

Metacomprehension for educationally relevant materials: Dramatic effects of encoding–retrieval interactions

AYANNA K. THOMAS

Colby College, Waterville, Maine

AND

MARK A. MCDANIEL

Washington University, St. Louis, Missouri

As the metacomprehension literature has grown, important discoveries pertinent to education have been made. For example, as students are better able to assess their knowledge and implement appropriate study strategies, presumably their acquisition and retention of course material improves. Accordingly, we consider the metacomprehension literature with an emphasis on factors that impact metacomprehension accuracy. Several studies have demonstrated that metacomprehension prediction accuracy will improve to the extent that people engage in enriched-encoding activities. More recently, research by Thomas and McDaniel (in press) has suggested that enriched-encoding manipulations interact with retrieval to impact both retention and metacomprehension and, in turn, the effectiveness of controlling subsequent study. Thus, matching enriched-encoding activities with the criterial test plays a critical role in metacomprehension accuracy and control of studying.

Learning and memory have been central topics in human experimental psychology. Not surprisingly, human memory is a fundamental concern of our education system, since students are required to commit to memory vast quantities of information. Much of human memory research over the past 30 years has focused on the conditions that lead to optimal retention. More recently, researchers have focused attention on the role that metacognitive processes play in the mediating of memory and learning. The goal of the present article is to examine selected findings and themes in the metacognitive literature that appear to have direct implication for education.

Two metacognitive components that presumably are fundamental for learning in the education setting are (1) the online monitoring of what one knows, which provides a basis for predictions of subsequent retention, and (2) the control processes that determine subsequent study activity (Nelson & Dunlosky, 1991). Students must constantly make decisions about acquisition of material in order to study effectively. Thus, learning is a dynamic process in which a learner monitors progress toward a learning goal and uses this information to regulate study (e.g., Metcalfe & Kornell, 2003; Nelson & Narens, 1990; Thiede & Dunlosky, 1999; Winne & Hadwin, 1998).

In the present article, we examine metacognitive processes associated with educationally relevant text materials. Metacognition applies to education insofar as

the assessment of knowledge and implementation of appropriate study strategies has important implications not only for test performance, but also for the student's ability to articulate problems in his or her understanding. We will refer to these metacognitive processes for educationally relevant material as *metacomprehension*. Formally defined, metacomprehension is the processes of monitoring the online learning of text material (Maki & Berry, 1984). For the purposes of the present article, metacomprehension accuracy will be examined by comparing predictions about future memory to actual retention of that material.¹

Early research into metacomprehension showed that college students in general performed no better than chance in predicting future memory performance associated with studied text material (Glenberg & Epstein, 1985; Glenberg, Sanocki, Epstein, & Morris, 1987). Maki and Berry (1984) found accuracy of metacomprehension predictions to be limited to better students. Although the evidence is correlational, this finding is consistent with the possibility that able students are those who can accurately assess what they have learned and can then implement effective control processes to acquire the material not yet learned. In a parallel vein, less able students' low test performances may be due, in part, to an inability to assess what information has been adequately learned during exposure to the target material and what knowledge is important. Accordingly, one

A. K. Thomas, ayanna.thomas@tufts.edu

line of work that has application to education is research illuminating factors that improve metacomprehension.

Improving Metacomprehension

Recently, researchers have demonstrated that under specific circumstances, students demonstrate good metacomprehension ability, with predictions of future test performance more closely representing actual test performance. Methodological parameters influencing metacomprehension accuracy include how the prediction question is posed (Weaver & Bryant, 1995), how test questions are asked (Maki, Foley, Kajer, Thompson, & Willert, 1990), and how the test question is scored (Maki et al., 1990; Weaver 1990).

Metacomprehension improves with enriched encoding. More relevant to education, enriched encoding seems to play an important role in metacomprehension accuracy. For example, research has demonstrated that metacomprehension improved after participants summarized text material and when participants generated key words after reading text material (Thiede & Anderson, 2003; Thiede, Anderson, & Therriault, 2003; Thiede, Dunlosky, Griffin, & Wiley, 2005). Additional enriched-encoding tasks that have been shown to affect metacomprehension accuracy are illustrated in Table 1.

