

What does it mean to learn from and understand science text?

Jennifer Wiley, Keith Thiede & Thomas Griffin

Since learning from text is so central to educational practice (especially expository texts in science and social studies instruction), the ability of students to monitor their own learning and understanding from expository texts is important. However, research suggests that readers lack an ability to track their own comprehension, thus fail to make optimal choices in study strategies (Maki, 1998; Thiede, Anderson, & Theriault, 2003).

Why are readers this poor at monitoring their own level of comprehension? How can we approve students' ability to gauge their own learning?

Improving Monitoring Accuracy

The majority of investigations exploring factors that affect the accuracy of metacognitive monitoring have been done using associative learning tasks (e.g., learning word pairs). This research has shown that monitoring accuracy improves when a person monitors learning (a) after a delay rather than immediately after studying an item (Dunlosky & Nelson, 1992, 1997; Kelemen & Weaver, 1997; Nelson & Dunlosky, 1991; Weaver & Kelemen, 1997), (b) after practice monitoring his or her learning (Vesonder & Voss, 1985), (c) after items are actively generated during study rather than passively read during study (Mazzoni & Nelson 1993), and (d) following a practice test of the material (King, Zechmeister, & Shaughnessy, 1980; Lovelace, 1984; Shaughnessy & Zechmeister, 1992). Furthermore, under certain circumstances, people attain near perfect levels of monitoring accuracy. For example, monitoring accuracy, operationalized as the correlation between predicted performance and actual performance, was $+0.93$ in Nelson and Dunlosky (1991).

Monitoring accuracy is far less impressive when the task involves learning from written texts. Glenberg and Epstein (1985) developed a paradigm to assess the accuracy of comprehension monitoring that mirrors the paradigm used in the aforementioned research (using an associative learning task). They had participants read 16 texts. Participants then rated their comprehension of each text and completed a test of reading comprehension for each text. A person's monitoring accuracy was operationalized as the intra-individual correlation between a person's comprehension ratings and his or her test performance; thus, greater monitoring accuracy was indexed by a stronger correlation. A standard term for this predictive accuracy is metacomprehension accuracy (Maki, 1998a).

In sharp contrast to accuracy of monitoring during associative learning tasks, metacomprehension accuracy has typically been quite low. For instance, Glenberg et al. (1987) concluded the following, "Data from our laboratory has almost uniformly demonstrated poor calibration" (p. 120). Likewise, Maki (1998a) reported that the mean intra-individual correlation between comprehension ratings and test performance across 25 studies from her lab was only $+0.27$.

In our research, we have approached this problem by trying to understand what factors may lead to the poor levels of metacomprehension accuracy that are usually observed. The main issue that we have explored is that students may not understand what it means to "comprehend" an expository text. The texts that students learn from in classroom contexts are often expository rather than narrative, and comprehension of expository text typically involves the generation of inferences and the construction of a representation of the text that goes beyond memory for the ideas that were presented. This level of meaning representation has been referred to as the situation model (Kintsch, 1998).

If students do not understand what it means to "comprehend" an expository text, then this is one major factor that may contribute to their poor metacomprehension accuracy. It will cause them to make study judgments based on memory cues instead of comprehension cues, and to read the text with the goal of trying to remember it, rather than trying to understand how or why a phenomenon occurs. Reading for memory of ideas may be a reader's default setting, and it is certainly important for some subject matter and learning contexts. However, if we want students to gain understanding from science text, then it is important to override the "reading for memory" setting. To achieve accurate meta-*comprehension*, we need to clearly indicate to students with specific instruction: that their goal for reading is comprehension, that they will be given tests on their comprehension, and that they should make judgments of their understanding of text in that context. In short, we seek to improve metacomprehension by focusing readers on their situation model since that is an appropriate cue on which to base judgments of metacomprehension. The accuracy of metacomprehension judgments varies as a function of the degree to which the cues used in assessing comprehension empirically correlate with performance on comprehension tests

(Koriat, 1997). If we can make the inferences that readers make about their own understanding more valid, then monitoring accuracy will increase as the available cues more highly correlate with subsequent test performance. Our most recent research, supported by a grant from IES (CASL), has tackled the issue of trying to make the inferences that students make about their own understanding more valid. That is, we have created interventions that make the situation model more accessible to readers and improve metacomprehension accuracy.

