Metacognition Learning
DOI 10.1007/s11409-009-9052-6

Individual differences in relative metacomprehension accuracy: variation within and across task manipulations

Evelyn S. Chiang · David J. Therriault · Bridget A. Franks

Received: 27 March 2009 / Accepted: 11 November 2009

Abstract In recent decades, increasing numbers of studies have focused on metacomprehension accuracy, or readers' ability to distinguish between texts comprehended more vs. less well. Following early findings that suggested readers are fairly poor at doing so, a number of studies have identified specific tasks to supplement a single reading of text that have resulted in greater metacomprehension accuracy. One assumption underlying these studies is that, in the absence of such tasks, metacomprehension accuracy is uniformly poor, and given their implementation, readers uniformly improve. Here we describe the individual variation that occurs both in the absence (e.g., within a *single* text reading manipulation) and presence (e.g., within a *rereading* or *selective rereading* task manipulation) of these supplementary tasks (N=214), in order to make a case for greater attention to individual differences in metacomprehension accuracy. We also introduce a new manipulation in metacomprehension research, selective rereading, and argue that certain types of tasks may be more likely to reveal individual differences in metacomprehension accuracy due to the nature of the task being more or less demanding on working memory capacity.

Keywords Metacomprehension accuracy · Rereading · Selective rereading · Working memory · Individual differences

In order to successfully learn from text, students must be able to estimate their level of text comprehension with at least some degree of accuracy. Accurate estimations of comprehension or *metacomprehension accuracy* permit more effective regulation of study. For example, identifying content that has been mastered in contrast to content that remains puzzling is necessary for students to selectively focus attention on problematic text material.

E. S. Chiang · D. J. Therriault · B. A. Franks
University of Florida, Gainesville, FL, USA

E. S. Chiang (✉)
Department of Psychology, University of North Carolina at Asheville, CPO 1630,
One University Heights, Asheville, NC 28804, USA
e-mail: echiang@unca.edu

Studies of metacomprehension accuracy have consistently reported that readers are poor at distinguishing between texts they have comprehended more vs. less well (e.g., Dunlosky and Lipko 2007; Maki 1998). Maki (1998), for instance, reported a mean metacomprehension accuracy of only .27 across 25 studies conducted in her laboratory. Given that metacomprehension accuracy ranges between −1 (achieved when the reader is wholly inaccurate in identifying texts comprehended more vs. less well) and +1 (achieved when the reader perfectly distinguishes between texts comprehended more vs. less well), a mean of .27 can be considered quite low and not only perplexing but perhaps even alarming in a population of presumably proficient, college-aged readers.

One response to the finding of low average metacomprehension accuracy across numerous studies has been an attempt to identify ways to improve metacomprehension accuracy. A number of manipulations have been quite successful in this regard. Specifically, tasks that supplement a single reading of text have resulted in greater mean metacomprehension accuracy. Rawson et al. (2000) and Dunlosky and Rawson (2005) demonstrated that rereading, or directing participants to read texts twice in entirety in succession, resulted in greater metacomprehension accuracy. Other tasks such as summarizing text content (Thiede and Anderson 2003) and generating keywords for text (Thiede et al. 2003) were successful as well in yielding greater mean metacomprehension accuracy among study participants. More recently, Dunlosky and Lipko (2007) took a new tack to exploring metacomprehension accuracy by having participants make more specific judgments (e.g., about specific text terms) rather than global ones (e.g., about the entire text). They found that directing participants to verbally recall term definitions resulted in greater metacomprehension accuracy compared to when participants were not directed to recall the definitions.

Altogether, these manipulations have focused on task-oriented variables that have globally led to either inferior or superior performance. For instance, Rawson et al. (2000) and Dunlosky and Rawson's (2005) finding that rereading leads to superior performance compared to single reading was explained in terms of a rereading effect, which posited that greater accuracy results from the availability of cues at second readings that are more predictive of performance than those at first readings. Similar explanations were offered for improved performance resulting from generating keywords for text (Thiede et al. 2003) and summarizing text content (Thiede and Anderson 2003). And, having participants recall definitions by typing them out provided greater metacomprehension accuracy compared to when they did not (Dunlosky and Lipko 2007).

Underlying these experimental manipulations has been a basic assumption: In the absence of directives to engage in activities supplementing a single reading of text, readers are on the whole fairly poor in their metacomprehension accuracy; however, given particular tasks to perform, readers uniformly improve. In other words, the assumption is that there is a one-size-fits-all means of improving metacomprehension accuracy. Virtually any classroom teacher, however, would question whether this is the case in learning: Learners vary, and different strategies or tactics work for different students.

