Contemporary Educational Psychology 34 (2009) 30–41

Contents lists available at ScienceDirect

Contemporary Educational Psychology

journal homepage: www.elsevier.com/locate/cedpsych

The limited benefits of rereading educational texts

Aimee A. Callender *, Mark A. McDaniel [1]

Psychology Department, Box 1125, Washington University, 1 Brooking Drive, St. Louis, MO 63130, USA

ARTICLE INFO

Article history:
Available online 31 August 2008

Keywords:
Rereading
Comprehension
Study methods
Text comprehension
Reading

ABSTRACT

Though rereading is a study method commonly used by students, theoretical disagreement exists regarding whether rereading a text significantly enhances the representation and retention of the text's contents. In four experiments, we evaluated the effectiveness of rereading relative to a single reading in a context paralleling that faced by students in the classroom. Participants read educational texts (textbook chapters or a *Scientific American* article) under intentional learning instructions. Learning and memory were tested with educationally relevant summative assessments (multiple choice, short-answer questions, and text summaries). With only several exceptions, rereading did not significantly increase performance on the assessments. We also found that reading comprehension ability did not alter this pattern. It appears that when using ecologically valid materials such as a textbook chapter, immediate rereading may have little or no benefit for improving performance on educationally relevant summative assessments.

1. Introduction

In educational settings, one of the main avenues for learning information is reading a textbook. A frequent assertion is that we reread in order to gain a better understanding of the text (Stine-Morrow, Gagne, Morrow, & DeWall, 2004), and it is generally thought that rereading improves memory for the information stated in the text (Amlund, Kardash, & Kulhavy, 1986; Anderson, 1980; Bromage & Mayer, 1986; Durgunoglu, Mir, & Anno-Marti, 1993; Dyer, Riley, & Yekovich, 1979; Haenggi & Perfetti, 1992; Howe & Singer, 1975; Krug, Davis, & Glover, 1990; Mayer, 1983; Millis & King, 2001; Rawson & Kintsch, 2005). In the present study, massed rereading (reading twice in a row without the use of further study aids such as notetaking) was compared to a single reading to determine its effectiveness for improving learning and retention. Additionally, we used materials and assessments that are reflective of those used in the classroom. Textbook chapters were used as the texts and short-answer questions, multiple choice questions, and summaries were used to assess learning and memory.

Prior to outlining the theoretical approaches to the possible consequences of rereading, it is important to delineate the specific context in which we were most interested. Specifically, we were interested in investigating the effectiveness of reading a text twice versus reading a text once in a learning situation paralleling that faced by students in classes. In this context, readers are generally aware that they will be tested on the text's contents, when they will be tested, and what type of test they will be taking. Accordingly, in the present study participants were informed of the assessment test as well as the format of the test. Further, we used the types of assessments that students commonly encounter: multiple choice questions, short-answer questions, and a summary task. The focus of our experiments thus differs from previous research that has used free recall tasks (Amlund, Kardash, & Kulhavy, 1986; Howe & Singer, 1975; Krug, Davis, & Glover, 1990) or used cloze tasks (Rothkopf, 1968), which require the reader to recall the text verbatim, to evaluate the possible benefits of reading a text twice (relative to reading once). Our assessments, which focus more on the gist or understanding of the text, are more typical of the assessments used in the college classroom (Bol & Strage, 1996; Jackson, 2005).

2. Theoretical approaches to rereading

The question of whether reading a text twice versus reading a text a single time improves comprehension and memory for the text is a theoretically important issue that has not yet been resolved. Below, we develop the theoretical stances surrounding rereading. The issue of how rereading impacts retention also has important implications for education. Students report that one of the more common study methods they use is rereading the textbook (Carrier, 2003; Goetz & Palmer, 1991). In a survey of 170 students at a select private institution, 84% of students reported using

* Corresponding author. Fax: +1 314 935 7588.
E-mail addresses: aaduprie@artsci.wustl.edu (A.A. Callender), mmcdanie@artsci.wustl.edu (M.A. McDaniel).
[1] Fax: +1 314 935 7610.

0361-476X/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.cedpsych.2008.07.001

repeated reading as a study method, and 55% listed rereading as their number one strategy (Karpicke, Butler, & Roediger, in preparation). Rereading is also advocated as a study method by researchers (Amlund et al., 1986; Barnett & Seefeldt, 1989; Howe & Singer, 1975; Krug et al., 1990; Mayer, 1983).

Our general theoretical approach assumes that during comprehension, a reader constructs a mental representation of the text (Kintsch, 1988). This representation contains multiple levels, ranging from surface level information (the verbatim representation of the text) to a higher level representation of the text that integrates prior knowledge into the representation (the situation model). More complete processing of the text leads to a richer mental representation resulting in better understanding of and memory for the text than a representation that only contains surface level information (see Kintsch for full description). The central issue herein is whether rereading supports a richer representation (e.g., perhaps an improved situation model) than that constructed from one reading alone. An appeal to the literature suggests two distinct possibilities. The first possibility is that a second reading might create a richer, more memorable representation of the text. The second possibility is that the mental representation of the text does not significantly change with a second reading of the text. These possibilities are elaborated below.

2.1. Rereading stimulates richer representations

One position claims that the initial reading of the text is used to establish the surface and textbase levels of representation, and a second reading then fosters construction of a situation-model representation (Millis, Simon, & ten Broek, 1998; Rawson, Dunlosky, & Thiede, 2000; Stine-Morrow et al., 2004). According to Millis et al. (1998), readers have a limited pool of resources to allocate to different reading processes. The different processes require varying amounts of resources and the allocation of resources depends on the level of the process; lower level processes are allocated resources first. The initial reading involves allocating more attention to the lower level processes, allowing the reader to obtain the surface form and create the textbase. Resources are no longer needed for these lower level processes during the second reading, freeing the resources to be allocated to constructing the situation model (Millis et al., 1998; Rawson et al., 2000; Stine-Morrow et al., 2004).

A related idea, the Deactivation Hypothesis, also suggests the possibility of different processing during a second reading (Krug et al., 1990). This hypothesis states that full processing of the text occurs only when the textual representation is not present in working memory. For longer texts, like those that were used in this study, it is possible that on a second reading the text will be fully processed, and possibly processed in a different way from the first reading, because the text will have left working memory by the time the reader encounters it the second time. In support of these ideas, some studies in the educational psychology literature (Amlund et al., 1986; Anderson, 1980; Barnett & Seefeldt, 1989; Glover & Corkill, 1987; Haenggi & Perfetti, 1992; Howe & Singer, 1975; Mayer, 1983) suggest that rereading benefits retention of the contents of the text relative to a single reading.

The theoretical underpinnings and empirical findings described above allow formulation of several predictions for the classroom-inspired context in which we were interested, as outlined at the outset. The most straightforward prediction is based on the hypothesis that rereading produces a stronger or more complete representation of the information conveyed in the text than a single reading of the text. On this view, rereading should enhance performance relative to a single reading on all of the summative assessments we administered. If this prediction were obtained it would support students' (and some educators') intuitions: rereading improves learning (see Karpicke et al., in preparation).

A somewhat different set of predictions is based on the view that a second reading primarily produces a richer representation of the situation model than a single reading (Millis et al., 1998; Stine-Morrowet al., 2004). Thus, it is possible that a second reading would improve performance on criterial tests that are most sensitive to the quality of the situation-model level representation. Because multiple choice questions that focus on factual information stated in the text (such as used in the present experiments) are arguably not especially sensitive to situation-model changes (for instance, see Schmalhofer & Glavanov, 1986), this perspective would anticipate little or no benefit of rereading on the multiple choice questions that we used. In contrast to these multiple choice questions, we included short-answer questions that required description of processes, explanation of concepts or application of read content in Experiments 1, 2, and 4, and we required readers to provide a summary in Experiment 3. Assuming that the situation-model representation supports the learner's ability to describe processes (e.g., psychological or biological processes presented in textbooks in psychology or biology classes) and explanation/application of concepts (e.g., see Mayer & Gallini, 1990), then according to this view we expected that rereading would benefit performance on the short-answer questions. Finally, because the situation model is viewed as integral to summarizing the content of text (cf. Kintsch & Kozminsky, 1977), it would be expected that perhaps the most robust effects of rereading would be observed on the summary task in Experiment 3.