Interventions That Focus Readers on Their Situation Model

In our project funded by CASL, we have explored several kinds of interventions designed to make the situation model more accessible to students as they read expository texts: Delayed Keyword/Summary Generation, Self Explanation, and Concept Mapping.

Delayed Keyword and Summary Generation Tasks

Our first line of studies built on previous work done by Thiede and his colleagues with delayed summary and keyword tasks. In several studies, Thiede found that having students engage in delayed generation tasks produced unprecedented levels of metacomprehension accuracy. In Thiede and Anderson (2003), metacomprehension accuracy was higher (around .60 as shown in the first two columns of Table 1) for students who wrote summaries after a delay, and then judged their performance, than for groups who either wrote a summary immediately after reading or wrote no summaries. This result occurred across two experiments (one with shorter texts and one with longer texts).

Table 1. Metacomprehension Accuracy by Group Across Studies

	T&A 2003	TA&T 2003	
	Summaries	Summaries	Keyword
Delayed	.61 (.04)	.63 (.04)	.71 (.04)
Immediate	.25 (.04)	.25 (.04)	.24 (.04)
Control	.27 (.06)	.27 (.06)	.27 (.06)

Note. Entries are mean intra-individual gamma correlations computed across subjects in a group. Values in parentheses are standard error of the means.

Similarly, Thiede et al. (2003) showed that a delayed generation task, where students were asked to a list of five keywords that captured the essence of a text, produced the same boost in metacomprehension accuracy as had generating summaries (see the third column of Table 1). Moreover, the effect was again influenced by the timing of generation. That is, generating keywords after a delay produced dramatic improvement, whereas generating keywords immediately after reading did not affect accuracy compared to a control group.

Linking Delayed-Generation Task Effects to the Situation Model. When assessing comprehension, a person monitors various cues (Koriat, 1997) that are produced by comprehension processes, such as the fluency of text processing (Begg, Duft, Lalonde, Melnick, & Sanvito, 1989; Dunlosky & Rawson, 2005; Rawson & Dunlosky, 1999). A judgment of comprehension is then based on an inference about those cues (Schwartz, Benjamin, & Bjork, 1998). In the case of fluency, people presumably judge that their understanding of a text is better when the text is easily processed than when it is difficult to process (Begg et al., 1989). According to this framework of metacognitive monitoring (Koriat, 1997), the accuracy of comprehension judgments is a function of the degree to which the cues used in assessing comprehension empirically correlate with performance on comprehension tests. More specifically, assuming the inferences are valid, metacomprehension accuracy will increase as the available cues more highly correlate with subsequent test performance.

To discover cues that will be highly predictive of performance on tests of comprehension, it is necessary to understand the comprehension processes that provide many of those cues. According to the construction-integration model of comprehension (Kintsch, 1988), various processes jointly construct a text representation, which may include three levels: a lexical or surface level, a textbase level, and a situation model (see Graesser, Millis, & Zwaan, 1997, for a discussion of others). The lexical level, containing the surface features of the text, is constructed as the words and phrases appearing in the text are encoded. The textbase level is constructed as segments of the surface text are parsed into propositions (Just & Carpenter, 1992), and as links between text propositions are formed based on argument overlap and other factors (Britton & Gulgoz, 1991; Miller & Kintsch, 1980).