This inattention to individual differences in metacomprehension accuracy may seem surprising, but on the whole, attempts to identify personal factors associated with metacomprehension accuracy have not been particularly fruitful. Maki and colleagues (Maki et al. 1994; Maki et al. 2005), for example, found no evidence for a relationship between verbal ability and metacomprehension accuracy, and early work by Glenberg and Epstein (1987) indicated that, contrary to expectations, readers with greater expertise in a domain did not benefit from their prior knowledge in terms of being more accurate.

Another individual differences variable, working memory capacity (WMC), has been tied to many cognitive tasks (Engle 2002) such as reading comprehension (Daneman and

Carpenter 1980) but has been largely unexplored with regard to metacomprehension accuracy until very recently (Griffin et al. 2008). Engle (2002) attributed the relationship between WMC and higher-order cognitive tasks such as reading to domain-free executive attention, or the ability to control attention. This ability to control attention can be thought of either in terms of the ability to selectively attend to relevant information while ignoring what's irrelevant or the ability to attend simultaneously to multiple sets of stimuli (Colflesh and Conway 2007). In the context of metacomprehension, both views are applicable: The reader must be capable of selective attention (i.e., attend to cues indicative of adequate vs. insufficient understanding by maintaining the goal of monitoring comprehension accurately) and divided attention (i.e., comprehend and monitor comprehension at the same time) (see Griffin et al. 2008).

Using Nelson and Narens's (1990) model of metacognition as a guide, we assume that cognitive processes are occurring at two different levels, i.e., the meta-level and the object-level. Under this model, it would be reasonable to expect differences in metacognitive performance as a function of WMC. In the context of text comprehension, metacomprehension would involve processes at the meta-level (e.g., making a judgment of learning) and comprehension at the object-level (e.g., constructing a situation model or text representation). Information flows both from the object-level to the meta-level, as well as from the meta-level to the object-level. Nelson and Narens described the former as monitoring and the latter as control. Monitoring with regard to reading text would involve processes such as gauging one's level of understanding, whereas control would involve processes such as slowing down or looking back in text when a breakdown in comprehension occurred. Relative metacomprehension accuracy, then, would depend greatly on the flow of information from the object-level to the meta-level, as the latter contains a representation of and is informed by the former. Theoretically, higher span readers should be more accurate than their lower span counterparts, because they would be better at shifting resources between comprehension and metacomprehension processes. However, differences in WMC do not necessarily translate into differences in performance in all tasks, as tasks vary in terms of the demand placed on working memory.

For instance, using a Stroop task, Kane and Engle (2003) found that low WMC individuals' performance suffered given conditions that involved high congruence between word name and ink color but not under conditions of low congruence. In the low congruence condition, individuals were constantly being reminded to suppress reading the word to name the color because the two rarely matched; in contrast, in the high congruence condition, the two often matched and so this goal was not reinforced, leading to more error for low span individuals in the cases of mismatch. Similarly, Colflesh and Conway (2007) found WMC effects in a dichotic listening task when task demands were higher but no effects when task demands were lower.

In terms of metacomprehension accuracy, Griffin et al. (2008) proposed that metacomprehension is a concurrent but secondary process to comprehension. Simply put, readers with higher WMC should have greater metacomprehension accuracy compared to their lower span counterparts because they are better able to maintain the goal of monitoring comprehension as well as shift attention between comprehension and metacomprehension. However, given a condition that reduces the demands of comprehension or that facilitates metacomprehension, WMC effects should disappear. This is what Griffin et al. (2008) found: When low span readers were able to reread, they performed as well as their higher span counterparts in terms of metacomprehension performance; however, when they read texts a single time, low span readers' performance suffered. Griffin et al. (2008) also tested a self-explanation manipulation, in which all readers benefited. They explained that self-explanation highlighted cues more relevant to global comprehension and that readers then based their judgments on these cues.

Whereas Rawson et al. (2000) and Dunlosky and Rawson (2005) found that rereading improves metacomprehension accuracy compared to single reading of text, Griffin et al. (2008) found that only low span readers benefit from rereading. Griffin et al. explained that rereading reduces the need for concurrent processing, which allowed the low span readers to perform as well as their high span counterparts.

Our purposes here are to make an argument for individual differences in metacomprehension accuracy. The commonly accepted practice is to average gamma across study participants in the same condition or experimental manipulation; hence it is assumed that a particular manipulation improves performance uniformly. Our results indicate this is not the case: We observed substantial variation in accuracy within each study condition. Our secondary purpose is to highlight a context in which individual differences are more likely to emerge: selective rereading, due to the greater task demands present with selective rereading compared to either single or rereading. As expected, working memory effects emerged here whereas they did not in the other two conditions. We conclude by advocating not only greater attention to individual variation in metacomprehension accuracy but also the identification of different types of tasks to improve metacomprehension accuracy for different students.