Metacomprehension and desirable difficulty. Like enriched-encoding manipulations, increasing the difficulty of initial encoding has been shown to impact subjective experiences associated with the learning of new material. Challenges experienced by the learner at encoding also have been shown to improve retention. These findings have been categorized as *desirable difficulties*, a term Bjork (1994) used to describe principles for designing instruction that make learning *seem* more difficult during acquisition but lead to increased retention. With regard to the retention of educationally relevant materials, researchers have found that participants show better retention after generating material than after simply reading the material (Einstein, McDaniel, Owen, & Coté, 1990; McNamara & Healy, 1995; and see Metcalfe & Kornell, 2007, for an educational application). With regard to subjective monitoring, deWinstanley and Bjork (2004) showed that participants were able to benefit from the experience of generation and, thereby, improve learning to the level stimulated by generation even for read items. This effect is contingent on participants' ability to recognize the ben-

efits of generation (through monitoring) and to implement internal generation on read trials (through control).

Although these apparent challenges during encoding have been shown to positively affect memory and metacomprehension in laboratory settings, the relationship between these manipulations and the retention of educationally relevant material may be far more complex. For example, when educators design study materials with embedded questions that require student responses, they find that conceptual comprehension improves (Chi, 2000; Davis & Linn, 2000; Palincsar & Brown, 1984); however, with these learning materials, the role of conceptual tests that tap complex relational and integrative processing is emphasized. In contrast, there are numerous situations in which students are held accountable for knowledge of unconnected item-specific information (e.g., objective examination questions in large introductory college courses, preparation for standardized testing, or second-language acquisition). In these cases, the processing stimulated by desirable difficulties could be incongruent with the information targeted in the criterial tests.

A key and novel question raised by the observation above is whether educators simply need to utilize enriched-encoding manipulations or promote desirable difficulties to improve metacomprehension. Appealing to existing findings, the answer to this question seems to be yes. Tasks that have been shown to facilitate memory the most are those that create a durable representation in memory and flexible access to the to-be-remembered information. Tasks shown to facilitate the greatest memory benefits are those that force students to process information in multiple ways, creating several routes of access to that information. These tasks also seem to improve metacomprehension accuracy in the laboratory. For example, generation has been shown to improve retention, monitoring, and control (deWinstanley & Bjork, 2004). However, we next will present very recent work that has revealed that the overlap between the encoding stimulated by desirable difficulties and the information targeted by the test is a critical factor for determining the positive effects of desirable difficulties on metacomprehension.

Transfer-Appropriate Processing and Metacomprehension

Basic cognitive research over the last 25 years has demonstrated the mnemonic advantages of emphasizing cog-

Table 1
Enriched Encoding Improves Metacomprehension

Type of Encoding	Study
Delayed summaries	Anderson & Thiede (2007)
Diagrams	Cuevas, Fiore, & Oser (2002)
Question generation	Davey & McBride (1986)
Rereading	Dunlosky & Rawson (2005), Rawson, Dunlosky, & Thiede (2000)
Letter reinsertion	Maki, Foley, Kajer, Thompson, & Willert (1990)
Inserted questions	Mitsuda (1988, Experiment 1)
Provided perspective	Mitsuda (1988, Experiment 2)
Elaborated guided reading	Schmitt (1988)
Summarization	Thiede & Anderson (2003)
Delayed key word generation	Thiede, Anderson, & Therriault (2003), Thiede, Dunlosky, Griffin, & Wiley (2005)
Embedded questions	Walczyk & Hall (1989)

nitive activities during encoding that will be required at retrieval (cf. McDaniel, Friedman, & Bourne, 1978; Morris, Bransford, & Franks, 1978). For present purposes, the key implication is that the application of desirable difficulties or enriched-encoding manipulations in the classroom may demonstrate benefits in memory and metacomprehension performance only in situations in which the cognitive activities utilized for encoding are congruent with those required by the test. We will describe new work supporting this implication but first will provide a real-world example that demonstrates that issues of transfer-appropriate processing can be operative in education.

Consider the example of the high school history class instructor related to us by a colleague (McDaniel, 2007). This instructor required his students to engage in an integrative study activity that fostered the analysis of political, economic, religious, and scientific dimensions across cultures. To assess learning, the instructor typically administered a multiple-choice test on the details of the cultures. According to the transfer-appropriate processing framework, students in this class might not have benefited, in terms of test performance, from this presumably desirable study procedure, because the processes instantiated during acquisition (conceptual) might not have corresponded to those that were instantiated by the multiple-choice test of historical details. The more uncertain outcome that, up to now, has not been addressed is whether students' metacomprehension may also suffer, possibly leading to ineffective self-study and exacerbating the negative consequences when study activities and test requirements are mismatched.