The construction of the situation model also involves linking propositions. However, the integration of propositions here involves connecting text information with the reader's prior knowledge (McNamara, Kintsch, Songer, & Kintsch, 1996). It is a person's situation model that largely determines his or her performance on tests of comprehension (Kintsch, 1994; McNamara et al., 1996; Wiley et al., 2005). Therefore, getting people to focus attention on their situation model at the time they judge their comprehension should increase the predictive accuracy of judgments. We hypothesize that generating summaries or keywords for texts after a delay can get people to do just this.

Writing a summary may allow a reader to reflect on how successfully he or she had retrieved information during summarization (cf. the modified feedback hypothesis

described by Glenberg et al. 1987). Accordingly, a text may receive a high rating of comprehension if the person had been able to retrieve a great deal of information about the text during summarization; whereas, a text may receive a low rating of comprehension if the person struggled to retrieve information about the text during summarization.

Further, the timing of summarization is critical. Kintsch, Welsch, Schamlofer, and Zimny (1990) showed that memory for the lexical and textbase level representations of texts decays over time, whereas memory for the situation model is robust to such decay. Thus, when writing a summary immediately after reading, a person may have easy access to the lexical and textbase representations and the person can use this information to generate a summary. However, using this summary as a basis of a judgment of comprehension may not improve accuracy because the performance on the comprehension test and the judgment of comprehension are based on different representations of the text (i.e., performance on the test being determined largely by the situation model, whereas, the summary being determined more by the lexical and textbase representations). In contrast, when writing a summary after a delay, the findings by Kintsch et al. suggest that a person will likely have access primarily to the situation model of a text. Thus, using the summary, written after a delay, as a basis of a judgment of comprehension should improve accuracy because the comprehension judgment and performance on the test are both based of the situation model.

Delayed Generation Tasks and Remedial Readers

In our recent grant, we extended this previous work in several important ways. In one line of research, we explored whether the delayed-summary instruction might be effective in a less-skilled reader population (i.e., college students enrolled in a remedial reading course) as well as more-skilled reader population (i.e., college students enrolled in a psychology course).

Using a within-subjects design, Anderson and Thiede (under review) replicated the previous finding that accuracy was greater in the delayed-summary condition than in the other conditions (see Figure 1). An important extension was that the effect held for typical college students, as well as less-skilled college readers.

We also found that readers focused on different aspects of the summaries in the different summarization conditions. In particular, in the delayed-summary conditions metacomprehension judgments were more strongly correlated with gist idea units than detail idea units; whereas, in the immediate-summary condition metacomprehension judgments were more strongly correlated with detail idea units than gist idea units.

This adds some converging evidence for our theory that delayed conditions are improving access to the situation model, and that this is the mechanism by which metacomprehension accuracy is improved. It also provides an important practical finding that extends the effectiveness of an intervention into a new population.

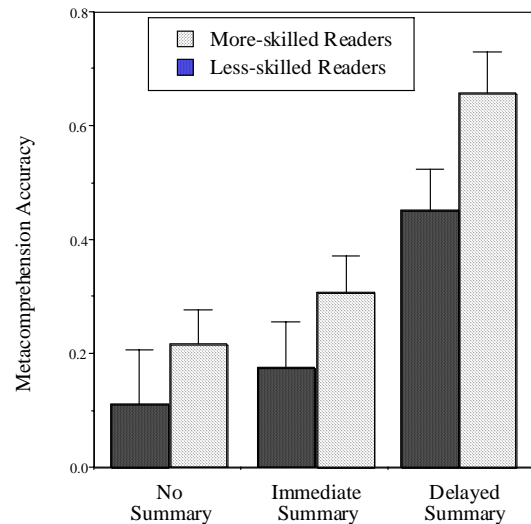


Figure 1: Findings from Anderson & Thiede (under review)