Method

Participants

Two hundred fourteen college undergraduates (159 female, 55 male) enrolled at a large, southeastern public university participated for partial course credit or extra credit. Fifteen percent were college freshmen, 25% were sophomores, 34% were juniors, and 25% were seniors. Sixty-nine percent of the participants reported their race as Caucasian, 11% as African-American, 6% as Asian/Pacific Islander, and 9% as Hispanic; 4% self-identified as a category “Other” than those listed, and one participant declined to report race. The majority of participants ranged in age from 18 to 23 years; in addition, there was one participant each at 24, 25, 32, and 40 years of age. Mean age was 20.04 years.

Conditions

We employed three manipulations in this study: single reading, rereading, and selective rereading. Rereading, or reading texts in entirety twice in succession, is a manipulation commonly used in the metacomprehension literature that is usually compared to single reading. We wished to examine the variation that occurs in both single reading and rereading. We also introduce a new manipulation to the metacomprehension literature: selective rereading, or directing readers to actively reinstate previously read text in order to improve comprehension. In this manipulation, readers used keyboard arrows to look back or search through previously read text as they read and monitored their understanding. Many studies of metacomprehension accuracy have restricted readers to a forward-only progression through text. That is, text has been typically presented line-by-line on a computer screen, and after reading a sentence and moving on to the next screen, the reader is unable to return to previously read text. The selective rereading manipulation not only provides readers with the ability to look back at previously read text but also actively encourages them to do so. As a manipulation, it is presumably more demanding because readers are being directed to actively monitor their reading. The majority of participants (over 92%) in the selective rereading condition employed the look back procedure during reading.

Texts

Six expository texts adapted from a GRE preparation manual (Dunlosky and Rawson 2005; Rawson et al. 2000) were used in this study. The texts cover a range of content, including topics on politics, literature, inventions, intelligence, guilt, and obesity. Six comprehension questions accompany each text. Half of the questions tap information that is explicitly stated in the text (memory-based questions); the other half focuses on information that could be inferred from the text (inference questions; Dunlosky and Rawson 2005). Sample memory-based and inference questions are presented in the "Appendix". One text (politics) was used as a practice text and the remaining five as experimental texts. The same five texts were used as the experimental texts across all conditions. The number of words in the experimental texts ranged from 358 to 601, with an average of 497.6 words. The number of sentences ranged from 16 to 27, with an average of 20.8 sentences. Texts were presented line-by-line on a computer screen using E-prime software. Participants controlled the rate of sentence presentation by pressing a keyboard button to advance to the next screen. In the selective rereading condition, pressing another keyboard button permitted readers to return to previously read sentences within the text. In the other two conditions (single reading, rereading), readers were unable to return to previously read sentences once they had advanced to the next sentence.

Ratings

Participants predicted their test performance by responding to the question, "How well do you think you will be able to answer a test question over this material in about 20 min? 0 (definitely won't be able), 20 (20% sure I will be able), 40 (40% sure I will be able), 60 (60% sure I will be able), 80 (80% sure I will be able), and 100 (definitely will be able; Rawson et al. 2000)." Ratings were coded on a Likert-type scale such that 0 (definitely won't be able) corresponded to 1 and 100 (definitely will be able) corresponded to 6.

Procedures

Participants were recruited from psychology or educational psychology classes. Upon arrival at the study session, participants were randomly assigned to one of three conditions: single reading, rereading, or selective rereading. Study sessions lasted approximately 75–90 min. In the single and rereading conditions, readers were able to advance forward only through the texts; they were not able to return to previously read sentences while reading each text. In the selective rereading condition, readers were able to return to previously read sentences within a text. All participants first completed a practice session in which they read one shortened text, predicted test performance, and answered two comprehension questions (one inference question and one memory-based question). Feedback was not provided. In the single and selective rereading conditions, participants then read each of the five experimental texts. Immediately after reading each text, they rated their learning of that text. After reading all of the texts, they completed six comprehension questions for each text. In the rereading condition, participants read the five texts; immediately after finishing the fifth text, they read each of the texts again in the same order. They rated their learning of each text immediately after reading it for the second time. After reading all of the texts twice, they answered the comprehension questions for all texts. There were no time limits placed on either reading or the comprehension test. Prior to reading the experimental texts, participants were told they would either be reading the texts a single time but could not return to previously read sentences within the text (single reading condition), twice in succession without the ability to return to previously read sentences within the text (rereading condition), or a single time with the additional directive to actively reinstate text whenever necessary for thorough comprehension (selective rereading condition).