We (Thomas & McDaniel, in press) directly examined whether transfer-appropriate considerations affect memory performance and metacomprehension accuracy when people engage in activities that promote desirable difficulties at encoding, by manipulating the processes required by the encoding tasks and the processes required by the retrieval tasks. Participants read educational texts under normal presentation conditions or under conditions that promoted desirable difficulties. The expository texts, taken from Levy (1981), had an average word count of 326 words. These passages were chosen so that the participants would be able to make six predictions per passage and so that six cued recall questions could be answered at the end of reading the passage. For each passage, six detail-oriented questions and six conceptual questions were generated. The detailed questions assessed individual item information, and the conceptual questions assessed global thematic information. The thematic questions related information that was presented across multiple sentences. Depending on condition, the participants received either detail-oriented question or conceptual questions for a given passage.

To promote desirable difficulties at encoding, two study activities previously shown to facilitate specific processes were implemented. Specifically, we used a sentence-sorting task, which has been shown to encourage relational processing of material, and a letter reinsertion task, which has been shown to encourage specific processing of details (McDaniel, Hines, Waddill, & Einstein, 1994,

Experiment 1). One of two different kinds of cued recall tests were used to assess learning after study: a conceptual or a detail-oriented test. Processing between encoding and retrieval was congruent when the participants engaged in letter reinsertion and were tested with the detail cued recall test or when sentence sorting was paired with the conceptual cued recall test. An example of a detailed question is as follows: *The walls of ice in Kanchenjunga range from _____ feet high. (The answer is 600–1,000.)* An example of a thematic question is as follows: *The downward speed of snow is much faster in Kanchenjunga than the Alps because _____.* (The answer is *there is more snowfall.*)

One important finding was that cued recall of text information was enhanced when the type of generative study activity focused on information congruent with that required in the cued recall test. The participants who engaged in letter reinsertion and then were given a detailed test performed better ($M = .60$) than did the participants who engaged in the same difficulty-encoding manipulation but were given a conceptual test ($M = .30$). On the other hand, the participants who engaged in sentence sorting and then were given a detailed test performed significantly worse ($M = .43$) than did the participants who engaged in that same encoding task but were given a conceptual test ($M = .63$). Interestingly, task–test congruency yielded test performance benefits even beyond those derived from simply reading passages ($M = .57$), and reading yielding statistically identical test performance for detailed and conceptual tests. These results are in line with the basic transfer-appropriate processing literature, in which typical laboratory stimuli, such as word lists, have been used. These studies, like Thomas and McDaniel (in press), have shown that memory performance is enhanced when the orienting activities stimulate the processing of information that is required in the test task (e.g., Blaxton, 1989; McDaniel et al., 1978; Morris et al., 1977; Nairne & Widner, 1987). These results are important to education in that they suggest that there are limitations to the use of enriched-encoding tasks at encoding (i.e., the desirability of enriched-encoding activities). Enriched-encoding activity yielded improvement in performance beyond that derived from the simple reading of texts only when the testing situation required processing congruent with that instantiated by the enriched-encoding study task. These results underscore the importance, in educational settings, of assigning study activities that produce processing that is congruent with the information being emphasized during testing.

More novel are the metacomprehension findings. Metacomprehension monitoring was measured by having participants make section-by-section predictions about future test performance. The relationship between metacomprehension predictions and retention is used to gauge the accuracy of those predictions and can be measured in two ways: (1) resolution, which is based on the relationship between individual item predictions and individual item test performance and is computed using Goodman–Kruskal gamma correlation, and (2) calibration, which is based on how well average predictions across items corre-

spond to average performance across items. For example, for a given passage presented by Thomas and McDaniel (in press), participants made a total of six predictions. The predictions were made after specific sections in the passage. Performance on test questions associated with those specific sections in the passage was then correlated with those predictions in order to assess resolution. Calibration was examined by averaging predictions for a given condition and comparing those averages with average test performance for that same condition. As in previous studies (e.g., Glenberg et al., 1987; Maki & Serra, 1992; Thiede et al., 2003; Weaver, 1990), resolution was assessed by computing a gamma correlation between predictions and actual performance for each participant. This method of analysis has been shown to not be affected by an individual's level of test performance or absolute threshold of prediction (for further discussion, see Nelson, 1984).