Understanding the Benefits of a Delay

A stronger test of our theory was to determine which delays might be responsible for the delayed keyword effect. In our theorizing, we claim that the critical delay for better metacomprehension would be the one between reading and summary or keyword generation, because this delay causes readers to access their situation model prior to judgment. The generation task forces them to access and use their text representation, and doing so at a delay after reading means that their representation is more likely to constitute a situation model than a surface memory representation that would have decayed during the delay. Although our interpretation fits the findings of Thiede, et al., 2003, there were actually several features of the original study's design and procedure that were confounded. The delayed condition was created by having readers generate keywords for all texts after reading all the texts, rather than immediately after each text. This produced not only a delay between reading and keyword generation, but also decreased the time lag between keyword generation and judgments of comprehension, and decreased the time lag between each keyword generation by having all generation tasks performed in immediate succession. In contrast to the reading-keyword delay, these other two time lags could account for the superior accuracy in the delay group without implying anything about access to the situation model.

In a second line of research under the recent grant, we replicated the Thiede et al. (2003) finding and showed that the delay between reading and generating keywords was critical to improving metacomprehension accuracy (Thiede, Dunlosky, Griffin, & Wiley, 2005). Since altering the order of tasks in the protocol will always affect more than one time lag, we constructed a set of two experiments, where each one evaluated the possible effects of one of the confounding lags. Both studies replicated the reading-generation delay effect, and Experiment 1 showed that the keyword-judgment lag was not critical, while Experiment 2 showed that the lag between keyword tasks was not critical.

In addition, the act of generating keywords not only forces readers to access their text representation, it also simply gets them thinking generally about the text and re-exposes them to the main text ideas via their own keywords they generate. Thus, Experiments 3 and 4 showed that the critical factor was actually generating keywords at a delay and not merely thinking about the text or being re-exposed to key ideas.

This set of 4 experiments provides a compelling case that the act of accessing and making use of one's representation via a generation task at a delay after reading are the critical factors that produced one of the largest improvements in metacomprehension accuracy observed to date. These findings provide strong support for our situation-model approach, because the presumed importance of accessing situation-model cues provides a plausible account of why both the reading-keyword delay and the generation component of the keyword task combine to produce the observed improvements in metacomprehension accuracy.

Self-Explanation Tasks

In a third line of research, we investigated how we might encourage the use of situation models during metacognitive monitoring more directly. In order to put readers in a position to use their situation models as a basis for comprehension judgments, we considered which instructional contexts might promote the construction of situation models. In previous work, Wiley and her colleagues (Wiley & Voss, 1999, Wiley, 2001) found that giving students the instruction to write an explanation or argument of what caused an event, lead to better integration of text material, more causal connections between ideas from the texts in the essay, as well as better recognition of true inferences in post-tests, than when students were instructed to write summaries or narratives about what caused an event. These effects were found both with texts on the Irish Potato Famine and texts on Volcanic Eruptions. Further, explanation and argument instructions were also found to lead to the construction of more advanced models of volcanic activity as compared to instructions to write a summary (Hemmerich & Wiley, 2002), and were found to promote more strategic rereading in an eye-movement analysis (Wiley, 2001).

Chi, DeLeeuw, Chiu, and Lavancher (1994) found similar benefits from a self-explanation reading task. In a study on learning from an expository text about the circulatory system, Chi et al. found that when readers were given general instructions to explain each sentence in terms of what new information it provides, how it relates to previous sentences, if it gives you a new insight on the system, and whether it raised questions, plus specific prompts to explain the functions of components in the circulatory system, this instruction led to better performance on tests of knowledge-based inferences and transfer to other contexts than was found in an unprompted control group of readers. In discussion of both of these studies, the authors suggested that the argument and explanation tasks helped readers to construct and pay attention to the underlying causal, situation models of the subject matter. Attention to the situation model promoted more coherent understanding of the system, and better performance on tests of comprehension.

Based on these previous findings, we hypothesized that a self-explanation instruction might be a way to get readers to focus on their situation models and therefore should improve metacomprehension accuracy (Griffin, Wiley & Thiede, *in press*). To test this, we randomly assigned participants to one of three groups. One group read a text and then self-explained the connections between parts of the text as they read it a second time. Another group read the text twice and a third group read the text once. As seen in Figure 2, the group that had self-explained had significantly better metacomprehension accuracy than the other groups.