After completing the reading portion of the study, participants completed the Automated Operation Span (Aospan) task (Conway et al. 2005; Unsworth et al. 2005). This task is an automated version of the Operation Span task that presents participants with math operations to solve followed by letters to recall. Participants must recall the string of letters in order after each set of math operations; sets range from three to seven items. Participants who fail to solve the math operations with at least 85% accuracy are eliminated from analyses, as this failure implies they were not engaged in both parts of the task (solving operations, recalling letters) (Unsworth et al. 2005). At the conclusion of the task, absolute and total working memory (WM) span scores are calculated. The total score reflects the total number of letters recalled in its correct position within a particular string. The absolute score takes set size into consideration. Only those sets with all letters recalled correctly are included in the absolute score. In this study, absolute scores were used as a measure of WM span.

Metacomprehension accuracy

Metacomprehension accuracy was computed by calculating a gamma correlation coefficient for each participant. When calculating a gamma coefficient, dyads are the unit of analysis (Gonzalez and Nelson 1996). Gamma is calculated by subtracting the number of discordant pairs from the number of concordant pairs, and then dividing this difference by the sum of the number of concordant and discordant pairs, or $G=(C-D)/(C+D)$.

Gonzalez and Nelson (1996) defined a discordant pair as the event in which one member of the pair exceeds the other member of the pair in terms of prediction value but then is exceeded by the second member in terms of criterion performance. A concordant pair, in contrast, is the case in which both the prediction and the criterion of one member of the pair exceed that of the other member of the pair. All dyads that contain ties, on either the predictor or criterion variable, are not considered. For each text in this study, the judgment of learning was the prediction variable and the number of comprehension questions answered correctly the criterion variable. With five experimental texts, there were 10 possible dyads.

Results

Data were discarded for one participant who was not fluent in English and one participant who failed to follow directions in the reading portion of the study. In all, 69 participants read texts a single time, 75 participants read texts twice in succession, and 70 participants read texts selectively.

Gamma correlations cannot be computed for participants who make the same prediction for all texts. Nor can they be computed when participants achieve the same comprehension score on all texts. Four participants in the single reading condition had indeterminate gammas, 5 in the rereading condition, and 4 in the selective rereading condition. In all, there were 65 participants in the single reading condition, 70 in the rereading condition, and 66 in the selective rereading condition with calculable gamma correlations.

For all analyses involving WM span scores, data from participants who responded incorrectly to more than 15% of the math problems were discarded. Thirteen individuals had a math operation error rate above 15%. For all analyses including WM span scores, this left a total of 58 participants in the single reading condition, 66 participants in the rereading condition, and 64 in the selective rereading condition with calculable gamma correlations (final $n=188$).

Within-condition variation

As noted previously, many studies on metacomprehension accuracy have centered around identifying a task or manipulation with potential for improving metacomprehension accuracy. Participants in these studies were assigned to either engage in the task or not, and resulting mean metacomprehension accuracy was compared between the conditions. This approach has been fruitful in identifying tasks that improve overall metacomprehension accuracy. But our purposes here are to draw attention to the individual variation that occurs within tasks or conditions rather than compare average performance across conditions.

We report results of a one-way analysis of variance (ANOVA) to illustrate the need to attend to variation with contexts. In this ANOVA, we compared mean metacomprehension accuracy of readers who read texts a single time, twice in entirety in succession, and selectively. Results indicated that the groups did not differ in metacomprehension accuracy, F (2, 198)=.757, p=.470. Table 1 displays the means and standard deviations for metacomprehension accuracy across the three reading conditions.

The ANOVA results would suggest that there were no differences among the three reading conditions in terms of metacomprehension accuracy, or in other words, a failure to replicate the rereading effect found by Rawson et al. (2000) and Dunlosky and Rawson (2005), as well as the failure to establish selective rereading as a means of improving metacomprehension accuracy. However, upon closer examination, the large standard deviations within each reading condition indicate substantial individual variability within each condition.

