SPIRGEL, ARIE, Ph.D. Assessing Summary Writing as a Memory Strategy. (2011)
Directed by Dr. Peter Delaney. 94 pp.

Writing a summary has been described as the most effective text-retention strategy. However, a review of the literature suggests that existing experiments on summarizing fail to include some of the most productive methods for improving text-retention. The current set of six experiments was designed to overcome these deficits and to create conditions in which text-summarization is optimized. Experiments 1 and 2 investigated how distributing summarizing sessions influences text-retention, and Experiment 3 investigated the influence of difficult retrieval conditions on summarizing and text-retention. Motivated by the large increase in retention associated with including an idea unit in a summary in Experiments 1 – 3, Experiment 4 was designed to examine the special status of included idea units by comparing summarizing to underlining. Experiment 4 replicated the claim that including an idea unit in a summary is associated with an increase in retention relative to underlining or not including an idea unit. To test if including an idea unit in a summary in fact causes an increase in retention, Experiment 5 included a condition in which participants were instructed to identify and include at least 10 of the most important points in their summaries. Because participants had difficulty identifying the important elements of the text, which had been identified a priori by the researcher, it was not possible to examine the causal effects of writing on memory. A different approach was thus taken in Experiment 6, in which participants read a text containing information meant to distract them from the main ideas, creating

conditions in which it could be evaluated how writing a summary influences people's ability to identify most important elements of a text.

ASSESSING SUMMARY WRITING AS A MEMORY STRATEGY

by

Arie Spirgel

A Dissertation Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2011

Approved by

____
Committee Chair

Thank you to my family for always encouraging me to pursue my interests.

APPROVAL PAGE

This dissertation has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair ______________________________
Peter Delaney

Committee Members ______________________________
Rob Guttentag

____
Lili Sahakyan

____
Paul Silvia

____
Dayna Touron

____
Date of Acceptance by Committee

____
Date of Final Oral Examination

ACKNOWLEDGEMENTS

Thank you to my advisor, Peter Delaney; if it were not for his creativity and desire to share it, this dissertation would not have been possible. And thank you to the members of my committee – Rob Guttentag, Lili Sahakyan, Paul Silvia, and Dayna Touron – whose combined expertise helped guide my research. My research was also shaped by Matt Meier, whose continious interest and constructive criticism of the work reported here greatly improved its quality.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

Writing across the curriculum (WAC) is the result of an idea that emphasizes the process of learning rather than the product. WAC programs, which are present in 38% of colleges and universities (McLeod & Shirley, 1987), stress that writing is a vital part of learning, regardless of the subject matter (Bazerman, Little, Bethel, Chavkin, Fouquette, & Garufis, 2005). Of the five generally agreed upon principles of WAC programs, one is that "writing promotes learning" (Palmquist et al., n.d.).

WAC emphasizes that most types of writing have a desirable impact on most measures of learning, but this paper will focus on one type of writing that has received considerable attention as a way of enhancing learning – summary writing – and investigate its effects on memory. A summary, which is the condensation of a topic's main points (Brown, Day, & Jones, 1983; Hidi & Anderson, 1986; Winograd, 1984), has received extensive praise for its effect on learning and memory. Some are so convinced by the evidence that they concluded that there is no better way to promote long-term text retention than summarizing (Westby, Culatta, Lawrence, & Hall-Kenyon, 2010). Summarizing has been said to make writers behave like archaeologists because both "must dig for information, make sense of it, and attach meaning to it" (Wormeli, 2004, p. 6). Wade-Stein and Kintsch (2004) say that summarizing results in deeper understanding than rereading a text because it requires integration of new knowledge with existing knowledge. Others have argued that summarizing results in improved retention and deep comprehension because writers must consider the entire passage, and determine its important elements (Friend, 2002; Radmacher & Latosi-Sawin, 1995).

