To Understand Your Understanding, You Must Understand What Understanding Means

Jennifer Wiley (jwiley@uic.edu)
Department of Psychology, 1007 W. Harrison Street (M/C 285)
Chicago, IL 60607 USA

Thomas D. Griffin (tgriffin@uic.edu)
Department of Psychology, 1007 W. Harrison Street (M/C 285)
Chicago, IL 60607 USA

Keith W. Thiede (keiththiede@boisestate.edu)
College of Education, (mail stop 1745), 1910 University Drive
Boise, ID 83725 USA

Abstract

Although critical for the regulation of many reading and studying behaviors, metacomprehension accuracy is generally observed to be quite low. The present research examined how metacomprehension accuracy would be affected by practice tests designed to give readers expectations about the kind of tests they would be given, and self-explanation instructions to give readers access to valid cues for their metacomprehension judgments. Both manipulations improved readers' ability to accurately judge their own level of comprehension for expository texts.

Keywords: metacomprehension, text comprehension, learning from text, testing effects.

Monitoring Understanding of Text

Reading text is a primary means by which people learn new information. However, a great deal of research has shown that readers lack an ability to track their comprehension of expository texts. Metacomprehension accuracy is defined as the ability to predict how well one will do on a test of comprehension after reading a text. Although this is a critical skill for the regulation of many reading and studying behaviors, the typical finding from research on metacomprehension is that accuracy is generally quite low. Typically correlations between predictive judgments and test performance hover around .27 (Dunlosky & Lipko, 2007; Maki, 1998). Further, it has been shown that as a result of poor metacomprehension accuracy, readers fail to make optimal decisions about what to re-read (Maki, 1998; Thiede, Anderson, & Therriault, 2003).

Why are readers so poor at monitoring their own level of comprehension? If students do not understand what it means to "comprehend" an expository text, or what a test of comprehension will be like, then this may be one major factor that could contribute to poor metacomprehension accuracy. It could cause readers 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 text 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 expository texts, then it is important to teach them how to override the "reading for memory" setting. To achieve accurate metacomprehension, comprehension goals need to be clearly signaled. Readers need to be aware that they will be given tests on their comprehension, and that they should make judgments of their understanding of text in that context. Further, readers need to use appropriate cues to judge whether or not they have understood a text. In short, the present studies explore whether we can improve metacomprehension accuracy by clarifying what is meant by comprehension, and by making appropriate cues based in situation models more salient. If we can make readers attend to both comprehension and relevant cues, the inferences that they make about their own understanding should become more valid.

Kintsch (1998) among others views the act of comprehension as occurring on multiple levels. The first level (the surface model) involves forming a memory representation of the exact words that are read, while the next level (the textbase model) encodes the semantic meaning of individual propositions. Only at the deepest level of representation, the situation model, are important connective and causal inferences generated via integration of multiple text propositions with each other as well as with prior knowledge. Thus, it is the creation of the situation model that can be seen as the process of deeply understanding a text, and it is the quality of the situation model representation that determines whether new information will be used in novel contexts including tests of comprehension (Kintsch, 1994; McNamara, Kintsch, Songer, & Kintsch, 1996).

For example, if a reader 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. If this is the kind of knowledge that is tested on the comprehension test, then this is the level of

understanding that we want students to monitor when they make metacomprehension judgments (Wiley, Griffin & Thiede, 2005). However, under normal circumstances students are more likely to consider their memory for the text as a basis for their monitoring judgments (Thiede, Griffin, Wiley & Anderson, *in press*). In a sample of 87 undergraduates, only 7 spontaneously reported using explanation-based cues to make their judgments. The other cues that students reported using included their memory for the text or exact words, their prior knowledge or interest in the topic, and how difficult the text seemed to read. All of these cues may sometimes relate to comprehension test performance. For example, cues based in surface memory may predict performance on some memory-based tests quite well, but will tend to be less predictive when tests tap inferences or understanding of text. Thus, the question at hand is how we can get readers to use appropriate cues, such as whether or not they can explain the phenomena they read about, as a basis for their judgments of metacomprehension.