The metacomprehension findings were telling. Metacomprehension resolution was improved when the type of enriched-encoding study task and the type of test were congruent (gamma correlation $M = .50$), as compared with when the participants were not given an explicit enriched-encoding study task but merely were instructed to read ($M = .19$). More startling was the finding that metacomprehension was completely disrupted when the type of enriched-encoding study task and the type of test were incongruent (i.e., letter-reinsertion-conceptual test; sentence-sorting-detailed test) ($M = -.20$).

Three aspects of the profound metacomprehension impairment observed in the incongruent study-test conditions are notable. First, it was not tied specifically to reductions in memory. The sentence-sorting condition ($M = .43$) produced cued recall levels identical to those in the read condition ($M = .43$) when detailed questions were given at test. Second, it was not a consequence of inappropriate modulation of absolute prediction levels. That is, the mean prediction level across items corresponded to mean cued recall performance. Thus, although resolution, as measured by gamma correlations, was poor, calibration was good. On the basis of the observations above, a plausible interpretation is that on a more global level, participants can modulate predictions on the basis of both congruent and incongruent study-test processing. That is, on average, participants may be aware that they will be able to recall less information in incongruent than in congruent conditions; however, when asked to make more fine-grained section-by-section judgments, participants are unable to predict performance in incongruent conditions. It is important to note that the participants were aware of the type of test questions to expect before making predictions of future test performance. In addition, the participants took part in a practice phase in which they practiced the encoding tasks and were exposed to both conceptual and detail-oriented test questions. These procedures were undertaken so that the participants would expect the appropriate type of test questions when making predictions. With this type of practice, it is unlikely that the impairment found in the incongruent task-test conditions was due to inexperience with the type of test questions.

In addition to the two demonstrated benefits of congruency (improved retention and metacomprehension), these findings also lead to the provocative possibility that at least a third positive benefit might accrue to test-appropriate study activities, a benefit that has significant educational implications. Specifically, given the high level of metacomprehension accuracy for test-appropriate study tasks, these tasks may promote efficacious control of subsequent learner-initiated study activities. In contrast, the decrement in metacomprehension accuracy when test-inappropriate study activities were used may extend to prompt ineffectual and misguided control of subsequent self-study.

We evaluated this possibility in a second experiment (Thomas & McDaniel, in press, Experiment 2). The experimental procedure was similar to that in Experiment 1; however, the participants were re-presented with each passage in sections for restudy after metacomprehension predictions had been made. The sections were the same as those used when metacomprehension predictions were collected. Before the presentation of the first section of a given passage, the participants were told that they could spend as much or as little time restudying the section as they deemed necessary. Because a given passage had already been processed two times (once during the initial encoding phase and a second time during the metacomprehension prediction phase), if the participants chose, they could skip rereading a given section of a passage by simply pressing the space bar. On the other hand, the participants could spend as much time as they chose trying to master the material in a given section of the passage. Control was measured by restudy time.

When restudy time was correlated with metacomprehension predictions, we found that the participants selected items to restudy that were judged to be less likely to be recalled. These findings are consistent with those in much of the study time allocation literature, which has shown that learners allocate more study time to items judged difficult (for a review, see Son & Metcalfe, 2000). Although our findings are consistent with those in the literature, their implications are somewhat troubling. Specifically, in the incongruent condition, the participants allocated time to items judged harder; however, the gamma analyses between predictions and performance showed that the items judged to be more difficult were not necessarily more difficult (i.e., resolution at chance). Thus, in the incongruent conditions, because monitoring was so inaccurate, the subsequent controlled study process may have been ineffective. In fact, when cued recall results across Experiments 1 and 2 were compared, improvements in performance after restudy were evidenced in congruent conditions, whereas decrements in performance were found in incongruent conditions. Table 2 shows these cued recall results.

