

Concept mapping improves metacomprehension accuracy among 7th graders

Joshua S. Redford^a, Keith W. Thiede^{a,*}, Jennifer Wiley^b, Thomas D. Griffin^b

^aBoise State University, United States

^bUniversity of Illinois at Chicago, United States

ARTICLE INFO

Article history:

Received 31 January 2011

Received in revised form

24 September 2011

Accepted 10 October 2011

Keywords:

Metacomprehension

Concept map

ABSTRACT

Two experiments explored concept map construction as a useful intervention to improve metacomprehension accuracy among 7th grade students. In the first experiment, metacomprehension was marginally better for a concept mapping group than for a rereading group. In the second experiment, metacomprehension accuracy was significantly greater for a concept mapping group than for a control group, while a group of students who were given already constructed concept maps had accuracy between these two groups. In both experiments, control groups had poor metacomprehension accuracy. That is, they performed worse on tests they predicted better performance and performed better on tests they predicted worse performance. Although constructing concept maps did not produce the same high level of accurate monitoring previously reported in the literature, it still reduced the illusion of knowing.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Models of self-regulated learning suggest that learners monitor their progress toward a goal and use this information to regulate their study (e.g., Ariel, Dunlosky, & Bailey, 2009; Metcalfe & Kornell, 2003; Nelson & Narens, 1990; Thiede & Dunlosky, 1999; Winne & Hadwin, 1998). Accurate monitoring identifies which information is well-learned and which information requires additional study. The influence of accurate monitoring on learning has been empirically supported across a variety of domains (for a review, see Dunlosky, Hertzog, Kennedy, & Thiede, 2005). More relevant to the present research, accurate monitoring and effective regulation of study has led to better comprehension of texts (Thiede, Anderson, & Theriault, 2003). Therefore, it is important to find ways to improve metacomprehension accuracy (the accuracy with which one monitors his or her own comprehension of text). To understand how one might improve metacomprehension accuracy, it is important to combine theories of metacognitive monitoring and comprehension (Rawson, Dunlosky, & Thiede, 2000; Weaver, 1990; Wiley, Griffin, & Thiede, 2005).

The cue-utilization model of metacognitive monitoring (Koriat, 1997) and the construction-integration model of comprehension (Kintsch, 1994, 1998) provide a framework for understanding techniques for improving metacomprehension accuracy (Rawson et al., 2000). Consider the processes involved in judging one's

comprehension of texts. After reading, a person is asked to judge his or her comprehension of a text. According to the cue-utilization framework of metacognitive monitoring (Koriat, 1997), the metacomprehension judgment may be based on a number of cues, such as how easily the text was processed during reading (Dunlosky & Rawson, 2005; Rawson & Dunlosky, 2002), how successfully the material was retrieved at the time of the judgment (Baker & Dunlosky, 2006; Benjamin, Bjork, & Schwartz, 1998; Morris, 1990), the familiarity with the domain of the text (Glenberg & Epstein, 1985; Glenberg, Sanocki, Epstein, & Morris, 1987; Maki & Serra, 1992), or global characteristics of texts such as length or difficulty (Weaver & Bryant, 1995). Metacomprehension accuracy will tend to increase as the cues that are used to make a judgment more highly correlate with performance on a test of comprehension (for empirical evidence linking metacomprehension accuracy and cue use see Thiede, Griffin, Wiley, & Anderson, 2010).

The construction-integration model (Kintsch, 1994, 1998) suggests different cues that may be available for metacomprehension judgments. According to this model, a reader creates multiple representations of a text as he or she reads. For instance, the reader constructs a representation of the surface level (i.e., the exact words), a textbase level (i.e., the meaning of sentences), and the situation-model level (i.e., connections between ideas contained in the text, and the connection between these ideas and prior knowledge). A well-constructed situation model integrates the ideas contained in a text and allows the reader to form a causal model and inferences implied by the text. When tests of comprehension assess the quality of the situation model of a text (Kintsch, 1994; McNamara, Kintsch,

* Corresponding author. Boise State University, Education (MS 1745), 1910 University Drive, Boise, ID 83725-1745, United States.

E-mail address: KeithThiede@boisestate.edu (K.W. Thiede).

Songer, & Kintsch, 1996), metacomprehension accuracy should increase if readers use cues based on their situation model to judge their comprehension (for a detailed discussion of designing texts and tests for assessing the situation model see Wiley et al., 2005). In fact, many of the techniques shown to improve metacomprehension accuracy arguably focus readers on their situation model while judging comprehension. We next provide a brief review of the metacomprehension literature.

Initial metacomprehension work began in the 1980s (e.g., Glenberg & Epstein, 1985; Maki & Berry, 1984). Glenberg and Epstein (1985) developed a paradigm in which participants read 16 texts and rated their level of comprehension for each. Participants then completed a comprehension test for each text. Metacomprehension accuracy was operationalized as the intra-individual correlation between a participant's metacomprehension rating and his or her test performance computed across the texts. Although poor metacomprehension accuracy reported in early studies was in part due to poor measures of comprehension (Weaver, 1990), much of the work in this area suggested college students are not adept at monitoring their own comprehension. Maki (1998) found that the mean correlation between predicted and actual test performance (metacomprehension accuracy) across the studies conducted in her laboratory was just .27. Dunlosky and Lipko (2007) reviewed the metacomprehension literature and found the mean accuracy across all the reviewed studies was no better.

Recently, interventions have been developed to improve metacomprehension accuracy. For instance, Thiede et al., (2003) had college students write summaries of texts after reading but prior to judging comprehension. When summaries were written immediately after reading, the summaries contained more details from the text, which may have focused participants on the surface features of a text when judging comprehension, and metacomprehension accuracy did not improve. By contrast, when summaries were written after a delay encouraged readers to focus on the situation model because the surface features decays over time, and metacomprehension accuracy improved (see also Anderson & Thiede, 2008). Generating a list of five keywords that captured the essence of a text, instead of writing a summary, produced a similar boost in metacomprehension accuracy (Thiede et al., 2003). Again the timing of keyword generation was critical to improving metacomprehension accuracy. Thiede, Dunlosky, Griffin, and Wiley (2005) showed that the delay between reading and generating keywords was critical to improving metacomprehension accuracy.