2.2. Rereading promotes little change to the initial representation

The second possibility mentioned earlier is that during immediate rereading few, if any, processing changes presumed to improve the representation of the text's contents are implemented. Under some circumstances (depending on the type of text and age of the reader; Stine-Morrow et al., 2004), rereading may produce little additional processing of higher level components relative to reading once. For instance, Dunlosky and Rawson (2005) found that when reading passages such as those found in the Graduate Record Exam, allocation of resources to the situation model decreased significantly for immediate rereading and showed no significant change for delayed rereading for both low and high ability college age readers.

Other evidence suggests that readers construct rigid representations that are resistant to change. Fritz, Morris, Bjork, Gelman, and Wickens (2000) required participants to process a text three times (each a week apart) and then administered a final test (during the fourth week) and found that the readers' representations changed minimally across the four weeks. Performance on 80% of the items did not change across text presentations (correct items remained correct and incorrect items remained incorrect). Further, only 4% of items not remembered on the first test were remembered on the final test. Previous research obtained similar results (Howe, 1970; Kay, 1955). In line with these results, Mannes (1994) has proposed that when the same text is read multiple times, a rigid representation of the text is created. The same inferences are made during the second reading as the first reading, and the same connections between topics are formed each time the text is read. Wiley, Griffin, and Thiede (2005) noted that readers experience greater perceptual fluency during the second reading and they mistake this fluency for understanding. Accordingly, during this second reading, the text is not processed fully and the situation model may not be improved.

Congruent with the above theoretical orientation is the material appropriate difficulty (MAD) view (Einstein, McDaniel, Owen, & Cote, 1990; McDaniel & Einstein, 2005). According to MAD, a reader will not alter his or her processing of the text unless prompted by some external factor (e.g., a study adjunct). MAD would argue

that the representation extracted and created during reading is one that is supported solely by textual elements. That is, most readers may adopt a "lazy" approach to constructing a representation, avoiding processing that is not straightforwardly afforded by the text itself (e.g., see Fletcher & Bloom, 1988). Improvements in learning and memory relative to reading alone seem to require text presentations that force generation of information not spontaneously extracted by the reader. For example, when reading an expository text readers appear to more easily extract detailed, item-specific information than relational aspects of the text (McDaniel & Einstein, 1989). Thus, relational tasks (such as sentence reordering or outlining) that force the reader to construct relations among the text elements augment the representation ordinarily constructed for expository text, and accordingly improve recall of the text (Einstein et al., 1990). By contrast an individual item task, such as completing fragmented words, that promotes processing that is ordinarily engaged when a reader encounters an expository passage (item-specific processing) generally does not improve recall of expository passages, and this occurs despite the significant increases in processing time required by the assigned task (Einstein et al., 1990; McDaniel, Einstein, Dunay, & Cobb, 1986; McDaniel, Hines, Waddill, & Einstein, 1994).

For present purposes, the core idea reflected by the various frameworks and findings just reviewed is that rereading alone does not force the reader to process the text in a manner other than that in which he or she ordinarily proceeded when first reading the text. Based on this idea, it is expected that rereading will promote little improvement on the standard learning and memory tests we investigated here.

3. Overview of the experiments

To test the competing predictions developed above, we conducted four experiments that contrasted the mnemonic effects of immediate rereading of a text (without using additional study aids) relative to reading a text once. We used texts that students commonly encounter in educational settings, specifically, extended portions of textbook chapters. It is important to note that these texts were longer than texts typically used in past rereading studies (approximately 2000 words as opposed to 300–500). We tested learning and retention with summative assessments commonly used by educators (multiple choice, short answer, and summarization; as opposed to typical laboratory based tests such as free recall of the entire passage), and the participants were aware of the nature of these tests prior to reading. Following existing studies showing that massed rereading of shorter texts can be beneficial (Amlund et al., 1986; Krug et al., 1990; Rawson & Kintsch, 2005), in Experiments 1–3 we implemented massed rereading conditions (along with single reading conditions), and we administered assessment tests within a short interval after reading. Experiment 4 extended this paradigm by also using a 24 h delayed test.

4. Experiment 1

4.1. Method

4.1.1. Participants

Thirty-two undergraduate students (22 females, 10 males) between the ages of 18 and 25 from the University of New Mexico volunteered to participate in this study. Participants were recruited through the psychology subject pool and all participants received course credit for participating in this study. Participants were tested individually in the laboratory. Participants were assigned randomly to each group, with 16 participants in each study group (16 in the read once group and 16 in the read twice group).

4.1.2. Materials and procedure

An adaptation of a portion of a chapter (2663 words) from the Crooks and Stein (1988) Introduction to Psychology text was used in this study. Participants read the text either once or twice (massed rereading such that the participants read the text the second time immediately after completing the first reading), and then were administered a summative assessment that included 22 multiple choice questions and four short-answer questions.

Participants were instructed to read through the text either once or twice (read twice group) at their normal reading speed. They were informed at that time that they would be asked to answer questions about the reading when they finished. Participants were also instructed that they would be timed as they read the text, however, there was no time limit imposed. Participants were allowed to take as much time as the needed to read the texts. Those who read the text twice were instructed to begin reading the text again after finishing the first reading.

After reading the text, the participants were given a 2 min distracter task that consisted of multiple choice and division problems. When the 2 min had passed, the math problems were collected, and the participants were given the summative assessments. There were two sets of questions, multiple choice and application questions.

The assessments were designed to test memory for the text as well as understanding of the material. Memory for the text was largely tested through multiple choice questions. An example of a multiple choice question is:

Hairstyles, dress, and social customs are examples of:

a. culture;
b. groupthink;
c. attractiveness;
d. social conformity.

Understanding of the text, as well as text memory, was largely tested through the short-answer questions. These questions required the reader not only to remember the concept from the reading, but to apply what they learned to a new situation, for example:

Pam just got a math test back and saw that she scored an 85.

Describe how social comparison might work when students get test scores back.

Within each experimental group, the order of question set was counterbalanced. Upon completion of the criterial tests, participants were debriefed and dismissed. Similar instructions were used in all of the reported experiments.

4.2. Results and discussion

4.2.1. Reading times

An initial analysis was performed to confirm that participants spent more time reading the text when required to reread. Those who read the text twice (M = 21:59, SD = 7:00) did spend more time reading than those who only read once (M = 15:24, SD = 3:37), $F = 11.12$, $p = .002$, $\eta^2 = .27$.

4.2.2. Summative assessments

Application questions were scored on a three point scale (0, 1, or 2 points given). A zero was assigned if the answer was incorrect/no answer was given, a 1 was assigned if the answer was partially correct and a 2 was assigned if the answer was entirely correct. Some participants included extra information in their responses that displayed better understanding of the text, and these answers were also assigned a 2. This scoring scheme was used for all of the experiments. Analyses were performed on the proportion of questions answered correctly (multiple choice) or the propor-

tion correct relative to the total possible score (application questions). Two separate 2 (1 or 2 times read) × 2 (order of criterial tests) between-subjects analyses of variance (ANOVA) were used to analyze the multiple choice questions and application questions.

For the multiple choice questions, there was no benefit of reading twice ($M = .83$, $SD = .11$) versus reading once ($M = .84$, $SD = .10$; $F < 1$). There was no effect of test order ($F < 1$), and no interaction ($F(1,28) = 1.17$, $p = .29$). Reliability analyses of the multiple choice questions indicate adequate reliability (Cronbach's $\alpha = .55$; see Field, 2005). Reliability was calculated using participants from Experiments 1–3.

For the application questions, there was a main effect of the order of the criterial tests. Performance was higher on the application questions when multiple choice questions were answered first ($M = .75$, $SD = .14$) than when the application questions were answered first ($M = .65$, $SD = .13$), $F(1,28) = 4.78$, $p = .04$, $\eta^2 = .15$. More importantly, reading twice ($M = .70$) was no more effective than reading once ($M = .69$; $F < 1$), and this pattern held regardless of the order in which the application test was taken ($F(1,28) = 2.83$, $p = .10$). Given the marginal significance of the interaction, we computed simple effects analyses to confirm that there was no benefit of reading twice in either of the order conditions (largest $F(1,30) = 1.51$, $p = .23$).