In sum, these results highlight the importance of congruent processing at encoding and testing for retention, monitoring, and subsequent control of study, yet incongruency between encoding and retrieval may often occur in educational settings. Students may not adopt study strategies that are congruent with subsequent testing. Instructors may suggest or require enriched-encoding ac-

Table 2
Mean Cued Recall Performance as a Function of
Orienting Task and Type of Test Before (Experiment 1)
and After (Experiment 2) Restudy

Condition	Experiment 1		Experiment 2	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Detailed Test				
Congruent (letter reinsertion)	.60	.04	.84	.03
Incongruent (sentence sorting)	.43	.04	.37	.07
Conceptual Test				
Congruent (sentence sorting)	.63	.04	.85	.04
Incongruent (letter reinsertion)	.30	.03	.25	.04

Note—Adapted from Thomas and McDaniel, in press.

tivities without consideration for what information will be tested. The clear implication for education is that for efficacious metacognitive processing, students need to be guided toward study activities that are congruent with test objectives; simply implementing activities that promote desirable difficulties or enriched study activities will not necessarily improve memory or metacomprehension. More generally, research directed at the metacognitive consequences of the congruency between enriched (active, difficult, etc.) encoding activities and the criterial test appears to be a fruitful direction for applying cognition to education.

Testing to Promote Accurate Metacomprehension

In situations in which the students’ encoding activities are incongruent with the processing required by the testing situation, what can be done to overcome the negative cascade of impaired metacomprehension and ineffective subsequent control of study activities? We suggest that testing can be used as a monitoring mechanism if students are able to recognize failures to retrieve. That is, testing may be able to improve the efficiency of learning by helping students isolate poorly learned material for restudy.

In much of the metacognitive and metacomprehension research, practice tests have been investigated to the extent that they may be able to improve predictive accuracy. For example, when students learn translation equivalents (e.g., *pombe-beer*), practice tests that are delayed several minutes after study and are prompted by the stimulus alone have almost universally supported high levels of predictive accuracy (King, Zechmeister, & Shaughnessy, 1980; Lovelace, 1984). Practice tests have also yielded improvements in predictive accuracy for students learning lists of related terms (Kelemen, 2000). (As has been detailed in McDaniel, Roediger, & McDermott [2007], practice testing has benefits to memory retention as well.)

On the basis of the results above, we reasoned that testing might be used to reduce and/or ameliorate the negative cascade of incongruent task–test processing. To address this issue, we conducted an experiment using the methodology in Thomas and McDaniel (in press). Participants engaged in two enriched-encoding manipulations (letter reinsertion or sentence sorting). Immediately following the encoding manipulation, the participants were given

four cued recall test questions. After answering these questions, the participants were given the opportunity to restudy the previously encoded passage. During restudy, the participants were presented with intact passages in the correct sentence order. The passages were presented in sections, and the participants could spend as much or as little time restudying each section of the passage as they wished. Following restudy, the participants were given a second test associated with the previously studied material. It is important to note that the tests were either detail-oriented or conceptually oriented in order to reflect both congruent and incongruent conditions. In addition, the first and second tests consisted of different questions but retained the same type of question. The participants were told that the type of questions in Test 1 and Test 2 would be the same for a given passage. Thus, we assumed that the participants would be able to anticipate the type of test questions in Test 2 on the basis of what they had been presented with in Test 1.

Replicating Thomas and McDaniel (in press), the participants who engaged in incongruent task–test processing performed less well on Test 1 than did the participants who engaged in congruent task–test processing. More interesting, all the groups showed some improvement on Test 2. Most important, those improvements were greatest in conditions in which the participants had engaged in incongruent task–test processing (see Table 3).² When average restudy times were examined, we found that the participants who had engaged in incongruent task–test processing spent more time restudying passages than did those who had engaged in congruent task–test processing. These findings suggest that students may be able to use performance on tests as a guide to further study (as long as the questions on Test 2 are similar to those given on Test 1), which, of course, is the rationale for the customary practice of providing practice questions and tests. The important result in Thomas and McDaniel (2007) is the demonstration that taking (practice) tests helps guide further study that compensates for incongruent task–test processing, although this did not, in this experiment, completely override the negative effects of an initial incongruent study activity.