We have broken down the frequency of gamma values in each condition and displayed these values in Table 2. This frequency distribution is highlighted for several reasons. First, it is apparent that, regardless of condition, many readers have high metacomprehension accuracy. Given the typical findings in the metacomprehension literature, a score of .5 or above is considered high. By this measure, approximately half of the participants in each condition are quite accurate. Secondly, regardless of condition, a handful of participants are

Table 1 Means and standard deviations for metacomprehension accuracy (gamma) by reading condition

Condition	Gamma	
	M	*SD*
Single reading[a]	.30	.57
Rereading[b]	.41	.61
Selective rereading[c]	.32	.59

[a] $n=65$

[b] $n=70$

[c] $n=66$

highly, or rather perfectly, inaccurate, achieving gammas of −1. This leaves about half the participants in each condition who are neither accurate nor inaccurate.

What does this mean for education? We cannot prescribe one means of improving metacomprehension accuracy due to learner variability. What does this mean for metacomprehension research? We cannot continue making the statement that "readers are poor at metacomprehension accuracy" because about half of them just aren't—even in the absence of being directed to perform tasks intended to improve accuracy.

Given evidence that some readers are accurate whereas others are not, the question of why immediately follows. Attempts to identify sources of variation have not been particularly fruitful—but perhaps because certain tasks highlight individual differences whereas others do not. We turn now to focus on working memory span as an individual differences variable that emerges in some contexts but not others.

WM effects in certain contexts

Prior work indicates that even though working memory is an important individual differences variable to consider, it may not emerge as an explanatory variable depending on the nature of the task being performed. We posited that WM would emerge as an individual differences variable in the selective rereading but not the single or rereading conditions. Selective rereading is a more demanding task because readers are directed to actively monitor their comprehension and reinstate previously read text in order to maximize performance on a test of comprehension following the reading. In other words, with selective rereading readers actively engage in the task of monitoring and self-regulation as they read. Although it might be expected that monitoring comprehension is an assumed concurrent process to comprehension (i.e., readers should monitor their comprehension as they are reading), explicitly directing readers to monitor highlights the act of monitoring and draws greater awareness and attention to this process. In particular, the act of control, or flow of information from the meta-level to the object-level (Nelson and Narens 1990) is highlighted.

Metacomprehension accuracy Our main goal here was to examine the relationship between WM and metacomprehension accuracy in the conditions of single reading, rereading, and selective rereading. We predicted that there would be no relationship between working memory and metacomprehension accuracy in either the single reading or rereading conditions but that this relationship would be evident in the selective rereading condition due to the greater task demands of the selective rereading condition. As a manipulation, rereading allows lower span readers to compensate for their smaller spans; thus, the

Table 2 Frequency of gammas by reading condition

Gamma value	Reading condition		
	Single reading	Rereading	Selective rereading
1	14	27	15
0.5–0.8	16	9	16
−0.4–0.4	30	28	28
−0.8 to −0.5	1	1	2
−1	4	5	5

performance of higher and lower span readers is comparable (see Griffin et al. 2008). Therefore, we did not expect any variation in metacomprehension accuracy due to WM span in the rereading condition. For participants who read texts a single time, we also did not anticipate finding a relationship between WM span and metacomprehension accuracy. Given the argument that working memory effects would not emerge in a rereading condition due to the compensatory nature of rereading, it may seem reasonable to expect that these effects would be evident in a reading condition that does not permit similar compensation (i.e., the single reading condition). However, working memory effects are more likely to emerge under more demanding task conditions. Compared to selective rereading, single reading is less demanding. Selective rereading is far more taxing than single reading in that readers are directed to actively monitor as well as regulate their learning. As such, selective rereading should highlight working memory differences whereas single reading should not.

We ran three separate regression analyses in order to examine the relationship between WM span and metacomprehension accuracy. As the conditions varied in terms of variables of interest, running a single model with dummy codes for each condition would have been inappropriate. In the rereading condition, we anticipated no relationship between WM and metacomprehension accuracy but were interested in the relationship between reading time and metacomprehension accuracy. Specifically, in this condition, we had two reading time measures: the amount of time during the first pass at reading (initial reading) and the amount of time in the second pass at reading (rereading). For the single reading condition, we were also interested in the relationship between reading time and metacomprehension accuracy but had only a single measure of reading time as participants read texts only a single time. For the selective rereading condition, in addition to WM, we were interested in the relationship between metacomprehension accuracy and the number of lookbacks, or text reinstatements, that readers made.

Following our initial analyses for each study condition, we computed additional regression analyses in order to examine the possibility of interactions between the independent variables. Specifically, in the rereading condition, we examined whether there was an interaction between initial reading and rereading time, initial reading time and WM span, or rereading time and WM span. In the single reading condition, we looked for an interaction between reading time and WM span, and in the selective rereading condition, between number of text reinstatements and WM span. These analyses were performed to investigate the possibility that the relationship between metacomprehension accuracy and reading time or text reinstatements would vary as a function of WM span. For example, it is possible that high span readers would be accurate regardless of how many times they looked back in text, whereas the accuracy of low span readers would depend on the number of text reinstatements. However, none of the analyses examining potential interaction effects in any of the three conditions indicated that interactions occurred. Below we report only the results of our initial analyses.