The emphasis on the situation-model-level representation as the source of valid cues for metacomprehension has been called the *situation model approach* to improving metacomprehension (Griffin, Wiley & Thiede, 2008; Thiede, Dunlosky, Wiley & Griffin, 2005; Thiede, et al., *in press*; Thiede, Griffin, Wiley & Redford, *in press*; Wiley, et al., 2005). Several studies have provided support for the *situation model approach* by providing particular instructional contexts that are thought to give students access to more predictive cues (i.e. cues at the level of the situation model). Successful interventions have included giving self-explanation instructions (Griffin, et al., 2008) and concept mapping tasks (Thiede, et al., *in press*) to focus students on the quality of the connections and explanations they are constructing. Improvements have also been found with delayed summarization and keyword tasks (Thiede, Anderson & Therriault, 2003; Thiede, et al, 2005). These are thought to be effective because after a delay readers are less likely to rely on surface cues as a basis for judgments since such information decays quickly over time (c.f Kintsch, Welsh, Schmalhofer, & Zimny, 1990). Each of these interventions has been shown to improve students' ability to judge their own understanding to around the .60 range or beyond, and represent a marked improvement from typical metacomprehension accuracy levels around .27 (Dunlosky & Lipko, 2007; Maki, 1998).

The research discussed above has improved metacomprehension accuracy by increasing the salience of cues related to the situation model, so readers are more likely to use those cues to judge their comprehension. Interventions such as delayed generation, self-explanation, and concept mapping were all secondary tasks readers had to perform that happened to provide readers with greater and more salient access to cues about the quality of their situation-model representations. Because readers were also given comprehension tests that tapped the situation model of a text, these interventions improved the alignment between metacomprehension judgments and performance on the tests of comprehension.

An additional constraint on metacomprehension accuracy is readers' lack of understanding about the kinds of questions that are likely to appear on comprehension tests. Even if readers could access situation-model-based cues on their own and without aid of a secondary task, they would likely fail to do so, unless they had explicit knowledge and understanding that the tests to be given after reading would require situation-model level comprehension.

The goal of the present research was to test how metacomprehension accuracy would be affected by giving readers expectations about the kind of tests they would be given. In contrast to other studies that have directly provided readers with greater access to valid cues through interventions at the time of processing or judgment, expectation manipulations provided readers with knowledge about the level of comprehension they would be tested on from the outset. Thus, in the current studies, we manipulated the expectations that students had about the tests they would be given, to see how providing readers with knowledge of the kinds of questions they would be asked might help students attune their judgments to the correct cues for accurately assessing their comprehension.

In particular, we attempted to make it clear to learners that they needed to make judgments that predicted their performance on inference-based tests of comprehension. To do this, some readers were given an explicit instruction that informed them that they would need to "take tests based on their ability to make connections across different parts of the text" and then were also given practice test questions that required the verification of inferences from the text. Another group of students were told that they would be tested on their ability "to remember specific details of the text" and received practice tests with questions that required memory for details of the text. Readers made judgments of how well they understood each text and then were given both memory and inference tests on each text. We expected a testing effect such that by providing readers with both explicit instructions and specific examples of comprehension tests, they should be able to direct their own attention to valid cues, and this should improve metacomprehension accuracy. Metamemory accuracy (i.e. the ability to judge performance on memory-based tests), however, might suffer. On the other hand, memory-focused instructions and practice tests should harm metacomprehension accuracy, although this condition could improve metamemory accuracy.

Experiment 1

In previous research, it has been demonstrated that metacomprehension accuracy can be improved by providing contexts that get readers focused on their situation model when judging comprehension. Prior work has not, however, examined whether readers can achieve this focus on the situation model on their own, given information that they should do so via instructions and practice tests. To evaluate

this possibility, changes in metacomprehension accuracy were explored using a standard test-expectancy paradigm where different groups were given different expectancies before reading and making judgments, but the same sets of tests.

Method

Participants. Participants were 108 undergraduates who received course credit as part of an introductory psychology subject pool.