As with other forms of writing, research on summary writing has been conducted almost exclusively outside the domain of cognitive psychology, but from the perspective of the cognitive psychologist, summarizing is intuitively appealing. Writing a summary appears to be conducive to the type of spaced practice and memory retrieval that cognitive psychologists advertise should be used in schools (e.g., Pashler, Rohrer, Cepeda, & Carpenter, 2007). The spacing effect is the finding that spaced items, or item repetitions presented with intervening time or items, are better remembered than massed items, or item repetitions presented in immediate succession. The closely related testing effect is the finding that being tested on an item improves retention relative to restudying it. The spacing and testing effects are robust phenomena that have been extensively studied over the last 100 or so years (for reviews, see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Delaney, Verkoeijen, & Spirgel, 2010; Dempster, 1996; Roediger & Karpicke, 2006a). Spacing or testing effects have been obtained with materials as diverse as word lists (e.g., Delaney & Verkoeijen, 2009; Karpicke & Roediger, 2007), foreign vocabulary (e.g., Bahrick, Bahrick, Bahrick, & Bahrick, 1993; Karpicke & Roediger, 2008), math (e.g., Rohrer & Taylor, 2006), maps (e.g., Carpenter & Pashler, 2007), name learning (e.g., Carpenter & DeLosh, 2005) and history facts (e.g., Carpenter, Pashler, & Cepeda, 2009). Failures to obtain spacing and testing effects are so rare that null effects are considered theoretically informative (Delaney & Knowles, 2005; Verkoeijen, Delaney, Bouwmeester, Coppens & Spirgel, 2011). Although most theoretical advancements of the spacing and testing effect have been made with word list learning studies, research has shown that results from these experiments translate well to the types of text passages that will be used in the current experiments (e.g., Karpicke & Roediger, 2010; Krug, Davis, & Glover, 1990; Rawson & Kintsch, 2005; Storm, Bjork, & Storm, 2010; Verkoeijen, Rikers, & Ozsoy, 2008).

The Effects of Summarizing on Memory

Thus far, the effects of summarizing on retention have been presented as conjecture. The current section describes the evidence. Although writing a summary might have some effect on memory, it is more practical to know how writing a summary compares to other learning strategies. Thus, the following review examines studies evaluating the influence of writing a summary on memory and is broken up by what summary writing was compared to.

Summary Writing vs. Traditional Instruction

Several studies implemented summary writing in classroom settings and compared writing a summary to more traditional instruction. For example, Day (1994) required daily journals that prompted students to restate major concepts from that day's lecture and to relate them to experiences from their own lives. There were no improvements on four multiple choice/essay examinations given throughout the semester compared to a control group that did not write journals, but whose attendance was marked daily.

Carroll (2008) investigated the value of writing a summary in an introduction to psychology course by beginning classes with an explanation of a famous quote related to that day's lecture, and for some days the quote was not revisited (quote-only), and on other days students were given 5 min at the end of class to write a summary connecting the quote to the material from that day (quote with summary). On exam questions that were associated with quote-only lectures, quote with summary lectures, or neither, only the quote with summary questions were answered at a higher rate than the same questions from a previous semester that served as a control group.

In Radmacher and Latosi-Sawin (1995), students in a third year psychology class were given class time at least once per week to summarize a portion of the text. When summaries were due, as a class students would identify the main points from the assignment and write them on the board. Compared to a section that took the same final exam the previous semester, students in the writing class scored significantly higher. Similarly, Horton, Fronk, and Walton (1985) assigned students to summarize eight lectures in a general chemistry class. Compared to students in the same class who did not write summaries, the treatment group scored significantly higher on an exam measuring knowledge of the lectures.

Although these results are somewhat positive in favor of summary writing, the only conclusion that can be drawn from them is that writing has an effect beyond traditional instruction.

Summary Writing vs. Restudying

Restudying has been reported as students' most common study technique (Karpicke, Butler, & Roediger, 2009). It is also a relatively ineffective study technique (Callender & McDaniel, 2009), and thus is a lenient comparison group.