Again with an eye toward focusing participants on the situation model of a text when judging comprehension, Griffin, Wiley, and Thiede (2008) had participants self-explain the connections between the ideas in a text as they read. Self-explaining, which facilitates construction of a situation model (Chi, DeLeeuw, Chiu, & Lavancher, 1994), increased the salience of cues related to the situation model at the time of judging comprehension, and improved metacomprehension accuracy. In addition, Thiede, Wiley and Griffin (2011) showed that providing instructions about what kind of test to expect and giving practice tests focused college students on appropriate cues and produced more accurate monitoring. For a review of the metacomprehension literature and theory related to improving metacomprehension accuracy by focusing readers on their situation model see Thiede, Griffin, Wiley, and Redford (2009).

As described above, many of the techniques that have been shown to improve metacomprehension accuracy arguably focus readers on their situation model while judging comprehension. One additional technique, which was used in the present investigations, is instructing participants to construct concept maps of texts prior to judging comprehension. As a concept map is a graphic representation of the underlying structure of the text, the act of

constructing a concept map helps readers form connections among concepts in a text (Weinstein & Mayer, 1986). Put differently, concept mapping can help readers form an external representation of a situation model for a text, which can be particularly helpful for less-able readers (Stensvold & Wilson, 1990; for a review and meta-analysis on effectiveness of concept maps with low-ability learners, see Nesbit & Adesope, 2006). Moreover, constructing concept maps should increase the salience of cues related to the situation model when judging comprehension.

Two advantages may arise from testing concept map construction as an additional intervention to improve metacomprehension accuracy. First, as keyword generation requires a delay to be effective, improved metacomprehension accuracy could arguably be due to transfer-appropriate-monitoring. Specifically, delayed keyword generation may have improved metacomprehension accuracy because the participants employed the same long-term memory retrieval process during the rating and test phase. Finding an improvement in metacomprehension accuracy with concept map construction provides additional support that the critical contributing factor is whether the situation model of the text is being accessed for judgements. A second benefit is that concept map construction is more practical than delay-based interventions. Teachers would have an easier time integrating concept map construction into their standard pedagogical practice than some practice that requires students to pause for several minutes prior to engaging in the next phase of text processing.

In an initial study, Thiede et al. (2010) showed that constructing concept maps improved metacomprehension accuracy with college students enrolled in remedial reading courses, but this technique has not been evaluated with younger readers. We know very little about younger readers' ability to monitor comprehension.

To date, we know of only one other study examining metacomprehension with younger readers. de Bruin, Thiede, Camp, and Redford (2011) showed that generating keywords improved metacomprehension accuracy for 6th, and 7th graders, but had no effect on accuracy for 4th graders. Given these results, one might expect concept mapping to improve metacomprehension accuracy in early adolescents. Nonetheless, it is important to evaluate whether concept mapping improves metacomprehension accuracy for this age group.

Although concept mapping has been shown to improve metacomprehension accuracy for some college students, there are some critical differences between college students and 7th graders. First, college students have relatively more experience reading expository texts; whereas, younger students are only just beginning to learn how to learn from expository texts. Reading instruction in early grades is generally confined to the understanding of narrative texts, even though students are expected to eventually read for understanding from informational expository texts within particular subject-matter areas. Students have much less familiarity with the structures and types of non-narrative texts (Duke, 2000; Pearson & Duke, 2002; Venezky, 2000). Thus, younger students may not be sensitive to the unique demands of constructing a causal situation model from informational expository texts, which in turn may affect their access to and selection of cues that they use to judge their comprehension.

Second, even if lack of experience with expository texts does not affect metacomprehension accuracy, younger students may lack the cognitive skills or capacity required to accurately monitor learning while reading complex expository texts. Studies have shown that children are less sophisticated in related cognitive domains, such as memory regulation (e.g., Roderer & Roebbers, 2009), uncertainty confidence monitoring (e.g., Roebbers, von der Linden, Schneider, & Howie, 2007), judgments of learning (e.g., Roebbers et al., 2007), and inconsistency detection (Markman, 1977,

1979). Therefore, younger students may not possess the capacities required to monitor comprehension during reading.

If less experience with expository texts inhibits 7th graders' ability to attend to appropriate cues for judging comprehension or if 7th graders lack the metacognitive sophistication to use cues to accurately assess their own comprehension, concept mapping will not affect metacomprehension accuracy. By contrast, if the improved metacomprehension accuracy for 7th graders who generated keywords (de Bruin et al., 2011) is evidence that younger readers can monitor their comprehension of expository texts, then concept mapping should improve metacomprehension accuracy. We expect concept mapping to improve metacomprehension accuracy for two reasons (Hypothesis 1). First, concept mapping eases the working memory requirements of text comprehension. Instead of needing to maintain earlier content as they try to understand later portions of a text, they are able to review the previous content visually as they construct their maps. This enables more resources to be devoted to gauging their level of comprehension when they revisit their maps rather than having resources divided between metacognitive monitoring and comprehending the text (Griffin et al., 2008; Rawson et al., 2000). Second, we expect concept mapping to improve metacomprehension accuracy as it increases the salience of cues related to the situation model, which should improve metacomprehension accuracy.

2. Experiment 1

In the first experiment, we asked students to construct concept maps as they read a set of three texts. We expected that concept map construction would encourage situation model development as well as ease the cognitive demands required to monitor comprehension. If concept maps allow participants to access and select cues based in their situation models for judging comprehension, then constructing concept maps should improve metacomprehension accuracy.