Materials. The texts were nine explanatory texts that each described complex causal phenomenon from the natural or social sciences (i.e., Antibiotic use causing allergies, Biological evolution, Volcano formation and eruption, Racial differences on I.Q. tests, Ice ages, Monetary policy, Cheesemaking, Lightning formation, the Scientific Method). 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. For each text, one 5 item multiple-choice test was created with detail questions, and a second 5 item multiple-choice test was created with inference questions. Detail questions referred to specific ideas, in exact surface form, that appeared in a single sentence of the text. Inference questions tapped connections across sentences.

Design. The design was a 3 (expectancy: memory, comprehension, none) x 2 (type of test on critical texts: inference, memory) mixed design. The type of test given on the six critical texts was a within-subjects variable.

Procedure. Participants were given a general summary of the tasks and their order. Each group of participants read the set of three practice texts followed by the set of six critical texts. The memory group was told that they would be tested on their memory of specific details for each text. They read the first practice text, and immediately made their judgment based on the question "How many items do you think you will get correct on a five item test?" They were then given a five item test of memory for details. They did the same for the other two practice texts. Following practice, they read and made judgments each of the six critical texts. After making the last judgment, they completed the first set of tests. For the critical texts, readers completed BOTH the memory test and the inference test for each text. The tests were divided into two blocks and the order of the test type was counterbalanced, with half of the participants in each expectancy condition receiving the all tests of the expected test type first and then all tests of the other type; and vice versa for the other half of the participants.

The comprehension group completed a similar procedure to the memory group, except participants were told that they would be tested on "their ability to make connections across different parts of the text" and they were given inference tests for practice texts.

The no-expectancy group was only told that they would be "taking a test" for the critical texts with no indication of the nature of the questions they would receive. They read and judged each of the practice texts, but did not receive either practice memory or inference tests.

Results and Discussion

Following the standard practice of analysis for metacomprehension studies (Maki, 1998), intra-individual correlations between predicted and actual performance were computed as measures of monitoring accuracy (see Griffin, et al., 2008 for reasons why Pearson correlations were used as opposed to Gamma correlations). The correlation between predictive judgments and memory test performance across the six texts for each individual represents a measure of relative metamemory accuracy. The correlation between predictive judgments and inference test performance across the six texts for each individual represents a measure of relative metacomprehension accuracy. Figure 1 shows the average correlations that were found in each condition. A repeated-measures ANOVA revealed no main effects for test type or expectation condition, Fs < 1. However, there was a significant test type X expectation interaction, $F(2,102)=3.33$, $MSE=.111$, $p<.04$ where expectations selectively improved monitoring accuracy in expectancy-congruent conditions.

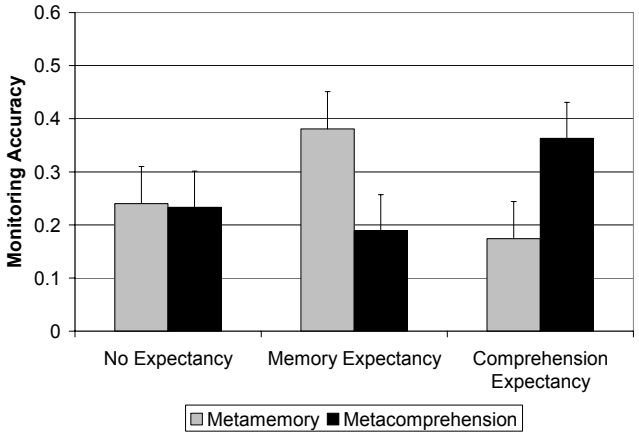

Figure 1: Monitoring Accuracy by Expectancy Condition for Experiment 1

The no-expectancy condition led to typical levels of poor accuracy, similar the .27 correlation noted in the literature. Also, when the test did not match expectations, metacomprehension accuracy was also poor. However, when participants were given expectations about the kind of test they would receive, and the tests they were given actually matched those expectations, their monitoring accuracy improved. The results suggest that part of the reason for poor monitoring accuracy in general may be due to a lack of clear expectations about the nature of the upcoming tests. Readers need specific information about

what sort of test they are preparing for, and when they have this information, their monitoring accuracy improves.