These results indicated that there was no benefit of reading a text twice over reading the text once, at least for tests commonly used in classroom settings. To establish the stability of these patterns, we attempted to replicate and extend the results to different participants and additional materials. Experiment 2 was conducted at a more selective institution. Perhaps top performing students have developed the necessary skills to extract more information from additional readings, and therefore will show significant benefits on the criterial tests from reading twice. Experiment 2 also added a second text, a chapter from an Industrial/Organizational (IO) Psychology textbook. Because IO is not taught in the institutions' psychology department (unlike social psychology), it was thought that participants would have less prior knowledge for the concepts and topics discussed in the text (Experiment 3 provides support for this assertion). Maybe rereading is valuable when the content is relatively unfamiliar.

5. Experiment 2

5.1. Method

5.1.1. Participants and design

Thirty-two undergraduates from Washington University in St. Louis participated in the study and received $10 for their participation. Fifteen participants were male and 17 were female. A between-subjects design (read once versus read twice) was used. To maximize the amount of data collected during each experimental session, all participants read both texts. Each participant read one text once and the other text twice.

5.1.2. Materials and procedure

Two texts were used in this study, the same Social Psychology text that was used in Experiment 1 and a chapter on "Organization Development" (IO) (3065 words) from *Applied Industrial/Organizational Psychology* (Aamodt, 2004).

The criterial tests used in Experiment 1 for the social psychology text were used again. A similar pair of tests with 22 multiple choice questions and four short-answer questions was constructed for the I/O text. Scoring of the short-answer questions was the same for both texts, with two points possible for each short-answer question.

All participants read through both the Social Psychology and IO texts. Each participant read one text once and one text twice. Both order of text presentation and number of times the texts were read were counterbalanced, resulting in four counterbalancing conditions with 8 participants assigned to each condition. Based on Experiment 1 findings, criterial tests were given such that the application questions were answered first followed by multiple choice. Additionally, the counterbalancing described above was used to prevent differences due to differential rapport with the experimenter, equalize time spent in the experimental session, and maintain similar levels of motivation.

Students were brought into the laboratory and given the first text to read. They were told to read through the text the assigned number of times, and to read it at a normal pace. Participants were timed as they read the texts. They then answered the criterial questions immediately after reading the text. Upon completing the test, they were given the second text to read, and the test was given immediately after finishing the reading. All participants were aware of the immediate tests.

5.2. Results and discussion

5.2.1. Reading times

Several participants did not time the readings correctly, and they were not included in the analysis. When participants read the IO text, they spent less time when reading once ($M = 11{:}42$, $SD = 5{:}28$) than when reading twice ($M = 20{:}02$, $SD = 4{:}26$), $F(28) = 41.29$, $p = .001$, $\eta^2 = .60$. Similar results were seen for the social text. Participants spent less time when reading once ($M = 12{:}58$, $SD = 2{:}07$) than when reading twice ($M = 18{:}41$, $SD = 4{:}39$), $F(28) = 17.89$, $p = .001$, $\eta^2 = .39$.

5.2.2. Summative assessments

Analyses were performed on the proportion of questions answered correctly (multiple choice) or the proportions correct relative to the total possible score (application questions). An α level of .05 was maintained for all analyses. Initial analyses of text counterbalancing order were performed on each text separately (as required by the experimental design) with ANOVAs performed on each of the summative assessments for each text. There were no main or interaction effects of counterbalancing order (largest $F = 2.25$). Accordingly, the analyses reported below focus on the experimental manipulation of rereading.

Two one-factor between-subjects ANOVAs were conducted on each text separately, one ANOVA for each summative assessment (multiple choice or short answer). As in Experiment 1, for the social text there were no effects of rereading for either type of assessment (both Fs < 1; see Table 1 for means).

For the IO text, MC questions showed a significant benefit of rereading, $F(1, 30) = 5.07$, $p = .03$, $\eta^2 = .15$ (see Table 1). For IO application questions, there was no effect of rereading the text ($F(1,30) = 1.19$, $p = .28$; see Table 1). Reliability was calculated using scores from participants in Experiments 2 and 3, Cronbach's $\alpha = .55$.

Experiment 2 replicated the results of Experiment 1 for the social psychology text, and extended this pattern of no rereading benefit to application questions for the IO text. The lack of a reread-

Table 1
Experiment 2 performance as a function of text and number of readings

	Multiple choice	Short answer
Social		
Read 1	.90 (.07)	.77 (.13)
Read 2	.91 (.06)	.83 (.14)
IO		
Read 1	.84 (.11)	.73 (.21)
Read 2	.91 (.08)	.80 (.14)

Note. Standard deviations given in parentheses.

ing effect in all but one of the assessments used in Experiments 1 and 2 disfavors those views in the text reprocessing literature, as well as the assumption held by many students (Carrier, 2003; Krug et al., 1990), that rereading text should improve comprehension and performance on summative assessments.

Still, rereading provided a small, but significant, benefit over reading once when answering multiple choice questions covering content in the IO text. One possibility for this inconsistent pattern is that social psychology may be somewhat familiar to college students (see Experiment 3 for supporting evidence), either because of high school courses or because of popularization of social psychology constructs. Such prior familiarity may allow readers to gain as much as they can from a first reading, thereby reducing the need for a second reading. By contrast, unfamiliar texts, like the IO text used in the present experiment, may allow a reader to gain more information during a second reading, resulting in the benefit of rereading observed (for the multiple choice test). As yet there is no evidence that content familiarity varied across the texts, and Experiment 3 was designed in part to confirm this assumption.

6. Experiment 3

We conducted this experiment to establish whether the rereading benefit obtained in Experiment 2 for the IO text (with the multiple choice assessment) was reliable, and to explore the possibility that less familiarity with the target content might modulate effects of rereading. With relatively unfamiliar content, readers presumably have a more difficult time extracting a coherent representation of the text; reading twice may allow construction of a more complete representation of the text, leading to better performance on the summative assessment.

We used the same social psychology and IO texts used in Experiment 2. Additionally, we used one of the same texts used by Rawson and Kintsch (2005) because they found that massed rereading of that *Scientific American* article improved performance on immediate comprehension and free recall tests (relative to reading once). We thought it possible that variation in content familiarity could reconcile both the somewhat divergent findings across texts in Experiment 2 and the divergence reflected in the rereading benefits reported by Rawson and Kintsch versus the absence of such benefits in Experiment 1. Accordingly, in the present experiment for each studied text we administered a survey to assess readers' familiarity with terms and concepts in each text, background knowledge for material, and perceived difficulty.

Another possibility is that rereading primarily benefits free recall. The idea here is that a second reading promotes a more coherent structural (e.g., situational) representation (Millis et al., 1998; Stine-Morrow et al., 2004) necessary for free recall (cf. Einstein, McDaniel, Bowers, & Stevens, 1984; Einstein et al., 1990) but not for tests relying on recognition (multiple choice) and cued recall (short answer; which facilitates access of the text representation). Though free recall of textbook chapters is not used in classroom summative assessments (the central interest here), writing a summary of a text requires free recall and is an assessment potentially adopted in some educational settings. Thus, in the experiment we required participants to write a summary of the text prior to a multiple choice assessment. If the possibility outlined here has merit, then rereading effect should be prominent for summaries of all of the texts.

6.1. Method

6.1.1. Participants and design

Thirty-six undergraduates ages 18–25 at Washington University in St. Louis participated in this study for course credit or $20. Twenty participants were female and 16 were male. The design included a between-subjects factor of the number of times the text was read (once or twice), and all participants read all of the texts. The order in which the texts were read was counterbalanced.

6.1.2. Materials and procedure

The same two texts (social psychology and IO) used in Experiment 2 were used in this study. A third text, an article on Greenhouse Gases (1770 words) adapted from a Scientific American article (see Rawson & Kintsch, 2005) was also used in this study. A set of 22 multiple choice questions was developed for this text, similar to that developed for the other texts. Additionally, a six question survey was constructed for participants to make judgments about each of the texts. Questions addressed the following factors: familiarity with terms, familiarity with concepts, difficulty of the text, how much they enjoyed reading the text, prior knowledge, and how much they learned from the text. Each question was answered on a five point Likert Scale, with five indicating a high level of familiarity, difficulty, etc., and 1 indicating a low level.

Participants came into the laboratory, and we obtained informed consent. Participants were then instructed that they would be reading three texts and taking a test after reading each text. They were also informed of the number of times (once or twice) they would read each text as well as the nature of the tests. They were also told that they would answer a short survey after reading each text.