3. Method

3.1. Participants and design

Fifty-nine 7th grade students (ages 12–13) from a local junior high school participated in this experiment. Of the 59, 33 were female and 26 were male; 45 were Caucasian, 14 were Hispanic or African–American. The school serves a low SES population, with 70% of students eligible for free or reduced lunch. All participants were treated in accord with APA ethical standards. Group assignment (concept mapping versus rereading/control) was done randomly between participants. We used a reread condition in order to equate the amount of time participants in the two groups worked with the texts. Of the 59 participants, thirty-eight participants were in the concept map construction condition and 21 participants were in the reread condition. We over assigned participants to the concept map construction group to increase the statistical power in examining the relation between the characteristics (content) of concept maps and metacomprehension judgments, test performance, and metacomprehension accuracy.

3.2. Materials

Junior high school science textbooks were used to create three science-based expository texts. Each text was approximately 430 words long. We chose topics that provided a framework for a causal model of a scientific phenomenon (water cycle, path of air in the circulatory system, and the visual system). The readability of the texts was similar from one text to another, with Flesch–Kincaid

grade levels ranging from 6.1 to 6.5. We also constructed a 5-question inference test for each text based in this causal model. As outlined in Wiley et al. (2005), questions were designed to assess the generation of inferences or connections implied by the text, and not simple memory for the facts contained in the text. A sample text and test can be found in Appendix 1.

3.3. Procedure

Prior to the experiment, all the participants received three 15–20 min lessons on concept mapping as part of their typical language arts class. All the instruction was done by the same teacher, using three classes. The training was done as a whole class activity. The first lesson described the general idea of concept mapping, and included discussion of how this might improve reading comprehension. The class then worked through one expository science text as a class. The subsequent two lessons involved working on practice texts. Students read and constructed concept maps. They then critiqued each other's maps and the whole class then walked through the mapping process with the teacher. In demonstrating mapping, the teacher put the main ideas of a text in nodes of the concept map and then used arrows to represent the relationship between concepts. The teacher also emphasized that constructing concept maps helped students understand texts by illustrating the relation among ideas in a text. Individual student's concept maps were not collected and graded, so students did not get individualized feedback on mapping.

On the day of the experiment, participants were instructed that they would be reading three texts, judging their comprehension of each test (i.e., predicting how they would do on a 5-item multiple choice test), and then completing a test for each text. After having an opportunity to ask questions about the experiment, participants read and constructed concept maps or reread each text (texts were present for reading and then immediately for rereading, that is, reading and rereading was *not* blocked for the three texts). They then judged their comprehension. The prompt for the metacomprehension judgment was, "Please indicate how many of the five test questions you think you will answer correctly on the text—TITLE OF TEXT." After making judgments, participants completed tests for each text. Students did not have access to the texts while judging or completing tests, but students in the concept map condition did have access to their concept maps during these activities.

4. Results

4.1. Metacomprehension judgments and test performance

As metacomprehension accuracy is the relationship between metacomprehension judgments and test performance, we first report data on these variables. The median of both metacomprehension judgments and test performances across the three texts was computed for each participant. We used the median because it is the recommended measure of central tendency for small sets of scores where extreme scores may have an undue influence on the mean (Gravetter & Wallnau, 1999). The mean of the medians was then computed across participants in each group, see Table 1. The mean magnitude of metacomprehension judgments did not differ across groups, $t(57) < 1.00$; nor did mean test performance, $t(57) = 1.16, p > .05$.

4.2. Metacomprehension accuracy

As suggested by Nelson (1984) metacomprehension accuracy was operationalized as a Goodman-Kruskal gamma correlation

Table 1
Test performance and magnitude of metacomprehension ratings.

Group	Test performance	Metacomprehension rating
Experiment 1		
Concept map (constructed)	2.45 (.12)	3.32 (.16)
Rereading	2.24 (.17)	3.10 (.15)
Experiment 2		
Concept map constructed	2.29 (.11)	3.16 (.18)
Concept map provided	2.32 (.10)	3.00 (.16)
Control	1.86 (.12)	2.97 (.18)

Note. Entries are mean across individual's median test performance and median comprehension ratings. Values in parentheses are standard errors of the means.

between judgments and test performance. Therefore, for each participant, we computed a gamma correlation between metacomprehension judgments and test performance across the three texts.¹ The mean gamma was then computed across participants in the respective groups. As seen in Fig. 1, the concept mapping group had a positive gamma, but this was not significantly different from zero, $t(27) < 1.00$. The control group had a negative gamma correlation, which was marginally different from zero, $t(16) = 1.95$, $p = .07$, which indicates that students in the control group gave higher metacomprehension judgment to texts they understood less. Perhaps more important, metacomprehension accuracy was marginally greater for the concept mapping group than for the rereading group, $t(43) = 1.70$, $p < .10$ (Hypothesis 1). Thus, constructing concept maps reduced the illusion of knowing.

Although the concept maps were fairly simplistic, we explored whether a relationship existed between map characteristics and metacomprehension accuracy or test performance. If particular map dimensions were related to either metacomprehension or test performance, future instruction could emphasize those factors in preparing students for reading in general and monitoring comprehension in particular. The properties that we examined are summarized in Table 2. We looked at the size of the concept maps (i.e., the number of nodes), how many nodes were repeated, how many reflect the paragraph structure of the text, and how many nodes had multiple links. A one-tailed Pearson correlation found only one significant relationship—an increase in redundant nodes was linked with a decrease in metacomprehension accuracy, $r = -.35$, $p < .05$. This suggests that participants were less equipped to estimate their level of comprehension when their maps were unnecessarily muddled with repeated nodes.