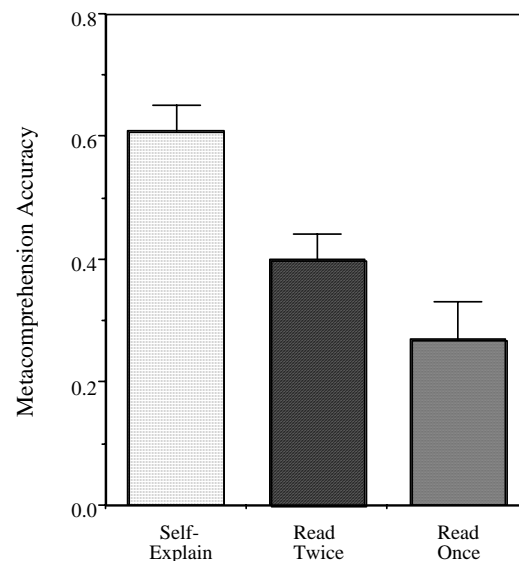


Figure 2: Findings from Griffin, Wiley & Thiede, *in press*

A key difference between this study and previous studies is that the boost in accuracy occurred even without introducing a delay between reading and judging comprehension. However, a practice text and test were necessary in order to get this effect. This suggests that readers were able to self-generate some kind of feedback from their performance on the first text/test, that allowed them to better calibrate on later tasks.

Concept Mapping Tasks

In a fourth line of investigation, we explored whether constructing concept maps would help readers to focus on their situation model. Weinstein and Mayer (1986) suggested that instructing students to create concept maps of texts during reading helps them to identify the connections among concepts in a text, which improves comprehension. In discussion of these studies, the authors suggest that argument, self-explanation, and concept mapping tasks help readers to construct and pay attention to the underlying causal, situation models of the subject matter. The advantage that these tasks have over a task like generating summaries or keywords is that these tasks create contexts where readers may be able to attend to their situation models on the first reading of the text. Since the text is available during the activity, concept mapping tasks may be especially appropriate for less-skilled, less-able or younger readers who may have difficulty engaging in a task that requires them to remember a text (Stensvold & Wilson, 1990). Hence, these tasks are more practical for meaningful classroom use, and may be especially appropriate for less-skilled readers.

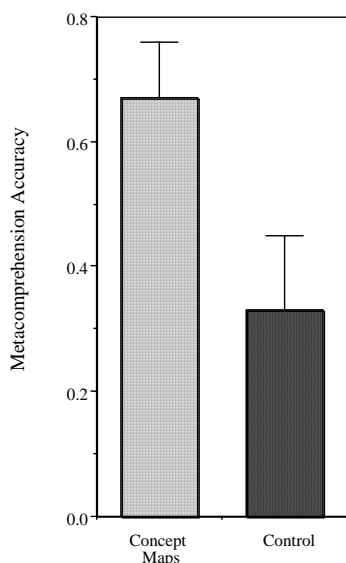


Figure 3: Data from Thiede, Anderson, Griffin & Wiley, under review

Using a within-subjects design, we (Thiede, Anderson, Griffin, & Wiley, under review) gathered pretest

metacomprehension data by running participants through the standard experimental procedure for assessing metacomprehension accuracy. That is, they read a series of texts, judged their comprehension of each text, and then completed a comprehension test for each text. Participants, who were 21 less-skilled college readers (i.e., college students in a remedial reading course), then received eight days of instruction on how to construct a concept map for a text with substantial feedback on their mapping. They then completed the procedure again, but this time constructed concept maps while reading the texts.

As seen in Figure 3, constructing concept maps improved metacomprehension accuracy. We interpreted these findings as further support that getting people to focus on their situation model when judging comprehension will improve metacomprehension accuracy. Again, the boost in accuracy occurred without introducing a delay between reading and judgments.