Summary

As the literature on metacomprehension continues to grow, ways to improve metacomprehension accuracy continue to unveil themselves. The importance of improving metacomprehension has direct relevance to education. The better students are at assessing what they have learned and what they have not, the more likely it is that

Table 3
Mean Cued Recall Performance as a Function of
Type of Encoding, Type of Test, and Test Order

Encoding	Detailed		Conceptual	
	T1	T2	T1	T2
Letter reinsertion	.80	.84	.41	.62
Sentence sorting	.58	.79	.81	.82

Note—T1, Test 1; T2, Test 2.

is that the students will be able to implement effective controlled processing to acquire more difficult material. This article has demonstrated several important factors that contribute to metacomprehension accuracy. In addition to enriched and elaborated encoding manipulations, highlighted by Thiede and colleagues, recent research by Thomas and McDaniel (2007, in press) suggests the need to account for the interaction between processes instantiated by the encoding manipulations and those required by the type of test. Study–test congruency improves metacomprehension accuracy, presumably because congruency yields a rich relevant knowledge base on which to base metacomprehension predictions. Study–test incongruency disables metacomprehension accuracy, thereby rendering subsequent control processes ineffective. Preliminary results in Thomas and McDaniel (2007) suggest that the negative cascade of incongruent task–test processing can be ameliorated if students are able to gain experience pertaining to their performances on representative test questions (e.g., practice tests). This research has direct implications for the way instructors teach classes, on the way text books are constructed, and on the way students are trained to study for tests. Instructors and students need to engineer class exercises and study activities to establish congruent relationships between encoding activities and retrieval tasks. In the event that congruent relationships cannot be established, practice testing may serve to ameliorate the negative cascade of incongruent task–test processing.

AUTHOR NOTE

This research was supported by a James S. McDonnell Foundation Collaborative Activity Award (Grant 220020041). We thank Adrian Gilmore and Katherine Price for their help on various aspects of this project, including subject testing and data scoring. Correspondence concerning this article should be addressed to A. K. Thomas, Tufts University, 490 Boston Ave., Medford, MA 02155 (e-mail: ayanna.thomas@tufts.edu).

REFERENCES

- ANDERSON, M. C. M., & THIEDE, K. W. (2007). *Why do delayed summaries improve metacomprehension accuracy?* Manuscript submitted for publication.
- BJORK, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185-205). Cambridge, MA: MIT Press.
- BLAXTON, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 657-668.
- CHI, M. T. H. (2000). Cognitive understanding levels. In A. E. Kazdin (Ed.), *Encyclopedia of psychology* (Vol. 2, pp. 146-151). Washington, DC: American Psychological Association.
- CUEVAS, H. M., FIORE, S. M., & OSER, R. L. (2002). Scaffolding cognitive and metacognitive processes in low verbal ability learners: Use of diagrams in computer-based training environments. *Instructional Science*, **30**, 433-464.
- DAVEY, B., & MCBRIDE, S. (1986). Effects of question-generation training on reading comprehension. *Journal of Educational Psychology*, **78**, 256-262.
- DAVIS, E. A., & LINN, M. C. (2000). Scaffolding students' knowledge integration: Prompts for reflection in KIE. *International Journal of Science Education*, **22**, 819-837.
- DEWINSTANLEY, P. A., & BJORK, E. L. (2004). Processing strategies and the generation effect: Implications for making a better reader. *Memory & Cognition*, **32**, 945-955.
- 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.
- EINSTEIN, G. O., MCDANIEL, M. A., OWEN, P. D., & COTÉ, N. C. (1990). Encoding and recall of texts: The importance of material appropriate processing. *Journal of Memory & Language*, **29**, 566-581.
- GLENBERG, A. M., & EPSTEIN, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, & 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.
- KELEMEN, W. L. (2000). Metamemory cues and monitoring accuracy: Judging what you know and what you will know. *Journal of Educational Psychology*, **92**, 800-810.
- 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.
- LEVY, B. A. (1981). Interactive processes during reading. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* (pp. 1-35). Hillsdale, NJ: Erlbaum.
- LOVELACE, E. A. (1984). Metamemory: Monitoring future recallability in free and cued recall. *Bulletin of the Psychonomic Society*, **22**, 497-500.
- MAKI, R. H., & BERRY, S. L. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 663-679.
- MAKI, R. H., FOLEY, J. M., KAJER, W. K., THOMPSON, R. C., & WILLERT, M. G. (1990). Increased processing enhances calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 609-616.
- MAKI, R. H., & SERRA, M. (1992). The basis of test predictions for text material. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 116-126.
- MCDANIEL, M. A. (2007). Rediscovering transfer as a central concept. In H. L. Roediger III, Y. Dudai, & S. Fitzpatrick (Eds.), *Science of memory: Concepts* (pp. 355-360). New York: Oxford University Press.
- MCDANIEL, M. A., & EINSTEIN, G. O. (2005). Material appropriate difficulty: A framework for determining when difficulty is desirable for improving learning. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 73-85). Washington, DC: American Psychological Association.
- MCDANIEL, M. A., FRIEDMAN, A., & BOURNE, L. E., JR. (1978). Remembering the levels of information in words. *Memory & Cognition*, **6**, 156-164.
- MCDANIEL, M. A., HINES, R. J., WADDILL, P. J., & EINSTEIN, G. O. (1994). What makes folk tales unique: Content familiarity, causal structure, scripts, or superstructures? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 169-184.
- MCDANIEL, M. A., ROEDIGER, H. L., III, & MCDERMOTT, K. B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review*, **14**, 200-206.
- MCDANIEL, M. A., & HEALY, A. F. (1995). A procedural explanation of the generation effect: The use of an operand retrieval strategy for multiplication and addition problems. *Journal of Memory & Language*, **34**, 399-416.
- METCALFE, J., & KORNELL, N. (2003). The dynamics of learning and allocation of study time to a region of proximal learning. *Journal of Experimental Psychology: General*, **132**, 530-542.
- METCALFE, J., & KORNELL, N. (2007). Principles of cognitive science in education: The effects of generation, errors, and feedback. *Psychonomic Bulletin & Review*, **14**, 225-229.
- MITSUDA, M. (1988). Effects of processing variables on comprehension-monitoring activities. *Psychologia*, **31**, 98-105.
- MORRIS, C. D., BRANSFORD, J. D., & FRANKS, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior*, **16**, 519-533.
- NAIRNE, J. S., & WIDNER, R. L. (1987). Generation effects with nonwords: The role of test appropriateness. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 164-171.