5. Discussion

Although these data suggest that concept map construction during reading may improve metacomprehension accuracy in young readers (Hypothesis 1), the absolute level of metacomprehension accuracy remained quite poor, and well below the levels of accuracy typically found with college students (Dunlosky & Lipko, 2007; Maki, 1998; Thiede et al., 2009) and below that found using delayed keyword generation with children (de Bruin et al., 2011). This poor level of accuracy was obtained even when

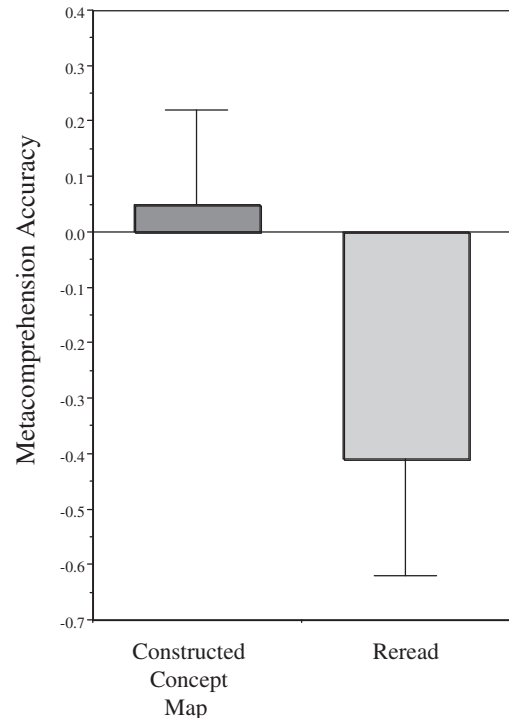


Fig. 1. Mean metacomprehension accuracy (gamma correlation) for Experiment 1. Error bars are the standard error of the mean.

students had the concept maps available to them during judging and testing. Perhaps more alarming is the poor accuracy for participants in the rereading group. This poor accuracy may be surprising because rereading has been shown to improve metacomprehension accuracy with older readers (Dunlosky & Rawson, 2005; Rawson et al., 2000). Thus, improving the metacomprehension accuracy among young readers may present a unique challenge. In our next experiment we attempt to pursue the potential benefits of concept map construction in two ways. First, we create two concept map groups where one group is provided concept maps, and the other group constructs them, to test for the effects of engaging in map construction, or whether access to a concept map during judgement/testing is sufficient to improve accuracy. Second, students may have been using the concept maps as a summarizing tool rather than as a tool for better understanding the text. Leopold and Leutner (2011) demonstrate the advantage of drawing (to understand) relative to summarization for text comprehension. Therefore, in the second experiment, we include a more elaborate set of concept mapping lessons for all students that includes both instruction in the mechanics of creating concept maps (as in Experiment 1), but also elaborates on the kinds of comprehension questions that students should expect, and how creating connections in the concept maps can help to answer those types of questions.

¹ Nelson (1984) recommended using a Goodman-Kruskal gamma correlation (Goodman & Kruskal, 1954) for these kinds of data. Gamma is computed by examining the direction of one variable relative to another. If one variable (e.g., metacomprehension judgment) is increasing from one text to another and the other variable (e.g., test performance) is also increasing across this same pair of texts, this is considered a concordance (C). By contrast, if one variable is increasing from one text to another and the other variable is decreasing across this same pair of texts, this is considered a discordance (D). Concordance and discordance is computed across all pairs of items. The total number of each is used to compute the correlation coefficient, $\text{Gamma} = (C - D)/(C + D)$.

Table 2
Concept map content for Experiment 1 and Experiment 2.

	Experiment 1	Experiment 2
Number of nodes	50.2 (5.6)	50.9 (5.4)
Redundant nodes	4.4 (1.4)	0.9 (.3)
Paragraph structure	0 (0)	1.5 (.3)
Multiple-link nodes	0.6 (.2)	1.9 (.5)

Note. Values in parentheses are standard errors of the means.

6. Experiment 2

In this experiment, we again examined the effect of concept mapping on metacomprehension accuracy for 7th graders. However, additional information about the benefits of concept mapping was provided when mapping was being taught, to provide metacognitive scaffolding for mapping (Pressley, Borkowski, & O'Sullivan, 1984). Moreover, students received explicit information on the kinds of questions they would be asked to answer (i.e., inference questions). This kind of information and experience with inference tests has been shown to improve metacomprehension accuracy of college students (Thiede et al., 2011). This additional instruction was hypothesized to further improve metacomprehension accuracy (Hypothesis 2).

As the quality of concept maps constructed by participants in Experiment 1 varied dramatically, with many maps being quite poorly designed, we added a new condition in this experiment. One group of participants was given a well developed concept map to refer to while reading. Although this should have provided students in this condition with a visual representation of the connections available in the texts, past research has shown that the act of generating situation model based cues is critical to improving metacomprehension accuracy (Thiede et al., 2005). However, if having access to a visual representation of the structure of a text is critical to metacognitive monitoring, then this condition should also improve accuracy.

7. Method

7.1. Participants and design

One-hundred-three 7th grade students (ages 12–13) from the same junior high school as in Experiment 1 participated in this experiment. Of the 103, 56 were female and 47 were male; 77 were Caucasian, 26 were Hispanic or African–American. All participants were treated in accord with APA ethical standards. Group assignment (concept map construction versus concept map provided versus control) was done randomly between participants. Thirty-two of these participants engaged in concept map construction and 39 participants were provided with a completed concept map for each text, and 32 were in the control group. Given that rereading texts produced poor metacomprehension accuracy in Experiment 1, the control group in this experiment simply read the text one time, which is the standard control condition in metacomprehension research.

7.2. Materials

We worked with the same teacher as in Experiment 1, but during a later academic semester. Although the teacher was expected to have completely different students, as a precaution to avoid possible exposure to the texts used in Experiment 1, we developed new texts. As in Experiment 1, junior high school science textbooks were used to create three science-based expository texts. Each text was approximately 400 words long. We chose texts that provided a causal model of a scientific phenomenon (functions of bacteria, breeding and cloning, and the causes of extinction). The readability of the texts was similar from one text to another, with Flesch-Kincaid grade levels ranging from 7.3 to 7.7. We constructed a 5-question test that asked about inferences designed to tap the situation model for each text.