Interestingly, when we tested an abbreviated version of this intervention with college students of normal reading ability, it did not improve metacomprehension accuracy. These results are consistent with a recent meta-analysis that found benefits of concept mapping were localized mainly among low ability and younger populations (Nesbit & Adesope, 2006). This differential result further emphasizes the need to explore the effectiveness of each intervention in specific student populations. But, they could also suggest that practice and feedback on self-explanation or concept mapping may be critical to maximizing their benefits.

Support for the Situation Model Approach

In sum, over the past several years we have discovered a variety of ways of improving metacomprehension accuracy of more-skilled and less-skilled college readers. All interventions that we tested were designed to focus readers on appropriate cues for judging their comprehension, and to direct their attention to the situation model. Consistent with our hypotheses based on this approach, the studies that have been conducted have demonstrated unprecedented improvement in metacomprehension accuracy.

We have explored answers to these questions using theories of text comprehension, and the hypothesis that readers tend to default to monitoring surface memory for text instead of deeper levels of comprehension (Wiley, Griffin & Thiede, 2005). Kintsch (1998) among others views the act of comprehension as occurring on three levels. The first two levels involve merely forming a memory representation of the surface words and exact form and of the semantic meaning of individual propositions. Only in the third and deepest level of

representation, the situation model, are important connective and causal inferences generated via integration of multiple text propositions and their implicit relations, and via integration of text ideas with prior knowledge. It is the creation of the situation model that can be seen as the process of deeply understanding a text, or comprehension, and determines whether new information can be used in novel contexts (Kintsch, 1994).

So if the student is given a text about blood circulating through the heart, then the situation model should capture the process of blood flow, and a reader with a good situation model should be able to recreate and explain the dynamic process of blood flow that was described by the text. This is the level of understanding that we want students to monitor when they make metacomprehension judgements. However, left uninstructed, students are more likely to consider their memory for the text as a basis for their monitoring judgments.

In addition, we are currently exploring how giving students practice tests, which provides them with knowledge of the kinds of questions that may be asked about the texts, may offer important feedback to students and help them to attune their judgments to the correct cues for accurately assessing their comprehension. All of these studies have been run in group contexts which shows potential for classroom application. We are also currently moving these interventions into middle-school science class contexts to test if they remain effective in an authentic instructional environment. The gains we have observed in metacomprehension accuracy thus far are among the largest ever found. These are promising results as the ultimate value of training students what it means to understand science text is that it can lead to better study behaviors and ultimately better learning in science.

REFERENCES

- Anderson, M. C. M. & Thiede, K. W. (under review). Why Do Delayed Summaries Improve Metacomprehension Accuracy?
- 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.
- Britton, B. K., & Gülgöz, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology*, 83, 329-345.
- Chi, M. T. H., DeLeeuw, N., Chiu, M., & Lavancher, C. (1994). Elicit self-explanation improves understanding. *Cognitive Science*, 18, 439-477.
- Dunlosky, J. & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition*, 20, 374-380.
- Dunlosky, J., & Nelson, T. O. (1997). Similarity between the cue for judgments of learning (JOL) and the cue for test is not the primary determinant of JOL accuracy. *Journal of Memory and Language*, 36, 34-49.
- 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.
- Glenberg, A. M., & Epstein, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 702-718.
- Glenberg, A. M., Sanocki, T., Epstein, W., & Morris, C. (1987). Enhancing calibration of comprehension. *Journal of Experimental Psychology: General*, 116, 119-136.
- Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology*, 48, 163-189.
- Griffin, T. D., Wiley, J., & Thiede, K. W. (in press). Resource-availability and misdirected monitoring as sources of metacomprehension failure. *Memory & Cognition*.
- Hemmerich, J. & Wiley, J. (2002) Do argumentation tasks promote conceptual change about volcanoes. *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122-149.
- Kelemen, W. L., & Weaver, C. A. (1997). Enhanced memory at delays: Why do judgments of learning improve over time. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 1394-1409.
- King, J. F., Zechmeister, E. B., & Shaughnessy, J. J. (1980). Judgments of knowing: The influence of retrieval practice. *American Journal of Psychology*, 93, 329-343.
- Kintsch, W. (1994). Learning from text. *American Psychologist*, 49, 294-303.
- Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language*, 29, 133-159.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: 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.