Participants were given the first text to read as well as a timer to time their reading. If they read once, they were instructed to start the timer when they began and to stop it when they finished the text. If they read twice, they were instructed to stop the timer after reading through the second time. They then returned the text and timer to the experimenter. After a short distracter task of math problems, participants were given the criterial tests and performed the summary task first, and the multiple choice test second. Participants were given a blank piece of paper for the summary task and were instructed to write a summary of the text they just read. Lastly, they answered the survey for the text. They then returned the materials to the experimenter and repeated the same procedure for the two remaining texts.

The summaries were scored by assessing the number of idea units in each passage. Because it was a summary task and not free recall, a large grain size was used in scoring, resulting in 61 idea units for the social text, 60 idea units for the IO text, and 54 idea units for the Greenhouse text. To illustrate the large grain size, examples from each of the texts are listed. From the social psychology text, which focused partly on group membership, idea units included: roles are different expected behavioral patterns connected with a particular group, roles allow us to anticipate the behavior of others, and the need to affiliate appears to be a basic human characteristic. The IO text, which discussed change within an organization, included idea units such as: fear can be reduced by describing benefits of change, employees are most responsive to change when kept informed, and change should occur in a timely fashion. Lastly, the Greenhouse text, which discussed carbon sequestration, included the following idea units: carbon dioxide can be injected into the ocean or underground, Sleipner is an offshore oil and natural gas field, and nature has stored carbon dioxide underground for years.

Summaries were also scored to evaluate the macrostructure of the responses. Hierarchical structures of the original texts were constructed by putting important concepts in nodes and connecting the nodes with appropriate links. Both content (nodes) and structure (links between the nodes) were given points in the scoring scheme. Links and nodes higher up in the structure were weighted more heavily than links and nodes in lower levels of the structure. All the texts had three levels in the hierarchy of the macrostructure. Information at the highest level was weighted by multiplying the score by three, the next level was multiplied by two, and the lowest level was given a weight of 1. A similar struc-

ture was formed out of each participant's response to each text and scored according to the scheme outlined above.

All summaries scored with the idea unit scoring scheme were scored by two independent raters. Interrater reliability was high (correlation coefficient was between .8 and .9), and discrepancies in scoring were resolved by discussion. Half of the summaries scored with macrostructure scoring were scored by two independent raters. Because interrater reliability was high ($r = .96$), the remaining summaries were scored by one rater.

6.2. Results

6.2.1. Reading times

Reading times were not collected for three participants, so these analyses were conducted with 33 participants. For all three texts, participants who read twice spent more time with the text than those who read once. The analysis of the social text revealed that participants spent 10 min 19 s with the text when reading it once, and 16 min 4 s when reading it twice, $F(1, 31) = 25.14$, $p = .0001$, $\eta^2 = .45$. For the IO text, participants spent 10 min 56 s when reading once and 17 min 5 s when reading twice, $F(1,31) = 28.21$, $p = .0001$, $\eta^2 = .48$. For the Greenhouse text participants spent 8 min 6 s reading the text once and 12 min 33 s reading the text twice, $F(1,31) = 31.59$, $p = .0001$, $\eta^2 = .51$.

6.2.2. Survey

We initially examined the results of the survey to confirm our intuitions that the IO and Greenhouse texts were less familiar to the participants than the social text. The analyses of the survey included an ANOVA for each question (with text as the independent variable), and planned comparisons between each text for each of the ANOVAs, using Tukey's HSD method for correcting for multiple comparisons. All analyses were two-tailed and performed at a .05 α level. Data for one participant were incomplete, as the participant did not fill out the survey for the Greenhouse text. Therefore, data for that individual were not used in the analyses.

Results from the survey (Table 2) showed that participants perceived the texts to be different from each other, as all ANOVAs but one (how much was learned from the text) were significant at the .05 level. Three questions addressed the participants' familiarity with the content of the text (familiarity with terms, familiarity with concepts, and prior knowledge). Comparisons for these three questions revealed that participants rated their knowledge as relatively high for the social text and lower for the IO and Greenhouse texts. The social text was also rated as more enjoyable than the other texts. The other significant ANOVA and comparisons indicated that participants perceived the Greenhouse text to be most difficult of the three texts.

Table 2
Experiment 3 survey results by text and question

Question	F and p values	Social	IO	GG
1. Familiarity of terms*	$F = 11.37$, $p = .001$	3.92 (.95)	2.89 (1.38)	2.86 (.76)
2. Familiarity of concepts*	$F = 11.98$, $p = .001$	3.86 (1.03)	2.91 (1.40)	2.63 (.77)
3. Difficulty of text**	$F = 4.19$, $p = .02$	2.49 (.79)	2.34 (.93)	2.91 (.85)
4. Enjoyment of text*	$F = 4.74$, $p = .01$	3.66 (.94)	3.06 (1.06)	3.00 (.97)
5. Prior knowledge*	$F = 15.89$, $p = .001$	3.34 (.75)	1.78 (.79)	2.20 (.16)
6. Amount learned	$F = 2.44$, $p = .09$	3.31 (.14)	3.60 (.14)	3.77 (.14)

Note. Standard deviations in parentheses.

* Social text rated significantly higher than IO and GG. No other significant comparisons.

** Social and IO texts rated significantly lower than GG. No other significant comparisons.

Table 3
Experiment 3 performance by text and number of readings

	MC	Summary	
		Idea units	Macrostructure
Social			
Read 1	.89 (.09)	.13 (.07)	.19 (.11)
Read 2	.93 (.07)	.17 (.05)	.22 (.07)
IO			
Read 1	.82 (.11)	.16 (.17)	.32 (.16)
Read 2	.83 (.10)	.20 (.17)	.44 (.18)
GG			
Read 1	.60 (.11)	.20 (.10)	.25 (.12)
Read 2	.64 (.15)	.20 (.07)	.25 (.08)

Note. Standard deviations given in parentheses.

6.2.3. Summative assessments

Table 3 displays the mean performances for the summative assessments. The results of each assessment were analyzed separately by text. An initial set of ANOVAs was conducted on each assessment to assess the effects of counterbalancing, and no significant effects were found.

For the social psychology text, performance on the multiple choice questions was not significantly affected by the number of times the text was read, ($F(1,34) = 2.50$, $p = .12$). Likewise, there was no effect of rereading on the summaries, when scored either by idea units ($F(1,34) = 2.64$, $p = .11$) or by the macrostructure ($F < 1$). For the IO text, there was no effect of rereading on the multiple choice questions ($F < 1$). For the summaries, when scored by number of idea units included, there was a marginally significant effect of the number of times the text was read, $F(1,34) = 2.93$, $p = .10$, $\eta^2 = .08$. When scored by the macrostructure, the effect was slightly larger and statistically significant, $F(1,34) = 4.65$, $p = .04$, $\eta^2 = .12$. Results for the Greenhouse text showed no effect of the number of times the text was read for either the multiple choice questions or the summaries when scored with either scheme (all Fs < 1).

6.3. Discussion

The results reinforced the patterns obtained in Experiments 1 and 2: rereading a text generally did not significantly improve performance on educationally relevant summative assessments. Two issues raised in the introduction require further discussion. First, it was suggested that the results of Experiment 2 (a benefit of rereading for MC questions when reading the IO text, but not for the social text) could be due to familiarity with the content. If this were the case, we should have observed a benefit of rereading for both the Greenhouse text and the IO text in this experiment, as both were rated significantly less familiar than the social text. This was not the case, and in fact, in contrast to Experiment 2, there was not a significant effect of rereading in this experiment when the IO text was tested with multiple choice. Thus, the rereading effect found in Experiment 2 was not robust, and neither apparently is the Rawson and Kintsch (2005) result.

Second, we speculated that rereading effects may be observed in free recall type tasks (as reported in Rawson & Kintsch (2005)), but not in multiple choice questions. With the idea unit scoring, there was not an effect of rereading when participants were asked to write summaries of the texts. There was one significant effect of rereading for the IO text when the summaries were scored by macrostructure. This scoring scheme was biased toward finding a rereading effect, as the main ideas of the text were assigned three times the importance than the smaller details. Thus,

if rereading does improve the situation model (Millis et al., 1998; Rawson et al., 2000; Stine-Morrow et al., 2004), it should have been found with this scoring method. That it was found with only one text indicates that rereading benefits, when found, are limited to certain texts. Moreover, for ecologically valid measures, massed rereading immediately prior to testing appears not to be a beneficial study method.