7.3. Procedure

Prior to the experiment, all the participants received eight 15 min lessons on concept mapping as part of their typical

language arts class. All the instruction was done by the same teacher as in Experiment 1, using six classes. The training was done as a whole class activity. The first lesson described the general idea of concept mapping, and included discussion of how this might improve reading comprehension. The class then worked through one expository science text as a class. In the second lesson, the whole class worked through a sample text and then used the concept map to answer practice test questions. The nature of the inference questions were explained during this lesson. The subsequent six lessons involved working on practice texts, and taking practice inference tests. The teacher emphasized that constructing concept maps helped students understand texts by illustrating the relation among ideas in a text, and also that the concept maps could help answer inference questions. The teacher put a greater focus on representing the text in the concept map and how to use the maps to answer questions. These enhanced instructions enabled the participants to construct qualitatively different concept maps (see Fig. 2). Concept maps in Experiment 1 were largely a string of nodes—often a node involved a number or adjective that was meaningless, per se. By contrast, concept maps in Experiment 2 often represented the ideas presented by paragraph and branching networks to represent the relationships embedded in the text. Individual student's concept maps were not collected and graded, so students did not get individualized feedback on mapping.

On the day of the experiment, participants were instructed that they would be reading three texts, judging their comprehension of each test (i.e., predicting how they would do on a 5-item multiple choice test), and then completing a test for each text. After having an opportunity to ask questions about the experiment, participants completed their assigned experimental procedure (i.e. they either read and constructed concept maps, read the texts while being given a provided concept map, or simply read the texts). As in Experiment 1, students did not have access to the texts during judgment or testing, but the students in the concept map conditions did have access to the concept maps.

8. Results

8.1. Metacomprehension judgments and test performance

Again, we first report data on metacomprehension judgments and test performance. The median of both metacomprehension judgments and test performances across the three texts was computed for each participant. The mean of the medians was then computed across participants in each group, see Table 1. The mean magnitude of metacomprehension judgments did not differ across groups, nor did mean test performance, both $F(2, 100) < 1$.

8.2. Metacomprehension accuracy

Metacomprehension accuracy was operationalized as in Experiment 1. The mean gamma was computed across participants in the three groups. Twenty-one participants had indeterminate gamma due to invariance in judgments or test performance. As seen in Fig. 3, the group that constructed concept maps had a positive gamma, which was significantly greater than zero, $t(19) = 2.08, p = .05$. Concept map construction contributed more to metacomprehension accuracy here than in Experiment 1. The group that was given concept maps also had a positive gamma correlation, but it did not differ significantly from zero, $t(29) = 1.06$. The control group had a negative gamma correlation but this was not significantly different from zero, $t(27) < 1.00$. More important, metacomprehension accuracy varied across the groups, $F(2, 79) = 3.81, MSe = .61, p = .03, partial eta squared = .11$.

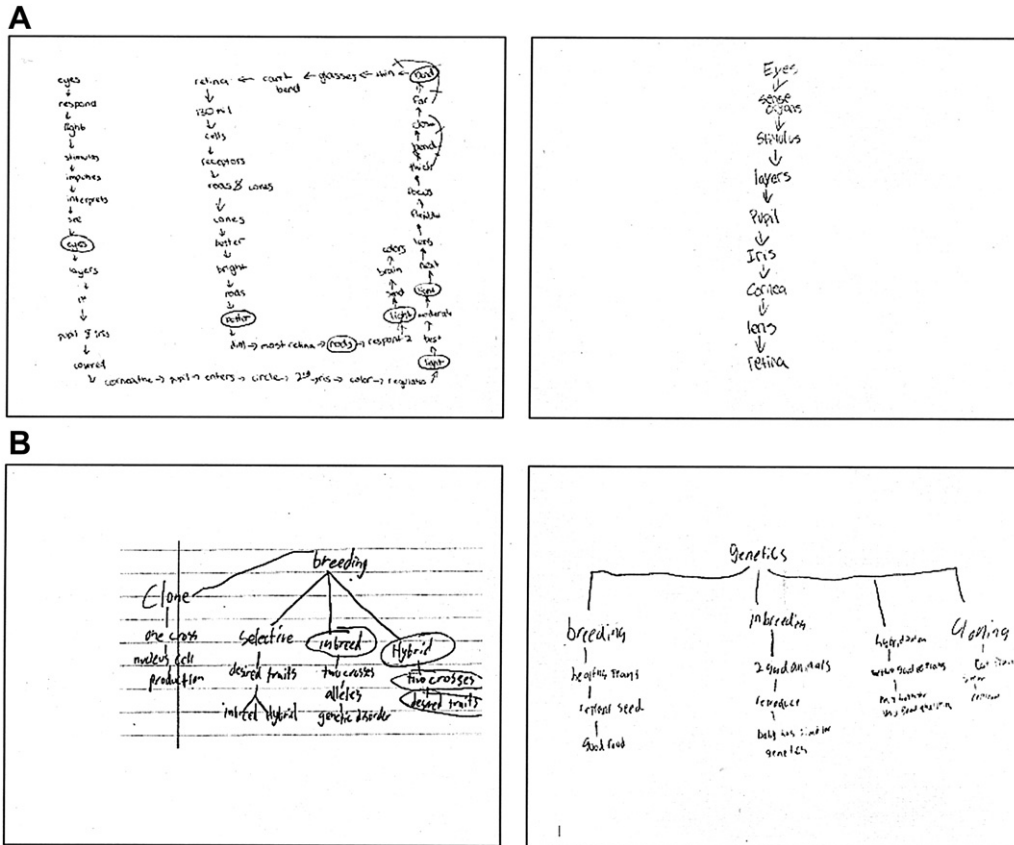


Fig. 2. A. Representative maps from Experiment 1. B. Representative maps from Experiment 2.