- Lovelace, E. A. (1984). Metamemory: Monitoring future recallability during study. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*, 756-766.
- Maki, R. H. (1998a). Test predictions over text material. In Hacker, D. J., Dunlosky, J., & Graesser, A. C. (Eds.). *Metacognition in Educational Theory and Practice*. (pp. 117-144). Hillsdale, NJ: LEA.
- McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, *14*, 1-43.
- Miller, J. R., & Kintsch, W. (1980). Readability and recall of short prose passages: A theoretical analysis. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 335-354.
- Nelson, T. O. (1984). A comparison of current measures of feeling-of-knowing accuracy. *Psychological Bulletin*, *95*, 109-133.
- Nelson, T. O. & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed JOL effect." *Psychological Science*, *2*, 267-270.
- Nelson, T. O., Dunlosky, J., Graf, A., & Narens, L. (1994). Utilization of metacognitive judgments in the allocation of study during multitrial learning. *Psychological Science*, *5*, 207-213.
- Nelson, T. O. & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation*, *26*, pp. 125-141. New York: Academic Press, Inc.
- Nesbit, J. C. & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research*, *76*, 413-448.
- 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., & Dunlosky, J., & Thiede, K. W. (2000). The rereading effect: Metacomprehension accuracy improves across reading trials. *Memory & Cognition*, *28*, 1004-1010.
- Schwartz, B. L., Benjamin, A. S., & Bjork, R. A. (1998). The inferential and experiential bases of metamemory. *Current Directions in Psychological Science*, *6*, 132-137.
- Shaughnessy, J. J. & Zechmeister, E. B. (1992). Memory monitoring accuracy as influenced by the distribution of retrieval practice. *Bulletin of the Psychonomic Society*, *30*, 125-128.
- Stensvold, M. S. & Wilson, J. T. (1990). The interaction of verbal ability with concept mapping in learning from a chemistry laboratory activity. *Science Education*, *74*, 473-480.
- Thiede, K.W. (1999). The importance of accurate monitoring and effective selfregulation during multitrial learning. *Psychonomic Bulletin & Review*, *6*, 662-667.
- Thiede, K. W. & Anderson, M. C. M. (2003). Summarizing can improve metacomprehension accuracy. *Contemporary Educational Psychology*, *28*, 129-160.
- 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., Anderson, M. C. M., Wiley, J. & Griffin, T. D. (under review). Does cue use influence the accuracy of metacomprehension judgments?
- 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, and 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
- Vesonder, G. T. & Voss, J. F. (1985). On the ability to predict one's own responses while learning. *Journal of Memory and Language*, *24*, 363-376.
- Weinstein, C.E. & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock (Ed.), *Handbook on Research in Teaching* (3rd ed., pp. 315-327). New York: Macmillan.
- Weaver, C. A. & Kelemen, W. L. (1997). Judgments of learning at delays: Shifts in response patterns or increased metamemory accuracy? *Psychological Science*, *8*, 318-321.
- Wiley, J. (2001) Supporting understanding through task and browser design. *Proceedings of the Twenty-third annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.
- Wiley, J., Griffin, T. & 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.

ACKNOWLEDGEMENTS

This research was supported by grant number R305H030170 from the Institute of Educational Sciences, Cognition and Student Learning Program, to Keith W. Thiede and Jennifer Wiley. Any opinions, conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect those of the funding organization.