One additional point that should be noted is that there were differences in the structures of the texts that could contribute to the significant rereading effect that was found. The IO text contained many headings and subheadings, and most of the nodes in the macrostructure were the identical to the subheadings in the text. In contrast, the Greenhouse text had few subheadings, and the social text had none. Thus, it is possible that rereading allowed the readers to acquire the explicitly stated structure of the IO text, but not the less marked structure of the other two texts.

7. Experiment 4

Two factors not yet considered might be important in producing benefits of rereading, that of reading ability and the delay between rereading and testing. Prior work has shown that rereading improved performance on questions requiring a situation model for high ability readers but not lower ability readers (Barnett & Seefeldt, 1989; Haenggi & Perfetti, 1992; Millis & King, 2001).

Barnett and Seefeldt (1989) found that repetition improved the situation model of high ability readers but only improved the surface and textbase representations of low ability readers (determined by a problem solving task). Two additional studies of relevance investigated rereading by comparing a test after the second reading against a test given after the first reading (Haenggi & Perfetti, 1992; Millis & King, 2001). While testing effects confound those results (see Roediger & Karpicke, 2006), rereading was associated with improved performance in both cases. Importantly, both studies found interactions between reading ability and rereading. High ability readers improved performance on questions that required a coherent situation model when rereading a text (Haenggi & Perfetti, 1992), and show a greater increase from the first recall test to the second test than for lower ability readers (Millis & King, 2001).

These findings raise the possibility that rereading effects could depend on reading ability. If so, then perhaps the absence of benefits (or minimal at best) found in the previous three experiments are a consequence of averaging across comprehension ability. In this experiment, we used two different reading comprehension measures to assess comprehension ability. The first measure, the Nelson-Denny Reading Test (Brown, Fishco, & Hanna, 1993) has been used widely by researchers. Uncertainty exists whether this test measures inferential processing and comprehension (Magliano & Millis, 2003), or if it primarily measures reading speed and efficiency (Carver, 1992) rather than situation-model construction. Accordingly, we used a second measure, the Multimedia Comprehension Battery (MMCB; Gernsbacher, Varner, & Faust, 1990) that assesses structure building ability (situation-model construction). Based on the view that rereading improves the situation model (Stine-Morrow et al., 2004), one would expect that comprehension ability as measured by the MMCB would interact with the number of times the text was read to predict performance on summative assessments. High ability readers (as measured by the MMCB) would benefit from a second reading, as they would have a good text-based representation from the first reading, allowing focus toward building the situation model on the second reading. Low ability readers, however, would not benefit from the second reading as they might not be able to construct a good structure out of the meaning of the text even with a second reading.

An additional factor considered in Experiment 4 was the delay of the assessment. Rawson and Kintsch (2004) found an interaction between spacing of reading and the delay of test on both short answer and recall performance. Massed rereading produced better performance on an immediate test, and spaced rereading produced better performance on a delayed test. We thought it possible that we had not obtained the optimal match between spacing of reading and time of test. While the rereading was immediate, it is possible that because the texts were longer (thereby making the reading times longer), the time between the first encounter with a specific piece of information and the second encounter could, in fact, be a spaced presentation (K.A. Rawson, personal communication, May, 2006). This could account for the differences in results obtained in previous research and our Experiments 1–3. Accordingly, in addition to an immediate test condition, we included a delayed test condition in which testing was 24 h after the readings took place.

A final feature of this experiment was that we sampled two new textbook chapters. These two chapters, one on action potentials and one on personality disorders, were taken from upper level psychology textbooks. We used these texts to preclude participants from having prior knowledge of the content, as well as to provide more generality to the results.

7.1. Method

7.1.1. Participants and design

Forty-eight undergraduates at Washington University in St. Louis participated in this study. Approximately 65% of the participants were female. All participants received either course credit or $30 for participating. None of the participants had taken a biopsychology or abnormal psychology course. (A questionnaire was administered to participants, and participants who were familiar with the topics were replaced). The design was similar to that of Experiment 2, in which the number of times a text was read was a between-subjects factor. To maximize data collection, each participant read both texts either once or twice, and the order of texts (and thereby the delay condition) was counterbalanced. The first text read was tested immediately, and the second text read was tested after a 24 h delay.

7.1.2. Materials and procedure

Two texts were used in this study: a biopsychology text (Kalat, 1998; 3167 words) on action potentials and an abnormal psychology text (Nolen-Hoeksen, 1998; 3226 words) on antisocial and borderline personality disorders. Both texts were taken from upper level psychology textbooks that the participants had not encountered in a course.

Two summative assessments were used for each passage: a 16 question multiple choice test and an eight question short-answer test. Additionally, two different reading measures were administered, the Nelson-Denny Reading Test (Brown et al., 1993) and the Multimedia Comprehension Battery (MMCB) (Gernsbacher et al., 1990). We used a computerized version of the reading portion of the MMCB which correlates well with the entire test (Gernsbacher et al., 1990).

On Day 1, participants were instructed that they would read through two texts either once or twice. They were also informed about the type and timing of the assessments that would be administered for each text. Participants read the first text the number of times that they were assigned to read it, and immediately after reading the text, participants took the assessments for that text. They then read through the second text the same number of times they read through text 1. After completing the second text, they were dismissed. Participants were timed as they read both of the texts.

Participants returned 24 h later to complete the second half of the experiment. On Day 2, participants completed both reading measures. The Nelson-Denny Reading Test was administered first, and all students were given 20 min to complete it. They then took the MMCB. Lastly, they took the criterial test for the second text that they read on Day 1. Participants were then debriefed and dismissed.

7.2. Results and discussion

7.2.1. Reading times

For both texts, those who read twice ($M_{biopsych}$ = 25:39, SD = 6:28; $M_{abnormal}$ = 27:02, SD = 7:31) spent significantly more time reading than those who read once ($M_{biopsych}$ = 17:03, SD = 4:04; $M_{abnormal}$ = 17:45, SD = 4:39), $F_{biopsych}(46) = 30.38$, $p = .001$, $\eta^2 = .40$, and $F_{abnormal}(46) = 25.37$, $p = .001$, $\eta^2 = .38$.

7.2.2. Summative assessments

Analyses were performed on the proportion of questions answered correctly (multiple choice) or the proportion correct relative to the total possible score (application questions). An α level of .05 was maintained for all analyses. An initial analysis was performed to examine the effects of counterbalancing. Two separate ANOVAs were conducted, one on each of the assessments (multiple choice or short answer), and there were no order effects (all $Fs < 1$).

Two sets of analyses were performed on the data (see Table 4 for means). First we examined the rereading effect, conducting similar analyses to those in Experiments 1 and 2. For each text, separate 2 (number of times read) × 2 (delay condition) between-subjects ANOVAs were conducted on each of the summative assessments. A set of regression analyses were also conducted to examine the influence of individual differences in comprehension ability and their possible interaction with rereading. These two sets of analyses will be discussed in turn.

For the biopsychology text the multiple choice questions showed a significant effect of delay such that performance was better on the immediate test than on the delayed test, $F(1,44) = 10.34$, $p = .002$, $\eta^2 = .19$. A marginally significant effect of the number of times the text was read was also found, $F(1,44) = 3.30$, $p = .08$, $\eta^2 = .07$. However, inspection of Table 4 shows that performance was better after reading once than twice. There was no significant interaction ($F < 1$). The short-answer questions also showed higher levels of performance for immediate testing than delayed testing, $F(1,44) = 12.40$, $p = .001$, $\eta^2 = .22$. There was no significant effect of number of times the text was read or a significant interaction (both $Fs < 1$).

For the abnormal psychology text the multiple choice questions showed a significant effect of delay with better performance after an immediate test than a delayed test, $F(1,40) = 4.13$, $p = .05$, $\eta^2 = .09$. There was no significant effect of the number of times the text was read and no significant interaction (both $Fs < 1$). Application questions showed a similar pattern, with a significant effect of delay, $F(1,40) = 16.40$, $p = .001$, $\eta^2 = .27$, and no effect of number of times the text was read and no interaction (both $Fs < 1$; see Table 4 for means).