Penrose (1992) had participants read a text and gave them up to 1 hr to either write an informative essay of the key concepts, or to study the key concepts. After each task, participants were asked to recall as much they could from the text. Even though the writing group spent 14 min longer on the task, they recalled less of the text than the restudy group. This finding suggests that even if there are benefits to writing, it may have the deleterious effect of taking time away from more efficient study strategies – a possibility consistent with a meta-analysis reporting that longer writing assignments are associated with a reduction in learning (Bangert-Drowns, Hurley, & Wilkinson, 2004).

Davis and Hult (1997) showed participants a video on the development of language that included three seven min segments and assigned people to one of three conditions: those in the summary group took notes during the video and wrote a summary during four min pauses after each segment; those in the pause group took notes during the video and reviewed the notes during the same four min pauses; a control group took notes throughout the video without pause. On a multiple-choice test given immediately after reading, there were no differences observed among the groups. When a multiple-choice test and a free recall test were given after a 12 day delay, the summary group outperformed the combined average of the restudy and control groups, suggesting the benefits of summarizing might not emerge until after a delay.

Although the previous two studies contradict each other, there are several explanations for the discrepancy. One is that Penrose (1992) permitted students in the restudy group to study how they pleased. It may be that allowing students to choose their own learning strategy results in better retention than controlling students' restudying, as Davis and Hult (1997) did. A second explanation is that Davis and Hult analyzed their data by combining the restudy and control group, which is problematic given that the control group was not given as much review time as the other two groups. Thus, the authors may have overestimated the benefit of summary writing relative to restudying. Lastly, the different retention intervals used in the studies hints that summarizing may slow the rate of forgetting relative to restudying, implying that writing a summary behaves like a test, which also slows the rate of forgetting (Roediger & Karpicke, 2006b)

Summary Writing vs. Generating Questions

When students are assigned to generate questions, they are generally asked to read a text, create an allotted number of questions based on the text, and to answer those questions. This method has been found to produce effects on memory comparable to testing (Weinstein, McDermott, & Roediger, 2010), and thus is a stringent comparison group.

In a study by King (1992), college students enrolled in a remedial reading and study skills class were assigned to one of three conditions over the course of eight 50 min sessions: generate and answer questions about a lecture; review lecture notes; or write a summary of the lecture. After a pretest in the first session, the self-questioning and summary group received practice and training in their respective strategies over the next four sessions. The next session was devoted to viewing a lecture, implementing their assigned strategy, and taking a comprehension test about the lecture. This was followed by a final comprehension test about the lecture 1-week later. On the immediate test, the generate-question and summary groups did not differ from each other on the number of questions they answered correctly, and both groups answered more questions correctly than the note-taking with review group. On the comprehension test 1-week later, only the generate-question group answered more questions correctly than the note-taking group, with summarizers falling numerically between the two other groups but not reliably differing from either.

King, Biggs, and Lipsky (1984) also gave participants extensive training on self-questioning and summarizing and compared these techniques for remembering a text passage to a note-taking control group. Participants were given a series of memory tests 48 hrs after reading and implementing their assigned strategy. On a free recall test, only the summary group remembered more than the control group (no other comparisons were significant). On an exam-like test (i.e., true/false, multiple choice, and short answer questions), the two treatment groups did not differ, with both remembering more than the control group. Taken together, these findings imply that the effects of summarizing on memory depend on the retention interval and the type of test used to measure memory.

The strength of the evidence in favor of writing a summary as a way to retain text does not match the level of enthusiasm for it. Although summary writing has the potential to serve as a case of spaced and tested practice, the evidence reviewed does not include these conditions.

Spacing, Testing, and Writing

The following review includes experiments that attempted to merge spacing or testing with writing. Although the experiments provide indirect evidence that spacing and retrieval with writing influence retention, they include major limitations, which will be reviewed here.