Rereading For the rereading condition, a standard multiple regression analysis was performed between the dependent variable (metacomprehension accuracy) and the independent variables (initial reading time, rereading time, and WM span).

Regression analysis revealed that the model significantly predicted metacomprehension accuracy, $F(3, 62)=3.05$, $p=.035$, $R^2=.13$, adjusted $R^2 = .09$. In terms of individual relationships between the independent variables and metacomprehension accuracy, initial reading time ($t=-2.72$, $p=.008$) significantly predicted metacomprehension accuracy but rereading time ($t=.452$, $p=.653$) and WM span ($t=.798$, $p=.428$) did not. In other words, spending less time reading during the first pass was associated with greater accuracy with regard to metacomprehension. It is possible that spending more time on reading is due to greater difficulties with decoding and comprehension and therefore less monitoring. Time spent on initial reading and rereading were moderately correlated ($r=.49$, $p<.01$). Table 3 summarizes the results of the regression analysis.

Single reading It was not expected that WM would be related to metacomprehension accuracy for readers in the single reading condition. A standard multiple regression analysis was performed between the dependent variable (metacomprehension accuracy) and the independent variables (reading time, WM span). Regression analysis revealed that the model did not predict metacomprehension accuracy, $F(2, 55)=.977$, $p=.383$. In other words, neither reading time nor WM span accounted for individual differences in metacomprehension accuracy in this condition.

Selective rereading The selective rereading condition highlighted the goal of monitoring by directing readers to actively reinstate previously read text whenever necessary. As a manipulation, it was more demanding compared to single reading or rereading. Thus we expected that WM would be positively related to metacomprehension accuracy in that readers with higher spans should be more accurate than those with lower spans. We were also interested in whether number of text reinstatements (lookbacks) would be related to metacomprehension accuracy. In the selective rereading condition, readers reinstated text at will. Five individuals did not reinstate text at all; they were removed from analysis because they did not reinstate text and instead read as did individuals in the single reading condition.

A standard multiple regression analysis with metacomprehension accuracy as the dependent variable and WM span and number of text reinstatements (lookbacks) as the independent variables was performed. Results indicated that the model significantly predicted metacomprehension accuracy, F(2, 56)=3.41, $p=.040$, $R^2=.11$, adjusted $R^2 = .08$. In terms of individual relationships between the independent and dependent variables, WM span ($t=2.19$, $p=.032$) significantly predicted metacomprehension accuracy and lookbacks ($t=-1.63$, $p=.11$) marginally predicted metacomprehension accuracy. The positive relationship between WM and metacomprehension accuracy was expected. The negative relationship between number of lookbacks and metacomprehension accuracy might be interpreted as follows: A greater number of lookbacks might be associated with greater breakdowns in processing text, and thus the necessity of focusing more on comprehension than metacomprehension. Table 4 summarizes the regression model.

Table 3 Summary of regression model for initial and rereading times (RT) and working memory (WM) predicting metacomprehension accuracy in the rereading condition

Variable	B	SE	β	*t*	*p*
Initial RT	−0.037	.014	−.368	−2.72*	.008
Rereading RT	0.009	.021	.061	.452	.653
Working memory	.003	.004	.093	.782	.437

$n=66$

*$p<.05$

Table 4 Summary of regression model for working memory (WM) span and lookbacks predicting metacomprehension accuracy in the selective reading condition

Variable	B	SE	β	*t*	*p*
Working memory	.010	.005	.274	2.13*	.037
Look backs	−.002	.001	−.180	−1.18	.242

n=59, *p<.05

Discussion

Results of this study reveal substantial individual variation in metacomprehension accuracy regardless of whether readers read texts a single time, twice in succession, or selectively with greater attention to monitoring and self-regulation. Evidence of this individual variation contrasts with earlier assumptions that all college-aged, presumably proficient readers have poor metacomprehension accuracy and must rely on particular tasks for improvement. Our results indicate instead that some readers are quite accurate whereas others are inaccurate regardless of whether they read texts once, twice, or selectively.

We explored the possibility that working memory may account for individual differences in metacomprehension performance. At the same time, we expected that certain study conditions would be more conducive than others in revealing individual differences due to working memory. In the selective rereading condition, a relationship between working memory and metacomprehension accuracy emerged, whereas in the single and rereading conditions these effects were absent.