The effect of comprehension ability on the benefits of rereading was analyzed using two separate regression analyses. Both analyses included delay condition, number of times the text was read, reading ability, and an interaction term of reading ability by the number of times the text was read as predictors. One analysis used the NDRT score as the reading ability predictor, and the other analysis used the MMCB score as the reading ability predictor. The analyses including the NDRT scores will be discussed first, followed by the analyses with MMCB scores. Because our focus here is on individual differences in comprehension ability and on possible interactions of rereading with individual differences we only report these effects. (Note, the results associated with the rereading and the delay predictors are redundant with those reported in the ANOVA's, and we only report these when the results diverge from the ANOVA.)

7.2.3. Individual differences in NDRT

For the biopsychology text the interaction term was not significant for either multiple choice or short-answer questions ($t(43) = -1.16$, $p = .25$ and $t(43) = 1.00$, $p = .32$, respectively). Reading ability significantly predicted performance on multiple choice and short-answer questions, $\beta = .26$, $t(43) = 2.04$, $p = .05$, $\beta = .27$, $t(43) = 2.14$, $p = .04$, respectively. As would be expected higher ability readers learned and retained more than lower ability readers.

For the abnormal text, the interaction was not significant for multiple choice questions ($t < 1$). For short-answer questions the interaction of reading score and the number of time the text was read was significant, $\beta = -3.50$, $t(43) = -2.03$, $p = .05$. Also, the number of times the text was read was a significant predictor, $\beta = 3.60$, $t(43) = 2.10$, $p = .04$. As Fig. 1 shows, participants who read twice performed slightly better than those who read once, with the rereading benefit being more pronounced for low ability than higher ability readers. Finally, as reading ability increased, performance on the multiple choice and short-answer questions improved, $\beta = .58$, $t(43) = 5.12$, $p = .001$ and $\beta = .74$, $t(43) = 4.03$, $p = .00$, respectively.

7.2.4. Individual differences in MMCB

For the biopsychology text the interaction term was not significant for either the multiple choice or short-answer questions, $t(43) = 1.15$, $p = .26$ and $t(43) = 1.80$, $p = .08$, respectively. Reading ability significantly predicted performance such that higher reading scores were associated with higher performance on both the multiple choice and short-answer questions, $\beta = .37$, $t(43) = 3.04$ and $\beta = .40$, $t(43) = 3.24$, $p = .002$, respectively.

For the abnormal psychology text, the interaction term was not significant for either the multiple choice or short-answer questions, (both $ts < 1$). The effect of reading score was significant for both multiple choice and short-answer question such that better performance on the MMCB was associated with better performance on the summative assessments, $\beta = .29$, $t(43) = 2.06$, $p = .05$ and $\beta = .27$, $t(43) = 2.15$, $p = .04$, respectively.

Using two new texts, we see again that rereading does not appear to be a potent study strategy on either an immediate or a delayed test. Performance on the delayed test was worse than performance on the immediate test, and this did not interact with number of times the text was read. Further, out of all of the regression models conducted, only one resulted in a significant interaction of reading ability and the number of times the text was read. Additionally, the interaction observed was contrary to that reported in previous research: low ability readers improved more with a second reading than did high ability readers.

Table 4
Experiment 4 performance as a function of text, number of readings, test, and delay

	Multiple choice		Short answer	
	Immediate	Delay	Immediate	Delay
Biopsychology				
Read 1	.58 (.13)	.43 (.14)	.38 (.16)	.23 (.04)
Read 2	.48 (.12)	.39 (.13)	.38 (.12)	.27 (.15)
Abnormal				
Read 1	.78 (.12)	.67 (.19)	.37 (.13)	.23 (.14)
Read 2	.79 (.17)	.70 (.20)	.39 (.08)	.27 (.09)

Note. Standard deviations given in parentheses.

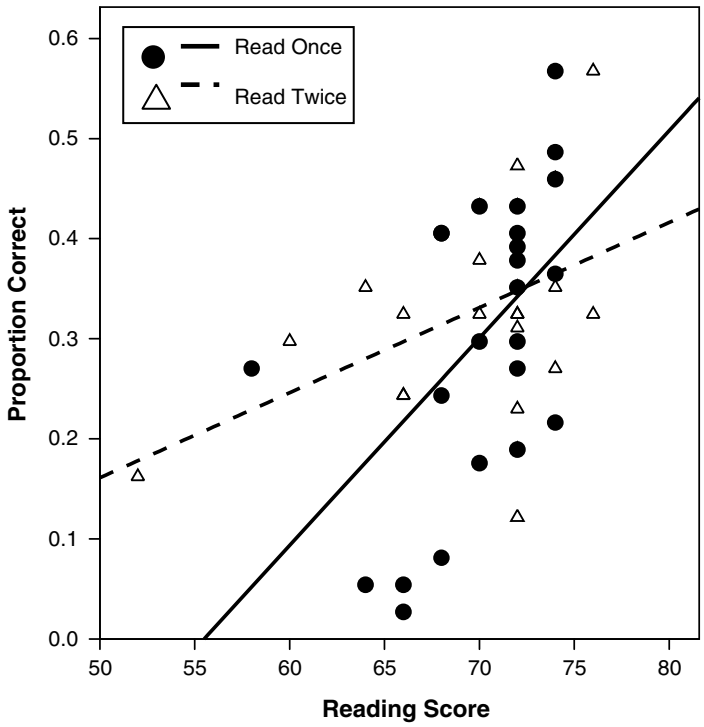

Fig. 1. Experiment 3, abnormal text short-answer performance as a function of Nelson-Denny Reading Score and reading condition.

8. General discussion

This series of experiments clearly shows that when using educationally relevant materials and summative assessments, immediate rereading is not a potent study method. There was a consistent absence of effects of rereading across all four experiments. There were no benefits of rereading an introductory social psychology chapter. There were also no effects for two chapters from upper level psychology textbooks, the biopsychology, and abnormal psychology texts. Even for a text which previously showed rereading effects, the Greenhouse gas article, again, there was no benefit of rereading. This consistent absence of effects of rereading collectively shows that immediate rereading does not reliably improve performance on summative assessments typically used in the classroom.

Two separate issues, motivation of the readers and power, could have affected how participants performed or how the results should be interpreted. First, we acknowledge that a reader's motivation in a laboratory setting is different from his or her motivation in a classroom setting. However, even with higher levels of motivation presumably present in a classroom, it is not clear that the readers would necessarily be able to use a second reading to augment their mental representations of the text sufficiently to improve performance on standard classroom assessments. To do so, the reader would likely have to possess a high level of metacomprehension ability, which many readers do not have (Maki, Jonas, & Kallod, 1994).

The second issue is whether there was adequate power to detect a difference in these experiments. Power analyses indicated that the power in each experiment by itself was not very high (power for Experiments 1 and 2 was .28 for medium effect sizes and .59 for large effect sizes, power for Experiment 3 was .31 for medium and .65 for large effect sizes, and power for Experiment 4 was .40 and .77 for medium and large effect sizes, respectively). To address this concern, we collapsed across the experiments that used identical texts and recomputed the analyses on each text and summative assessment. Performances on the multiple choice questions for the social text were collapsed across Experiments 1–3, which resulted in 50 participants per group. The means were virtually identical (M_{once} = .88, SD = .09; M_{twice} = .89, SD = .09), and the effect of rereading was not significant ($F < 1$). For the application questions for the social psychology text, we were able to collapse across Experiments 1 and 2 (this summative assessment was changed in Experiment 3), yielding 32 participants per group. Again, differences between the means were minimal (M_{once} = .74, SD = .14; M_{twice} = .76, SD = .16), and the impact of rereading was not significant ($F < 1$). We collapsed across Experiments 2 and 3 to examine performance on the multiple choice questions for the IO text, resulting in 34 participants per group. Similar to the social text analyses, the difference between the means was small (and in the opposite direction than one favoring rereading) and non-significant (M_{once} = .86, SD = .11; M_{twice} = .83, SD = .10; F = 1.62, p = .21).

We also performed a power analysis (Faul, Erdfelder, Lang, & Buchner, 2007) for each of the above analyses (cf. Fernandez & Glenberg, 1985), using both a medium and a large effect size to reflect effect sizes in prior research that has been conducted. Effect sizes in the literature vary from medium (for comprehension questions, Experiment 1, Rawson & Kintsch, 2005) to large (for recall, Experiment 1, Krug et al., 1990). The power to detect a rereading effect for the multiple choice questions for the social text ranged from .70 (medium effect size) to .98 (large effect size). Power for the application questions for the social text ranged from .50 to .88, for medium and large effect sizes respectively. Finally, for the IO multiple choice questions, power ranged from .53 to .90 for medium and large effect sizes, respectively. Thus, power was high to detect large effects due to rereading and at least in one case adequate to detect a medium size effect.