Spaced Writing. Although they did not test memory and therefore their study cannot be viewed as a confirmation of the benefits of spaced writing, Benton, Kiewra, Whitfill, and Dennison's (1993) findings on note-taking and delayed essay writing produced results that are consistent with the deficient processing theory of spacing (e.g., Hintzman, 1974). Benton et al. had participants watch a video-lecture and either immediately after (Experiment 3) or after a 1-week delay (Experiment 4) write an essay about it. Some participants were given notes on the lecture and others were not. In the delayed (i.e., spaced) condition, participants given notes wrote essays that were longer and more organized than those who wrote without notes, but no such differences emerged in the immediate (i.e., massed) condition. This discrepancy led the authors to posit that participants in the delayed condition, but not those in the immediate condition, made use of the notes. Presumably, for the participants in the immediate condition the lecture never left the short-term store, which is why they did not rely on their notes. Therefore, delaying writing appeared to create conditions suitable for retrieval, which is one mechanism that is thought to drive the spacing effect (Delaney, Verkoeijen & Spirgel, 2010).

Writing With Retrieval. Research that has implemented retrieval with writing has done so in absence of variables that create a testing effect. The few studies that exist on this topic, however, provide some evidence that introducing more difficult retrieval conditions during writing improves retention.

Dyer, Riley, and Yekovich (1979) made the retrieval demands of summarizing more difficult by limiting access to the source text while writing. They had participants first read a text while taking notes on it or not, then write a summary without access to the text or complete a distractor task, and next reread the text or engage in a distractor. All participants were then given a retention test after a 30 min delay and after a 1-week delay. Text-absent summarizing should have acted as a test, and subsequent rereading as feedback, thus resulting in improved retention relative to the other groups, at least at the longer delay. But because the retention interval was a within-subjects factor, all participants received a test before the 1-week delay, making it is unsurprising that no retention benefits of summarizing were observed. Research that has investigated the outcomes of text-absent summarization more in-depth offers promising results.

Spurlin, Dansereau, O' Donnell, and Brooks (1988) manipulated the retrieval demands of writing by giving participants a text and assigning some to stop four times to summarize it, some twice, and some to implement their own study strategy. The authors hypothesized that less frequent summarizing should require greater retrieval effort, resulting in better retention of the text. Their hypothesis was supported as those who summarized twice, but not those who summarized four times, included more correctly recalled idea units on a final essay test occurring 5 days later than those who implemented their own strategy. Similarly, participants in Foos' (1995) study who summarized a text once demonstrated better retention on a fill-in-the blank and multiple-choice test than participants who summarized twice or used their own study strategy. Even though Foos obtained this finding at a short, 5 min retention interval, presumably participants successfully retrieved a high percentage of the material on the initial writing opportunity, and the testing effect has been observed at short delays when there is a high success rate of initial retrievals (Kuo & Hirshman, 1996).

Kirby and Pedwell (1991) increased the retrieval demands during summarizing by having participants summarize without access to it (i.e., text absent). A comparison group summarized with access to the text (i.e., text present), and participants from both conditions returned a few days later to recall what they could from it. Recall rates between the text-absent and text-present groups did not reliably differ, but given that text-present writers included more idea units in their summaries, final recall for text-absent summarizers may have improved had they received feedback. The authors did find that summary content (i.e., number of important details, main ideas, and themes included) was positively related to recall in the text-absent condition but not in the text-present condition. This result, which resembles what has been found with children (Stein & Kirby, 1992), suggests that what was included in text-absent summaries benefited from memory retrieval and thus subsequent recall. When the text was present, memory retrieval was largely negated and better summary content did not lead to increased recall on the final test.

Overview of Experiments

In sum, even if summarizing is as effective of a retention tool as advocates claim that it is, it has the potential to be improved upon by capitalizing on spacing and testing manipulations**;** however, existing research has not taken advantage of them. The two overarching questions in the current set of experiments were (1) is summarizing in fact as effective of a memory strategy as it is claimed to be, and (2), what are the conditions in which writing a summary optimizes retention? The general plan was to use both effective control groups (restudy) and our knowledge of spacing and testing principles to maximize the chance that summary writing would produce a memory benefit.