- NELSON, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, **95**, 109-133.
- NELSON, T. O., & DUNLOSKY, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect." *Psychological Science*, **2**, 267-270.
- NELSON, T. O., & NARENS, L. (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.
- PALINCSAR, A. S., & BROWN, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition & Instruction*, **1**, 117-175.
- RAWSON, K. A., DUNLOSKY, J., & THIEDE, K. W. (2000). The rereading effect: Metacomprehension accuracy improves across reading trials. *Memory & Cognition*, **28**, 1004-1010.
- SCHMITT, M. C. (1988). The effects of an elaborated directed reading activity on the metacomprehension skills of third graders. In J. E. Readence & R. S. Baldwin (Eds.), *Dialogues in literacy research: Thirty-seventh yearbook of the National Reading Conference* (pp. 167-181). Chicago: National Reading Conference.
- SON, L. K., & METCALFE, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 204-221.
- 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., & 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.
- THOMAS, A. K., & MCDANIEL, M. A. (2007). *Testing to improve metacomprehension*. Manuscript in preparation.
- THOMAS, A. K., & MCDANIEL, M. A. (in press). The negative cascade of incongruent generative study-test processing in memory and metacomprehension. *Memory & Cognition*.
- WALCZYK, J. J., & HALL, V. C. (1989). Effects of examples and embedded questions on the accuracy of comprehension self-assessments. *Journal of Educational Psychology*, **81**, 435-437.
- WEAVER, C. A., III (1990). Constraining factors in calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 214-222.
- WEAVER, C. A., III, & BRYANT, D. S. (1995). Monitoring of comprehension: The role of text difficulty in metamemory for narrative and expository text. *Memory & Cognition*, **23**, 12-22.
- WINNE, P. H., & HADWIN, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277-304). Mahwah, NJ: Erlbaum.

NOTES

1. The theoretical construct of metacomprehension is closely related to the concept of *comprehension monitoring*, or people's awareness of whether they understand what they are reading (Mayer, 1998). In the present article, we make a distinction between comprehension monitoring and metacomprehension, with a focus on metacomprehension. For the present purposes, metacomprehension is analogous to metacognitive processes, the only differences being the type of material the learner is monitoring.
2. We acknowledge the possibility that ceiling effects may have masked the improvement found on Test 2 for congruent conditions. Further experiments are presently underway to remove this factor.