Griffin et al. (2008) explained that rereading allows lower span individuals to compensate for their lower spans and perform comparably with their higher span counterparts. Thus, as a manipulation, rereading would not elicit working memory effects. In contrast, the condition of selective rereading is one that is presumably more taxing, as it directs the reader to actively monitor comprehension and repair breakdowns by actively reinstating text. As such, the condition would be more likely to reveal working memory effects. This is what we found.

Furthermore, results of this study indicated a negative relationship between reading time and metacomprehension accuracy in the rereading condition, as well as a marginally negative relationship between number of lookbacks and metacomprehension accuracy in the selective rereading condition. The negative relationship between reading time and metacomprehension accuracy may appear puzzling, as it indicates that reading longer is associated with lower accuracy. It might be expected that increasing one's study time ought to result in more favorable results. However, the difference in processing between initial reading and rereading may account for this finding. Millis et al. (1998) described accessing word meanings, establishing a textbase, and building a situation model as three processes involved in understanding discourse. They proposed that compared to initial readings, readers focus less on the textbase and more on the situation model at rereading. If a reader requires greater amounts of time during a first pass at reading, it is likely due to difficulties in text decoding and comprehension. Given these difficulties, the reader would have fewer resources available for monitoring. However, rereading permits readers to focus on the situation model, or more global text representation, at the second pass. As noted by Griffin et al. (2008), the rereading paradigm allows readers to compensate for smaller working memory spans. This compensation, effected at the second pass at reading, does not change the characteristics of the first pass at reading (i.e., length of reading times). Thus, while metacomprehension accuracy is not tied to working memory, initial reading time can be.

With regard to the marginally negative relationship between lookbacks and metacomprehension accuracy, a similar explanation may apply. For instance, readers who look back more often might be those who experience more breakdowns in comprehension. Attending to comprehension breaks, in a limited resources paradigm, would leave little remaining for metacognitive monitoring. Thus, a negative relationship between lookbacks and metacomprehension accuracy would ensue. However, it is unclear why readers look back. Are readers striving to maintain local coherence or are they seeking a more global understanding? If the latter, then greater metacomprehension accuracy should result from increased lookbacks, as the reader should be more aware of having attained (or not) adequate overall understanding. However, if looking back is driven mostly by the need for repairs of breakdowns in comprehension between individual sentences rather than the text as a whole, then again this focus on comprehension would draw resources away from more effective monitoring.

Several models that have emerged out of research focusing on readers' judgments of difficulty may potentially provide direction for examining this issue. Metcalfe's (Metcalfe 2002, 2009) region of proximal learning model, for instance, suggests that learners allocate study time strategically in order to maximize the time available for study. Dunlosky and Hertzog's (1998) discrepancy-reduction model, on the other hand, suggests that study continues as long as the perceived degree of learning has not yet matched the desired degree of learning. With regard to selective rereading, the region of proximal learning model might suggest that readers look back in order to make the smallest, easiest repairs in comprehension whereas the discrepancy-reduction model might imply that efforts are devoted towards the more difficult items—in this case, global understanding (see Metcalfe 2009, for a comparison of the models). In the present study, all judgments were made at the conclusion of reading; however, these two models may be useful in understanding readers' decisions to look back in text as they read and how number of text reinstatements could be related to metacomprehension accuracy. Although more information is clearly needed, it may be possible that a negative relationship between look backs and accuracy may be attributed to readers' focus on repairing simple comprehension breaks rather than on achieving a more global understanding.

Implications

The results from this study argue that we should revisit the presumption that (generally) readers' metacomprehension accuracy is low. We obtained evidence across all study conditions (i.e., whether readers read once, twice, or selectively) that a number of readers were highly accurate. Furthermore, our findings suggest that supplementary activities designed to improve metacomprehension accuracy may be effective for some individuals but not others.

Our evidence suggests that, at least with regard to the selective rereading manipulation employed in this study, a task may tax the learner and impair rather than facilitate metacomprehension performance. Tasks that are likely to impair performance are those that are more demanding with regard to working memory. Rereading, on the other hand, appears to be an effective way for low span readers to compensate for their lower working memory spans.