Clearly, the power to detect a small effect size was not high; however, even if there were a small effect, the reliance on rereading as a potent study method would still be questionable. Other study methods such as embedded questions take less time than rereading and produce larger benefits on similar summative

assessments (Callender & McDaniel, 2007). Beyond the question of whether rereading produces an increase in performance, there is also an issue of efficiency, and whether particular study methods are a valuable use of a student's time. The implication from the present research is that rereading is not an especially effective use of a student's study time.

One difference between these experiments and previous studies that have found effects is that some of the effects were found with free recall tasks. However, tasks such as free recall are not typical of assessments used in the classroom. With common assessments like multiple choice and short-answer questions no benefit of rereading was observed, even with relatively difficult texts.

Importantly, in these experiments we also examined several factors that have previously been shown to moderate the effectiveness of rereading and comprehension in general: prior knowledge, delay of test, and individual differences in reading comprehension ability. Generally, all failed to significantly modulate or augment the effectiveness of rereading. Turning first to prior knowledge, we speculated that the degree of familiarity a reader had with the text content might influence the benefits of rereading. Prior knowledge is a theoretically interesting factor, as it has been shown to affect a number of comprehension processes and effects of text manipulations. For example, McNamara and Kintsch (1996) found that high knowledge readers (relative to target content) benefited from increasing the difficulty of text (by reducing its coherence) whereas low knowledge readers learned less when difficulty was increased. This pattern suggests that low knowledge readers may need more support when reading a text. Additionally, low knowledge readers are less able to inhibit irrelevant information, which may interfere with the situation model being constructed (McNamara & McDaniel, 2004). Here, again rereading might have been expected to overcome this limitation of low prior knowledge. Further, Haenggi and Perfetti (1992) found that when reprocessing a text, readers tend to use the second reading to pick up pieces of factual knowledge rather than elaborating on their existing schema. This perspective also might suggest that rereading would benefit a reader who has low knowledge (since there are more facts to learn), but would not be beneficial when a reader has high knowledge for the text.

Counter to these possibilities, we found no benefit of rereading on one text for which participants self-rated their prior knowledge as low (Greenhouse gas). We did find a modest rereading effect on multiple choice questions (Experiment 2) that was nonreplicable (Experiment 3) for another text on which participants rated their prior knowledge as low (Industrial–Organizational text). Clearly, these results are suggestive at this point, as we did not objectively assess prior knowledge. It could be that readers misperceived their level of familiarity with terms and concepts and their level of prior knowledge. However, because learners are more apt to overestimate than underestimate their prior knowledge levels (Koriat & Bjork, 2005), the present results do not support the premise that rereading would particularly benefit low knowledge learners.

A second possible moderating factor for benefits of rereading is the delay between the reading and the summative assessment. One extant hypothesis is that massed rereading results in improved performance on an immediate test but does not improve performance on a delayed test (Rawson & Kintsch, 2005). In Experiment 4 we confirmed that massed rereading does not affect performance on a delayed test. However, we were consistently unable to find a benefit of massed rereading on an immediate test, even when using one of the same texts used by Rawson and Kintsch. Thus, our findings appear to hold across delays at least as long as 24 h.

A third factor investigated concerning rereading was individual differences in reading ability. Reading ability was measured in two different ways, using two reading assessments which have been shown to capture different aspects of reading comprehension ability. The first reading assessment we used, the Nelson-Denny Reading Test, has been shown to measure reading efficiency, which involves both reading speed and reading accuracy (Carver, 1992). Further, this test has been shown to measure lower level reading skills such as decoding ability (McDaniel, Hines, & Guynn, 2002). The second assessment we used, the Multimedia Comprehension Battery (Gernsbacher et al., 1990) assesses a reader's structure building ability, or their ability to build a coherent mental representation out of the text (Gernsbacher et al., 1990; McDaniel et al., 2002). Thus, two different aspects of reading comprehension were assessed and used to predict performance on the summative assessments.

One prediction related to reading ability was that rereading would benefit high ability readers more than low ability readers. We found no evidence supporting this idea for the kinds of summative assessments used in classrooms. Another idea is that readers with decoding and efficiency problems might benefit from rereading. In one model we tested (short-answer questions for the abnormal text using the NDRT as a predictor) there was a significant interaction between reading ability and the number of times the text was read, such that low ability readers realized more of a benefit of rereading than the high ability readers. Barnett and Seefeldt (1989) report converging findings for this pattern in that they found improvements in performance due to rereading for both low and high ability readers when administered a free recall. However, there was not more of an advantage for low ability readers than high ability readers. The lack of a consistent pattern of rereading benefits associated with low ability readers (there were no effects for the biopsychology text in either summative assessment, and no effect for the multiple choice questions when reading the abnormal psychology text) in the analyses cautions against drawing any firm conclusions about individual differences and rereading.

Finally, concerning the initial predictions presented in Section 1, one idea was that a second reading would produce a richer representation, perhaps in terms of a more complete situation model. This prediction was not supported with the measures we used in our study. Two different types of measures were used that presumably are sensitive to the readers' situation model: application/short-answer questions (Experiments 1, 2, and 4) and a summary of the text (Experiment 3). Application questions by definition assess the situation model as they require the reader to use their own knowledge as well as knowledge from the text to answer the question. Summaries also assess the situation model, as they are an explicit statement of the reader's macrostructure of the text (see Franzke, Kintsch, Caccamise, Johnson, & Dooley, 2005). In all cases but one (macrostructure scoring of the summary for IO texts), there was equal performance across once-read and twice-read texts on tasks that presumably reflect situation-level representations. The one instance of a rereading effect suggests that text characteristics may contribute to rereading effects. The IO text, compared to the other texts sampled in this experiment, included many more headings and subheadings. Any observable improvements to the situation model as a consequence of rereading, then, may rely in part on signaling devices incorporated into the text being read.

The results are most consistent with the views outlined in Section 1 that assume that readers extract the same representation from the text both times it is read (e.g., that readers' representations are relatively rigid across readings, Mannes, 1994; the material appropriate difficulty view, McDaniel & Einstein, 1989, 2005). In the majority of analyses, we found that performance was generally the same whether the text was read once or twice. According to the MAD perspective, when a text is encountered multiple times

by a reader, it will be processed similarly each time, unless the reader is forced to process it differently. For instance, some type of task that makes reading more difficult, such as unscrambling sentences or filling in missing letters, can force the reader to mobilize processing he or she would not ordinarily engage (McDaniel et al., 2002). Immediate rereading does not appear to stimulate such additional processing. Most likely this is due to the higher level of fluency that readers experience when reading a text for a second time (Levy, 1993; Wiley et al., 2005). In fact, fluency does quite the opposite of imposing difficulty as it is often misunderstood by readers as improved understanding and comprehension of the text (Rawson et al., 2000). Not only is there no difficulty, but the task is actually perceived as easier. Thus, when read for a second time, the text may not receive qualitatively different processing, thereby leaving the representation, such as the situation model, relatively unchanged.

Also problematic is the uncertainty of whether a reader is aware of what it means to understand a text (Wiley et al., 2005). If a reader believes that building a textbase is what is necessary for understanding a text, then the reader will not use the second reading to build a coherent situation model. With textbook materials, it is possible that students believe that a textbase level of understanding is all that is necessary.

In sum, we have shown that benefits of rereading reported in the existing literature do not necessarily extend to educationally relevant materials and summative assessments. We have obtained this finding using five different texts, taken from textbooks and *Scientific American*, and at two different universities with different student populations. This finding converges with Rawson and Kintsch's (2005) finding that there are specific boundaries to the effectiveness of rereading a text. Extending their work, the present study suggests that immediate rereading is not effective when reading a textbook chapter, at least over the ranges of content that we sampled, and when summative assessments are those typically used in the classroom (multiple choice and short-answer tests). Prior to these experiments, there has been little evidence to challenge the common opinion that rereading, even immediate rereading, is an effective way to study a text (Carrier, 2003; Goetz & Palmer, 1991). The present findings caution against broadly embracing this opinion.