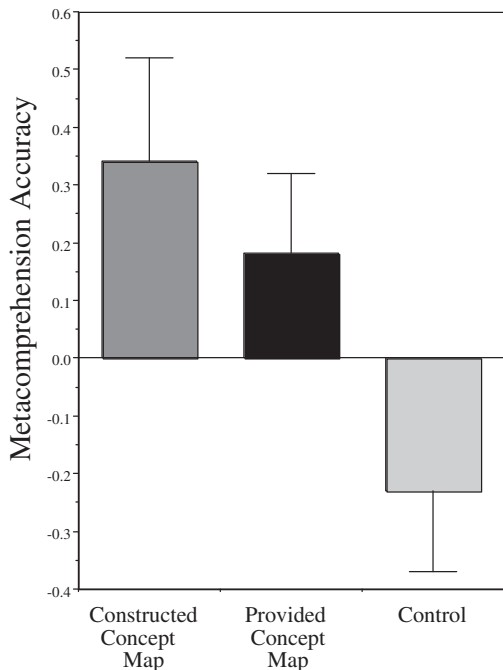


Fig. 3. Mean metacomprehension accuracy (gamma correlation) for Experiment 2. Error bars are the standard error of the mean.

Tukey HSD tests showed that accuracy was greater for the group that constructed maps than for the control group, $p < .05$, $q = 2.67$ (Hypothesis 2).

We also examined the concept map properties following the enhanced instruction. Participants constructed concept maps that were about the same size. However, unlike the maps from Experiment 1, these maps contained fewer redundant nodes, were often organized in a paragraph fashion, and contained more multi-link nodes. These modifications suggest that these maps better reflected the situation model of the texts and also provided a better guide for participants as they gauged their level of metacomprehension accuracy. A one-tailed Pearson correlation found two significant relationships. Test performance was positively correlated with paragraph use in maps, $r = .27$, $p < .05$, and marginally correlated with the number of nodes used, $r = .44$, $p = .07$. This suggests that the same amount of content as evidenced by the number of nodes was more useful when it was better organized.

9. Discussion

These results are promising in that metacomprehension accuracy was improved when classroom instruction provided students with evidence of the utility of using concept mapping, as well as practice with inference tests (Hypothesis 2)—which has recently been shown to improve metacomprehension accuracy in adult readers (Thiede et al., 2011). Moreover, this experiment provides additional evidence that information-generation is most useful to improving metacomprehension. Providing the same information does not improve metacomprehension accuracy to the same extent. This finding is consistent with the finding that simply providing

readers with keywords does not impart the same benefits as generating keywords (Thiede et al., 2005).

The results of the reading alone control condition show that concept map training was not enough to improve the meta-comprehension accuracy of these adolescent students. When students did not have the advantage of generating concept maps, and having them available during judgment and testing, students were not able to take advantage of the training. Having the concept maps available during judgments may help readers in both concept map condition to select relevant cues.

10. General discussion

Models of self-regulated learning describe the important role of accurate metacognitive monitoring in learning. Moreover, accurate monitoring of comprehension is critical to learning from texts (Thiede et al., 2003). According to the cue-utilization framework of metacognitive monitoring (Koriat, 1997), accuracy of metacognitive monitoring is driven by the cues used during monitoring. Use of cues that are predictive of future test performance, will increase monitoring accuracy; whereas, use of cues that are not predictive of test performance will decrease accuracy.

Comprehension of texts requires readers integrate ideas across a text as well as combine these ideas with prior knowledge of a topic. Thus, tests of comprehension should not assess knowledge of the surface features of a text (e.g., details contained in a text), but rather knowledge of the situation model of a text (Wiley et al., 2005). Therefore, getting readers to base their meta-comprehension judgments on the quality of their situation models, or at the very least not base their judgments on other cues such as their interest in the topic, or their memory for details contained in a text, should improve metacomprehension accuracy. Recent interventions that improve metacomprehension accuracy have done just this (for a review of this research see Thiede et al., 2009). For instance, Thiede et al., 2011 focused college readers on the situation model by informing readers of the nature of the comprehension tests and giving them practice with tests that assess knowledge of the situation model for a text. Griffin et al. (2008) focused college readers on the situation model by instructing students to self-explain during reading. These interventions improved metacomprehension accuracy for college students, but have yet to be evaluated with younger students.

The findings from the present research suggest that the meta-comprehension accuracy of 7th grade students also benefits from interventions designed to focus readers on more appropriate cues. Constructing a concept map arguably requires students to generate visual representation of a situation model, which should increase the salience of cues related to the situation model when judging one's comprehension. And using these cues for judging comprehension should improve metacomprehension accuracy. However, generating concept maps may not be enough to improve accuracy—as Experiment 1 showed only marginal improvements in meta-comprehension accuracy associated with constructing concept maps. The results of Experiment 2 show the added benefit of clarifying the nature of tests and getting explicit instruction on the utility of concept mapping. Additional gains may occur if students are guided during concept map construction with prompts such as those used to promote learning while writing. Nuckles, Hubner, and Renkl (2009) used prompts such as “Which main points have I already understood well?” as participants wrote about an observed videotaped lecture. These prompts improved learning beyond what occurred from writing alone. Prompts may offer similar benefits to participants as they construct concept maps.

Although the focus of the present research was meta-comprehension, it is important to comment on comprehension

itself—as concept mapping has been shown to improve comprehension (Nesbit & Adesope, 2006). In two experiments, test performance did not differ across groups. Put differently, concept mapping did not improve comprehension. Even with several short lessons on mapping (i.e., more focused, extended instruction), the concept maps were not particularly high quality—and certainly not universally high quality. Perhaps students needed individualized feedback to guide their mapping activities to become skilled at constructing well developed concept maps.