In terms of future research, our results indicate that there is substantial variation in metacomprehension accuracy regardless of whether readers read once, twice, or while actively monitoring comprehension and reinstating previously read text. Therefore, averaging gamma correlations across study participants in each condition led to null results when study conditions were compared. These findings are at odds with previous studies that have uncovered substantial effects in which readers directed to perform supplemental tasks achieved greater accuracy, on average, compared to those not similarly directed. By and large, these studies did not include a breakdown of the proportion of readers in each study condition that achieved high, moderate, or low metacomprehension accuracy. Instead, only averages for each study condition were reported. It is possible that there was variation in the control conditions (usually the single reading condition without any supplementary tasks), and that this variation averaged out once all participants' gamma correlations were computed. At the same time, it is possible that the supplementary tasks improved performance enough for enough participants that average performance was heightened.

The present work provides sufficient evidence of individual variation in metacomprehension accuracy to argue that these differences should not be ignored. Certain experimental conditions may be more vs. less likely to expose individual differences effects because of the nature of the tasks being employed, complicating the matter of identifying sources of systematic variation in performance. Therefore, researchers may want to revisit how they make use of averages of metacomprehension accuracy across study conditions.

Limitations

Several limitations to this work should be considered. First, the single and rereading conditions presented text in a forward-only manner. We did this in order to be consistent with previous studies of metacomprehension accuracy that employed similar text presentation methods. One might argue that this is not how readers typically read; in fact, there is evidence that readers naturally look back as they read (e.g., Rayner and Sereno 1994). The selective rereading condition, then, would be the closest approximation to natural reading conditions. However, the fact that metacomprehension accuracy was similar across the three conditions (i.e., in terms of the proportion of readers that were highly accurate, highly inaccurate, and neither accurate nor inaccurate in each condition) suggests that the forward-only method of text presentation is not as constraining as might be assumed. However, it is possible that readers adjust their processing in response to more vs. less constraining reading conditions, and future research might address this possibility.

Although the majority of participants in the selective rereading condition employed the look back procedure, a small percentage (about 8%) did not. It is unclear why these readers progressed through text in a forward-only fashion. Much remains to be explored regarding readers' reasons for looking back in text.

Appendix

Sample text comprehension questions

The following two questions accompany the text on obesity.

Memory-based question

"According to the statistics presented in the passage,

a. 20% of men and 30% of women in America are obese

b. 20% of men and 40% of women in America are obese
c. 30% of men and 30% of women in America are obese
d. 30% of men and 40% of women in America are obese (*)
e. 40% of men and 30% of women in America are obese" (Dunlosky and Rawson 2005)

Inference question
"It can be inferred from the passage that

a. atherosclerotic people also suffer from obesity
b. following a careful weight-loss diet is the only effective cure for obesity
c. bringing the body into a condition of negative nitrogen balance will assist the dieter in achieving weight loss
d. the roots of obesity are to be found in the feeding and eating problems of infancy and childhood (*)
e. psychiatric treatment can uncover the underlying causes of obesity" (Dunlosky and Rawson 2005)

References

Colflesh, G. J. H., & Conway, A. R. A. (2007). Individual differences in working memory capacity and divided attention in dichotic listening. *Psychonomic Bulletin & Review, 14*, 699–703.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*, 769–786.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior, 19*, 450–466.

Dunlosky, J., & Hertzog, C. (1998). Training programs to improve learning in later adulthood: Helping older adults educate themselves. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 249–275). Mahwah: Erlbaum.

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

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

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*, 19–23.

Glenberg, A. M., & Epstein, W. (1987). Inexpert calibration of comprehension. *Memory & Cognition, 15*, 84–93.

Gonzalez, R., & Nelson, T. O. (1996). Measuring ordinal association in situations that contain tied scores. *Psychological Bulletin, 119*, 159–165.

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.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132*, 47–70.

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: Erlbaum.

Maki, R. H., Jonas, D., & Kallod, M. (1994). The relationship between comprehension and metacomprehension ability. *Psychonomic Bulletin & Review, 1*, 126–129.

Maki, R. H., Shields, M., Wheeler, A. E., & Zacchilli, T. L. (2005). Individual differences in absolute and relative metacomprehension accuracy. *Journal of Educational Psychology, 97*, 723–731.

Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning? *Journal of Experimental Psychology: General, 131*, 349–363.

Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science, 18*, 159–163.

Millis, K., Simon, S., & tenBroek, J. (1998). Resource allocation during the rereading of scientific texts. *Memory & Cognition, 26*, 232–246.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory, vol. 26* (pp. 125–173). New York: Academic.

Rawson, K. A., Dunlosky, J., & Thiede, K. W. (2000). The rereading effect: metacomprehension accuracy improves across reading trials. *Memory & Cognition, 28*, 1004–1010.

Rayner, K., & Sereno, S. C. (1994). Eye movements in reading: Psycholinguistic studies. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 57–81). Mahwah: Erlbaum.

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.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*, 498–505.