While the above result is encouraging, the improvements seen here, especially in metacomprehension accuracy which is the primary focus of this research, are relatively modest. Assuming there are multiple sources of difficulty preventing readers from engaging in accurate metacomprehension, Experiment 2 takes a combined approach where both expectations and the availability of relevant cues were manipulated.

Experiment 2

In this experiment, the goal was to maximize metacomprehension accuracy by combining complementary manipulations from previous research. In particular, previous work has demonstrated that self-explanation during reading gives readers access to valid cues for metacomprehension judgments, and can improve metacomprehension accuracy (Griffin, et al., 2008). Importantly, this prior study used a re-reading comparison condition in order to rule out increased time on task as an alternative explanation for self explanation effects.

In Experiment 2, a self-explanation instruction was combined with the inference-based test expectancy manipulation from Experiment 1. The idea here is that explicit instructions and practice tests should clarify that metacomprehension judgments should be based on the ability to connect ideas of texts, while self-explaining should facilitate formation of the situation model and make cues about this level of comprehension more accessible and salient when judging comprehension.

Method

Participants. Participants were 144 undergraduates who received course credit as part of an introductory psychology subject pool.

Materials. The texts and tests were the same as used for Experiment 1.

Design. The design was a 2 (comprehension expectancy, none) x 2 (self-explanation, none) x 2 (type of test on critical texts: inference, memory) mixed design. The type of test given on the critical texts was a within-subjects variable.

Procedure. Participants in the self-explanation and combined self-explanation/comprehension expectation conditions were given an additional instruction (based on Chi, 2000; Griffin, et al., 2008; and McNamara, 2004). They were told "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 question 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. In all other respects, the procedure was the same as in Experiment 1.

Results and Discussion

Figure 2 shows the pattern of Pearson correlations that were found in each condition. A repeated-measures ANOVA revealed a main effect for test type, as metacomprehension accuracy was generally greater than metamemory accuracy in this experiment, $F(1, 140)=5.09$, $MSE=.146$, $p<.02$. The main effect for comprehension expectancy was also significant, $F(1, 140)=5.17$, $MSE=.244$, $p<.03$. There was no main effect for self explanation. However, there were two significant two-way interactions. Self-explanation instructions interacted with test type, $F(1, 140)=9.52$, $MSE=.146$, $p<.01$, as did comprehension expectancy, $F(1, 140)=5.28$, $MSE=.146$, $p<.02$. Both manipulations separately led to higher metacomprehension accuracy and lower metamemory accuracy.

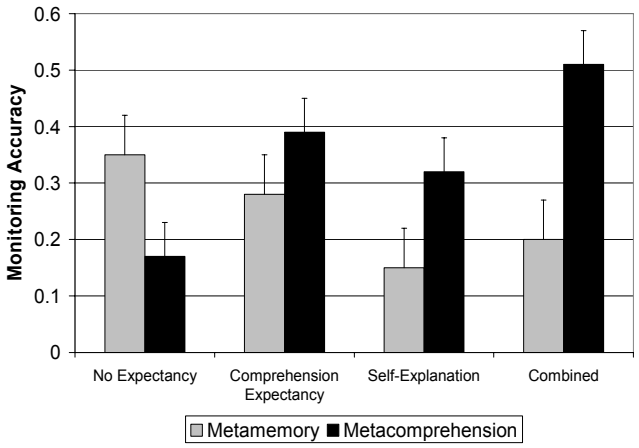

Figure 2: Monitoring Accuracy by Expectancy and Self Explanation Condition for Experiment 2

As a result, the best performance was seen the combined self-explanation and comprehension expectancy condition. The three-way interaction was not significant, suggesting the effects of self-explanation and test expectancy were additive. Readers benefited from both manipulations. Practice tests gave readers the expectancy that tests would tap their ability to generate and recognize inferences based on the text, and self-explanation instructions gave readers valid cues to judge their ability to perform well on these tests.