The specific question addressed in Experiment 1 was how distributing writing opportunities influences text retention. The primary question asked in Experiment 2 was the same, with the major addition being how this influence operates over a longer retention interval. Experiment 3 examined how increasing the retrieval difficulty of writing influences its effect on text retention. Experiment 4 was designed to directly ask the question, does writing confer a text-retention benefit beyond simply rereading? Experiment 5 was designed to test the causal effects of writing on memory. Experiment 6 evaluated the hypothesis that creating a summary helps writers pick out the important information from a text.

Table 1.
Overview of Experiments

Experiment #	**Design**	**# of Words in Text**	**Retention Interval**	**Type of Test**
1	2 (Massed vs. Spaced)[w] x 2 (Restudy vs. Write)[b]	256-258	Immediate	Short-answer
2	2 (Massed vs. Spaced)[b] x 2 (Restudy vs. Write)[b] x 2 (2 m vs. 48 hr)[b]	264	2 m or 48 hr	Short-answer
3	4 (Restudy vs. Text-Present Summary vs. Regular Test vs. Text-Absent Summary)[b]	≈2,500	7 d	Multiple-choice
4	2 (Underline vs. Summary)[b]	≈2,500	Immediate	Multiple-choice
5	4 (Regular Summary vs. Restudy vs. New Instruction Summary vs. Control)[b]	≈2,500	Immediate	Short-answer
6	2 (Restudy vs. Summary)[b]	967	Immediate	Free recall

[w] indicates that the variable was manipulated within subjects
[b] indicates that the variable was manipulated between subjects

CHAPTER II

EXPERIMENT 1

The first experiment was designed to test the hypothesis that spacing improves the effects of writing on memory relative to massing, and to compare restudying and writing as memory strategies. Participants were given a short text to read, and either immediately restudied it or immediately summarized it (massed), or they restudied it or summarized it after a 5 min delay (spaced). A short answer test was then given. Participants then repeated the same procedures, but were given a spaced presentation if the initial one was massed, or vice versa.

If the general effects of spacing on memory apply to the text passages used here, then a main effect of spacing should be observed, such that those in the spaced condition retain more than those in the massed condition. If the broad support for summary writing as a text retention strategy is warranted, then a main effect of task should be observed such that writers remember more than those who restudy. An interaction between spacing and task (i.e., write or restudy) was not expected.

Method

Participants

Eighty-eight UNCG undergraduates participated in this experiment for course credit.

Materials

I used two text passages, one about the sun (256 words) and another about Lake Okeechobee (258 words). They were obtained from the Test of English as a Foreign Language preparation book (Rogers, 2001) and Wikipedia, respectively. The sun passage was also used by Roediger and Karpicke (2006b). I created 11 short-answer questions testing people's memory for the text that accompanied each passage. A word search that contained 1-word movie titles was used as the distractor task.

Procedure

Participants began by reading one of the two text passages at their own pace. Then, participants worked on a word search for 8 min before beginning *writing* or *restudying* (*spaced*), or they immediately began *writing* or *restudying* (*massed*). When the task began, those in the restudy group were instructed to restudy the passage "with a focus on the parts you think are the most important, for example, what you think might appear on a later test." Those in the summary group were instructed to write a summary of the text and to "not simply list facts from it, but try and make the summary cohesive, like what you might read at the end of a textbook chapter. Include the parts that you think are the most important, for example what you think might appear on a later test." Participants were allowed to refer back to the text as much as they desired while writing. Both restudy participants and summary participants were given 5 min for their respective tasks. After 8 min of working on the word search, participants were given the test. After the test, those who first participated in the spaced condition first subsequently completed the massed procedures with the other text, and vice versa.

