of WMC on concurrent processing. The present predictions and interpretations presume only that WMC, as measured by span tasks, will relate to concurrent-processing efficiency and will have no implications for the ongoing debates between perspectives that share this presumption.

APPENDIX

Excerpt of Ice Ages Passage

An ice age is a period of time—usually millions or tens of millions of years—when vast glaciers cover as much as a third of the Earth's land surface. Average global temperatures can drop by as many as 12 degrees Celsius overall. The latest Ice Age began about 2.5 million years ago, and ended approximately 15,000 years ago. Average global temperatures decreased by approximately 8 degrees Celsius. Sea-level was lowered substantially due to the amount of water that was frozen in the glaciers. Ice core analysis indicated there were reduced amounts of carbon dioxide in the atmosphere. Giant ice sheets that originated at the North Pole advanced and retreated many times in North America and Europe. The movement of the glaciers coincided with cycles of warm and cold periods in the Earth's temperature. Throughout history, cycles of changes in global temperatures usually occur every 100,000 years or so. Each cycle consists of a long, generally cold period during which the entire Earth cools, followed by a relatively short warm period during which Earth warms up rapidly.

We are now in a warming period that has lasted more than 10,000 years, which is longer than many of the previous warming intervals. Warm temperatures over the last century have been attributed to the increased man-made release of carbon dioxide. CO_2 prevents long-wave radiation from escaping from the Earth into space. The more CO_2 there is in the atmosphere, the more long-wave radiation is kept from leaving the Earth. The more radiation that is trapped, the hotter the Earth becomes. This trapping of radiation works like a gardener's greenhouse, and this phenomenon is commonly known as the "Greenhouse Effect."

Example Inference Verification Questions (and Correct Answers)

Higher levels of CO_2 in the atmosphere lead to higher sea levels. (True)

The Earth is almost completely covered by glaciers during an ice age. (False)