Acknowledgments

Results were reported in part at the 2006 meeting of the Midwestern Psychological (Chicago, IL; May, 2006) and the 2007 meeting of the American Educational Research Association (Chicago, IL; April, 2007). This research was supported by a Collaborative Activity Grant from the James S. McDonnell Foundation.

References

Aamodt, M. G. (2004). *Applied industrial/organizational psychology* (4th ed.). Belmont, CA: Thomson-Wadsworth.

Amlund, J. T., Kardash, C. A. M., & Kulhavy, R. W. (1986). Repetitive reading and recall of expository text. *Reading Research Quarterly, 21*, 49–58.

Anderson, T. H. (1980). *Study strategies and adjunct aids*. Hillsdale, NJ: Lawrence Erlbaum.

Barnett, J. E., & Seefeldt, R. W. (1989). Read something once, why read it again? Repetitive reading and recall. *Journal of Reading Behavior, 2*, 351–360.

Bol, L., & Strage, A. (1996). The contradiction between teachers' instructional goals and their assessment practices in high school biology courses. *Science Education, 80*, 145–163.

Bromage, B. K., & Mayer, R. E. (1986). Quantitative and qualitative effects of repetition on learning from technical text. *Journal of Educational Psychology, 78*(4), 271–278.

Brown, J. I., Fishco, V. V., & Hanna, G. (1993). *Nelson-Denny reading test*. Itasca, IL: Riverside Publishing.

Callender, A. A., & McDaniel, M. A. (2007). The benefits of embedded question adjuncts for low and high structure builders. *Journal of Educational Psychology, 99*, 339–348.

Carrier, L. M. (2003). College students' choice of study strategies. *Perceptual and Motor Skills, 96*, 54–56.

Carver, R. P. (1992). What do standardized tests of reading comprehension measure in terms of efficiency, accuracy, and rate? *Reading Research Quarterly, 27*, 346–359.

Crooks, R. L., & Stein, J. (1988). *Psychology: Science, Behavior and Life*. New York: Holt, Rinehart and Winston, Inc..

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.

Durgonoglu, A. Y., Mir, M., & Arino-Marti, S. (1993). Effects of repeated readings on bilingual and monolingual memory for text. *Contemporary Educational Psychology, 18*, 294–317.

Dyer, J. W., Riley, J., & Yekovich, F. R. (1979). An analysis of three study skills: Notetaking, summarizing and rereading. *Journal of Educational Research, 73*, 3–7.

Einstein, G. O., McDaniel, M. A., Bowers, C. A., & Stevens, D. T. (1984). Memory for prose: The influence of relational and proposition-specific processing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 10*, 133–143.

Einstein, G. O., McDaniel, M. A., Owen, P. D., & Cote, N. C. (1990). Encoding and recall of texts: The importance of material appropriate processing. *Journal of Memory and Language, 29*, 566–581.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

Fernandez, A., & Glenberg, A. M. (1985). Changing environmental context does not reliably affect memory. *Memory & Cognition, 13*, 333–345.

Field, A. (2005). *Discovering statistics using SPSS*. Thousand Oaks, CA: Sage.

Fletcher, C. R., & Bloom, C. P. (1988). Causal reasoning in the comprehension of simple narrative texts. *Journal of Memory and Language, 27*, 235–244.

Franzke, M., Kintsch, E., Caccamise, D., Johnson, N., & Dooley, S. (2005). Summary street: Computer support for comprehension and writing. *Journal of Educational Computing Research, 33*, 53–80.

Fritz, C. O., Morris, P. E., Bjork, R. A., Gelman, R., & Wickens, T. D. (2000). When further learning fails: Stability and change following repeated presentation of text. *British Journal of Psychology, 91*, 493–511.

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*, 430–445.

Glover, J. A., & Corkill, A. J. (1987). Influence of paraphrased repetitions on the spacing effect. *Journal of Educational Psychology, 79*, 198–199.

Goetz, E. T., & Palmer, D. J. (1991). The role of students' perceptions of study strategy and personal attributes in strategy use. *Reading Psychology, 12*, 199–217.

Haenggi, D., & Perfetti, C. A. (1992). Individual differences in reprocessing of text. *Journal of Educational Psychology, 84*, 182–192.

Howe, M. J. (1970). Repeated presentation and recall of meaningful prose. *Journal of Educational Psychology, 61*, 214–219.

Howe, M. J. A., & Singer, L. (1975). Presentation variables and students' activities in meaningful learning. *British Journal of Educational Psychology, 45*, 52–61.

Jackson, N. E. (2005). Are university students' component reading skills related to their text comprehension and academic achievement? *Learning and Individual Differences, 15*, 113–139.

Kalat, J. W. (1998). *Biological psychology* (6th ed.). Pacific Grove, CA: Brooks/Cole Publishing Co..

Kay, H. (1955). Learning and retaining verbal material. *British Journal of Psychology, 46*, 81–100.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction integration model. *Psychological Review, 95*, 163–182.

Kintsch, W., & Kozminsky, E. (1977). Summarizing stories after reading and listening. *Journal of Educational Psychology, 69*, 491–499.

Koriat, A., & Bjork, R. (2005). Illusions of competence in monitoring one's own knowledge during study. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31*, 187–194.

Krug, D., Davis, T. B., & Glover, J. A. (1990). Massed versus distributed repeated reading: A case of forgetting helping recall? *Journal of Educational Psychology, 82*, 366–371.

Levy, B. A. (1993). Fluent rereading: An implicit indicator of reading skill development. In P. Graf & M. E. J. Masson (Eds.), *Implicit memory: New directions in cognition, development, and neuropsychology* (pp. 49–73). Hillsdale, NJ: Lawrence Erlbaum Associates.

Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition and Instruction, 21*, 251–283.

Maki, R. H., Jonas, D., & Kallod, M. (1993). The relationship between comprehension and metacomprehension ability. *Psychonomic Bulletin & Review, 1*, 126–129.

Mannes, S. (1994). Strategic processing of text. *Journal of Educational Psychology, 86*, 577–588.

Mayer, R. E. (1983). Can you repeat that? Qualitative effects of repetition and advance organizers on learning from science prose. *Journal of Educational Psychology, 75*, 40–49.

Mayer, R. E., & Gallini, J. K. (1990). When is a picture worth ten thousand words? *Journal of Educational Psychology, 82*, 715–726.

McDaniel, M. A., & Einstein, G. O. (1989). Material appropriate processing: A contextualist approach to reading and studying strategies. *Educational Psychology Review, 1*, 113–145.

McDaniel, M. A., & Einstein, G. O. (2005). Material appropriate difficulty: A framework for determining when difficulty is desirable for improving learning. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 73–85). Washington, DC: American Psychological Association.

McDaniel, M. A., Einstein, G. O., Dunay, P. K., & Cobb, R. (1986). Encoding difficulty and memory: Toward a unifying theory. *Journal of Memory and Language, 25*, 645–656.

McDaniel, M. A., Hines, R. J., & Guynn, M. J. (2002). When text difficulty benefits less-skilled readers. *Journal of Memory and Language, 46*, 544–561.

McDaniel, M. A., Hines, R. J., Waddill, P. J., & Einstein, G. O. (1994). What makes folk tales unique: Content familiarity, causal structure, scripts, or superstructures? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 169–184.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes, 22*, 247–288.

McNamara, D. S., & McDaniel, M. A. (2004). Suppressing irrelevant information: Knowledge activation or inhibition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 30*, 465–482.

Millis, K. K., & King, A. (2001). Rereading strategically: The influences of comprehension ability and a prior reading on the memory for expository text. *Reading Psychology, 22*, 41–65.

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

Nolen-Hoeksen, S. (1998). *Abnormal psychology*. Boston: McGraw Hill.

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

Rawson, K. A., & Kintsch, W. (2005). Rereading effects depend on time of test. *Journal of Educational Psychology, 97*, 70–80.

Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*, 181–210.

Rothkopf, E. Z. (1968). Textual constraint as function of repeated inspection. *Journal of Educational Psychology, 59*, 20.

Schmalhofer, F., & Glavanov, D. (1986). Three components of understanding a programmer's manual: Verbatim, propositional, and situational representation. *Journal of Memory and Language, 25*, 279–294.

Stine-Morrow, E. A. L., Gagne, D. D., Morrow, D. G., & DeWall, B. H. (2004). Age differences in rereading. *Memory & Cognition, 32*, 696–710.

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