Another possible explanation for the lack of improvement in comprehension may come from the metacognitive literature. Improving metacomprehension accuracy is important because better monitoring of learning can lead to more effective regulation of study, which in turn can lead to increased learning (e.g., Thiede, 1999; Thiede et al., 2003). In the present research, students were not given an opportunity to use their monitoring to guide regulation of study (e.g., select texts for restudy). Thus, the influence of monitoring was essentially removed. Begg, Martin, and Needham (1992) also showed that learning is not affected by monitoring accuracy when regulation of study is not controlled by the learner.

Metacomprehension research has long focused on college students. The present research suggests younger readers are not adept at monitoring their own comprehension. The negative gamma correlations in the control groups indicate that left to their own devices 7th graders are quite inaccurate at judging their comprehension, which argues for developing new interventions to help young readers develop monitoring skills, especially in the context of attempting to learn from expository texts. The present research is a first step in this direction. Additional research is needed to find new techniques to improve metacomprehension accuracy with younger readers. Further, additional research is needed to examine whether adolescents are able to use metacognitive monitoring to guide subsequent study (regulation of study). Although de Bruin et al. (2011) showed that adolescents can regulate their study more research is needed to further demonstrate this and to examine the effect of monitoring and regulation on reading comprehension.

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305B07460 to Keith Thiede, Jennifer Wiley, and Thomas Griffin. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Appendix 1. Text on viruses

A virus is a tiny, nonliving particle that invades and then multiplies inside a living cell. Viruses are not cells and do not have the characteristics of organisms. The only way in which viruses are like organisms is that they can multiply. Although viruses can multiply, they multiply differently than organisms. Viruses can only multiply when they are inside a living cell. The cells that viruses infect in order to multiply are called host cells. Despite their tiny size, viruses have the ability to cause a lot of damage to cells of other organisms.

All living things are capable of being infected by viruses. One of the best studied viruses infects bacteria. It is called a bacteriophage. In humans, viruses may cause relatively harmless diseases such as cold sores and colds, or life-threatening diseases such as polio and AIDS.

All viruses have two basic parts: a protein coat that protects the virus and an inner core made of genetic material. The proteins on the surface of a virus play an important role during the invasion of

a host cell. Each virus contains unique surface proteins that look like the proteins that the host cell normally needs. The virus attaches itself to special sites on the host that are usually reserved for these proteins. Like keys, a virus's proteins fit only into certain "locks," or proteins on the surface of a host's cells. Because the lock-and-key action of a virus is specific, a certain virus can attach only to one or a few types of cells.

Once a virus attaches to the surface of a host cell, it injects its genetic material into the cell. The virus's genetic material takes over the cell functions and the cell starts to produce the virus's proteins and genetic material. The proteins and genetic material assemble into new viruses that fill the cell. When it is full of new viruses, the host cell bursts open and dies as it releases hundreds of new viruses to infect other cells and the process starts over again.

Test on Viruses(asterisk denotes the correct answer)

- According to information in the passage, which of the following could not be infected by a virus?
 - A virus*
 - A house plant
 - A fungus
 - A dog
- If a virus contains genetic material but does not have a protein coat, then
 - The virus could infect host cells.
 - The virus could reproduce and burst a host cell.
 - The virus could only attach to specific host cells.
 - The virus could not attach to any host cell.*
- What would likely happen if a virus could attach to a host cell, but did not take over the host cell's functions?
 - The virus would not burst the cell.*
 - The virus would live in the host cell.
 - The host cell would produce the virus's proteins.
 - The host cell would produce the virus's genetic material.
- All living things are capable of being infected by viruses because
 - All living things are made of water.
 - The cells of all living things can serve as host cells.*
 - A single virus can infect all types of cells.
 - All living things contain bacteria.
- What is a correct order of processes underlying the spread of a virus?
 - A host cell bursts, a new virus attaches to host cell, a new virus injects genetic material*
 - A virus reproduces, a virus injects genetic material, a virus attaches to host
 - A virus attaches to a host, a host cell bursts, a virus takes over host cell
 - A virus injects genetic material, a virus attaches to host cell, a host cell bursts.