General Discussion

Recently, there has been a great deal of attention focused on *testing effects* and how they may lead to better learning

outcomes. In this emerging literature, researchers have demonstrated that taking a test on studied material promotes better remembering of that material on a final test, even when compared to students who spent additional time studying the target material (Roediger & Karpicke, 2006). These effects are presumably due to the act of recalling information from memory, which helps to cement the information to memory and thereby reduces forgetting (Carpenter, Pashler, Cepeda & Alvarez, 2007). By answering questions on quizzes, the student practices the act of recalling specific information from memory which improves the chances of retrieval on future tests.

The present research program also aims to use testing to ultimately improve study behaviors, but our approach differs from other testing approaches in two very important ways. First, in this research, we are concerned with the comprehension of phenomena from text instead of the acquisition of isolated facts. Second, all of the previous work on testing effects has been examining repeated test performance on the same content information. Even if the exact items differ from one test to another, in typical testing effect paradigms it is the same content that is being tested across different testing occasions. In our studies, however, the practice tests and target tests were on different topics. The goal behind our practice tests was to inculcate what we meant by comprehension in readers' minds. The practice tests were intended to give readers a sense of what they could expect on later tests. Even though our version of a "testing effect" paradigm is different than those used by others, we still observed striking consequences for having been exposed to a set of preliminary tests. The results showed that the practice tests did set up an expectation for readers, and that readers were able to transfer this sense of "comprehension" to judgments of their understanding on new topics. An interesting question for future research is whether these test expectancy and "transfer" effects on metacomprehension accuracy can be demonstrated with test formats other than multiple choice.

The addition of a self-explanation instruction, in combination with the practice tests, supported an even better understanding of understanding. Self-explanation is another paradigm that has received a good deal of recent attention. Self-explanation instructions have been shown to promote better learning among students reading expository text (Chi, 2000; McNamara, 2004), and this may be for many reasons. Self-explanation may encourage more elaborate, constructive, or extensive processing of information than other activities such as re-reading or summarization tasks. Any of these processes can serve to generate useful metacognitive cues. Our research suggests yet another advantage of self-explanation instructions is that they also support better metacomprehension by making the appropriate cues for judging understanding more salient (Griffin, et al., 2008).

In combination, both practice tests and self-explanation instructions helped readers to develop a better understanding of what it means to understand a text. Exploring the long-term effects of such interventions on learning is the next step. Ultimately, this new sense of understanding should allow readers to engage in better self-regulated learning as they attempt to comprehend information from expository texts.

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Cognition and Student Learning Grants R305H030170 and R305B07460 to the authors. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

References

Carpenter, S.K., Pashler, H., Cepeda, N.J., & Alvarez, D. (2007). Applying the principles of testing and spacing to classroom learning. In D.S. McNamara and J.G. Trafton (Eds.), Proceedings of the 29th Annual Cognitive Science Society, (p. 19). Nashville, TN: Cognitive Science Society.

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* (161-238). Mahwah, NJ: Lawrence Erlbaum Associates.

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

Griffin, T. D., Wiley, J. & Thiede, K. W. (2008) Individual Differences, Rereading, and Self-Explanation: Concurrent Processing and Cue Validity as Constraints on Metacomprehension Accuracy. *Memory & Cognition, 36,* 93-103.

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

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

Kintsch, W., Welsch, D., Schmalhofer, F. & Zimny, S. (1990). Sentence memory: Theoretical analysis. *Journal of Memory and Language, 29*, 133-159.

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: LEA.

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

McNamara, D. S., Kintsch, E., Butler-Songer, N., and 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.

Roediger, H.L. & Karpicke, J.D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249-255.

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

Thiede, K. W., Griffin, T. D., Wiley, J., & Anderson, M. (in press) Poor Metacomprehension Accuracy as a Result of Inappropriate Cue Use. *Discourse Processes*.

Thiede, K. W., Griffin, T. D., Wiley, J., & Redford, J. (in press) Metacognitive Monitoring During and After Reading. To appear in D. J. Hacker, J. Dunlosky & A. C. Graesser (Eds). Handbook of Metacognition in Education. Routledge.

Wiley, J., Griffin, T. D., & Thiede, K. W. (2005). Putting the comprehension in metacomprehension. *Journal of General Psychology*, *132(4),* 408-428.