References

- Anderson, M. C. M., & Thiede, K. W. (2008). Why do delayed summaries improve metacomprehension accuracy? *Acta Psychologica*, 128, 110–118.
- Ariel, R., Dunlosky, J., & Bailey, H. (2009). Agenda-based regulation of study-time allocation: when agendas override item-based monitoring. *Journal of Experimental Psychology: General*, 138, 432–447. doi:10.1037/a0015928.
- Baker, J. M. C., & Dunlosky, J. (2006). Does momentary accessibility influence metacognitive judgments? The influence of study-judgment lags on accessibility effects. *Psychonomic Bulletin & Review*, 13, 60–65.
- Begg, I. M., Martin, L. A., & Needham, D. R. (1992). Memory monitoring: how useful is self-knowledge about memory? *European Journal of Cognitive Psychology*, 4, 195–218. doi:10.1080/09541449208406182.
- Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: when retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, 127, 55–68. doi:10.1037/0096-3445.127.1.55.
- Chi, M. T. H., DeLeuw, N., Chiu, M., & Lavancher, C. (1994). Eliciting self-explanation improves understanding. *Cognitive Science*, 18, 439–477.
- de Bruin, A., Thiede, K. W., Camp, G., & Redford, J. R. (2011). Generating keywords improves metacomprehension and self-regulation in elementary and middle school children. *Journal of Experimental Child Psychology*, 109, 294–310. doi:10.1016/j.jecp.2011.02.005.
- Duke, N. K. (2000). 3.6 minutes per day: the scarcity of informational texts in first grade. *Reading Research Quarterly*, 35, 202–224. doi:10.1598/RRQ.35.2.1.
- Dunlosky, J., Hertzog, C., Kennedy, M. R. F., & Thiede, K. W. (2005). The self-monitoring approach for effective learning. *International Journal of Cognitive Technology*, 10, 4–11.
- Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension: a brief history and how to improve its accuracy. *Current Directions in Psychological Science*, 16, 228–232. doi:10.1111/j.1467-8721.2007.00509.x.
- 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. doi:10.1207/s15326950dp4001_2.
- Glenberg, A. M., & Epstein, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 702–718. doi:10.1037/0278-7393.11.1-4.702.
- Glenberg, A. M., Sanocki, T., Epstein, W., & Morris, C. (1987). Enhancing calibration of comprehension. *Journal of Experimental Psychology: General*, 116, 119–136. doi:10.1037/0096-3445.116.2.119.
- Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross classification. *Journal of the American Statistical Association*, 49, 732–764. doi:10.2307/2281536.
- Gravetter, F. J., & Wallnau, L. B. (1999). *Essentials of statistics for the behavioral sciences* (3rd ed.). Pacific Grove, CA: Brooks/Cole Publishing Company.
- 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. doi:10.3758/MC.36.1.93.
- Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, 49, 294–303. doi:10.1037/0003-066X.49.4.294.
- 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. doi:10.1037/0096-3445.126.4.349.
- Leopold, C., & Leutner, D. (2011). Science text comprehension: drawing, main idea selection, and summarizing as learning strategies. *Learning and Instruction*, 1–11. doi:10.1016/j.learninstruc.2011.05.005.
- 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. 117–144). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Maki, R. H., & Berry, S. L. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 663–679.
- Maki, R. H., & Serra, M. (1992). Role of practice tests in the accuracy of test predictions on text material. *Journal of Educational Psychology*, 84, 200–210. doi:10.1037/0022-0663.84.2.200.
- Markman, E. M. (1977). Realizing that you don't understand: a preliminary investigation. *Child Development*, 48, 986–992.
- Markman, E. M. (1979). Realizing that you don't understand: elementary school children's awareness of inconsistencies. *Child Development*, 50, 643–655. doi:10.2307/1128929.
- 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. doi:10.1207/s1532690xci1401_1.
- 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. doi:10.1037/0096-3445.132.4.530.
- Morris, C. C. (1990). Retrieval processes underlying confidence in comprehension judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 223–232. doi:10.1037/0278-7393.16.2.223.
- Nelson, T. O. (1984). A comparison of current measures of feeling-of-knowing accuracy. *Psychological Bulletin*, 95, 109–133. doi:10.1037/0033-2909.95.1.109.
- 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. 125–141). New York: Academic Press.
- Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: a meta-analysis. *Review of Educational Research*, 76, 413–448. doi:10.3102/00346543076003413.
- Nuckles, M., Hubner, S., & Renkl, A. (2009). Enhancing self-regulated learning by writing learning protocols. *Learning and Instruction*, 19, 259–271. doi:10.1016/j.learninstruc.2008.05.002.
- Pearson, P. D., & Duke, N. K. (2002). Comprehension instruction in the primary grades. In C. C. Block, & M. Pressley (Eds.), *Comprehension instruction: Research based best practices* (pp. 247–258). New York: Guilford.
- Pressley, M., Borkowski, J. G., & O'Sullivan, J. T. (1984). Memory strategy instruction is made of this: metamemory and durable strategy use. *Educational Psychologist*, 19, 94–107.
- 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. doi:10.1037/0278-7393.28.1.69.
- Rawson, K., Dunlosky, J., & Thiede, K. W. (2000). The rereading effect: metacomprehension accuracy improves across reading trials. *Memory & Cognition*, 28, 1004–1010.

- Roderer, T., & Roebers, C. M. (2009). Children's strategic regulation of memory accuracy. In M. R. Kelley (Ed.), *Applied memory* (pp. 253–274). Hauppauge, NY, US: Nova Science Publishers.
- Roebers, C. M., von der Linden, N., Schneider, W., & Howie, P. (2007). Children's metamemorial judgments in an event recall task. *Journal of Experimental Child Psychology*, *97*, 117–137. doi:10.1016/j.jecp.2006.12.006.
- 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. doi:10.1002/sce.3730740407.
- Thiede, K. W. (1999). The importance of accurate monitoring and effective self-regulation during multitrial learning. *Psychonomic Bulletin & Review*, *6*, 662–667.
- Thiede, K. W., Anderson, M. C. M., & Theriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, *95*, 66–73. doi:10.1037/0022-0663.95.1.66.
- 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. doi:10.1037/0278-7393.25.4.1024.
- 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. doi:10.1037/0278-7393.31.6.1267.
- Thiede, K. W., Griffin, T., & Wiley, J. (2011). Test expectancy affects metacomprehension accuracy. *British Journal of Educational Psychology*, *81*, 264–273.
- Thiede, K. W., Griffin, T. D., Wiley, J., & Anderson, M. C. M. (2010). Poor metacomprehension accuracy as a result of inappropriate cue use. *Discourse Processes*, *47*, 331–362. doi:10.1080/01638530902959927.
- Thiede, K. W., Griffin, T. D., Wiley, J., & Redford, J. S. (2009). Metacognitive monitoring during and after reading. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 85–106). New York, NY, US: Routledge/Taylor & Francis Group.
- Venezky, R. (2000). The origins of the present-day chasms between adult literacy needs and school literacy instruction. *Scientific Studies of Reading*, *4*, 19–39.
- Weaver, C. A. (1990). Constraining factors in calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 214–222. doi:10.1037/0278-7393.16.2.214.
- 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.
- 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.
- Wiley, J., Griffin, T., & Thiede, K. W. (2005). Putting the comprehension in metacomprehension. *Journal of General Psychology*, *132*, 408–428. doi:10.3200/GENP.132.4.408-428.
- 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). Hillsdale, NJ: LEA.