Results & Discussion

Final Test

To test the effects of spacing and task on memory, I ran a 2 Spacing (massed vs. spaced) x 2 Task (summary vs. restudy) ANOVA on the proportion of questions answered correctly. There was no main effect of Spacing, $F < 1$, nor was there a Spacing x Task interaction, $F < 1$. There was a marginal main effect of Task, $F(1,86) = 3.90$, $MSE = 7.56$, $p = .052$, $\eta^2 = .043$, such that the proportion of questions answered correctly was higher in the restudy condition ($M = .69$, $SE = .027$) than in the summary condition ($M = .61$, $SE = .027$). Figure 1 summarizes these findings.

A null spacing effect is rare with word lists, but a failure to produce one here suggests that text passages are more resistant to the effects of spacing than more traditionally used materials. This may be because even immediately rereading a text results in functional spacing, thus attenuating its benefit relative to massing. To use an extreme example, if a book was read, and then reread immediately, the amount of time that would elapse between initially reading the first sentence of the book and then rereading it would render it functionally spaced, but by the current definition would qualify it as massed. This is in stark contrast to massing in list-learning studies, when massed repetitions appear in immediate succession (e.g., *cat, cat, ball, ball*).

In contrast to the vast enthusiasm for summarizing as a memory strategy, restudying resulted in better memory than summarizing. This finding is consistent with prior research (Penrose, 1992), but conflicts with other research (e.g., Davis & Hult, 1997). Similar to Penrose (1992), who found that restudying results in better retention than summarizing, the current experiment used a short retention interval. Davis and Hult (1997) observed that summarizing led to better retention than restudying and included a more extended retention interval -- a manipulation that was implemented in Experiment 2.

Writing Analyses

For the following analyses, *idea unit* is defined as including text in a summary that could be used to answer a question on the final test. All coding for this and subsequent experiments was completed by Arie Spirgel. A paired samples *t*-test showed that massing did not result in including a different number of idea units ($M = 5.07$, $SE = .37$) than spacing ($M = 4.84$, $SE = .35$), $t < 1$. This amounts to including, respectively, 46.1% and 44.0% of the potential 11 idea units.

The number of idea units included was correlated with recall in the spaced condition, $r(44) = .42$, $p = .005$, but not in the massed condition, $r(44) = .21$, $p = .174$.

I calculated the probability of recalling an idea unit if it was included in a summary, and the probability of recalling an idea unit if it was not included in a summary and ran a 2 Inclusion (included vs. not included) x 2 Spacing (spacing vs. massing) Repeated Measures ANOVA on the probability of recalling an included vs. not included idea unit. This answers the question of whether including an idea unit in a summary increases the probability of later recalling that item, and whether that probability varies by spacing. There was a main effect of inclusion, $F(1,38) = 74.32$, $MSE = .052$, $p < .001$, $\eta^2 = .662$, such that the probability of recalling an included idea unit ($M = .79$, $SE = .030$) was higher than the probability of recalling a non-included idea unit ($M = .48$, $SE = .035$). There was no main effect of spacing, $F < 1$. There was not a Spacing x Inclusion interaction, $F(1,38) = 2.23$, $MSE = .041$, $p = .144$, $\eta^2 = .055$. These data are summarized in Figure 2.

Although the final test data suggest that spacing had no influence on memory and restudying was more effective than summarizing, the writing analyses provide a somewhat different picture. First, the number of idea units included in a summary correlated with final test accuracy in the spaced but not massed condition, which suggests spacing may have had a covert effect. When prior presentations of stimuli are retrieved (i.e. study phase retrieval), spacing should improve retention for those items relative to massing (Delaney, Verkoeijen & Spirgel, 2010). A prediction that derives from this is that as more idea units are included in a spaced summary, retention should improve because of successful study phase retrievals. No reliable relation would be expected between idea units included in massed summaries and recall. The correlations are consistent with this prediction.

On the final test, restudying resulted in better memory than summarizing, suggesting that for the purposes of text retention, summarizing is relatively futile. The inclusion data, however, show that including an idea unit in a summary is associated with an increase in the likelihood of later recalling that item, as compared to omitting an idea unit from a summary. Thus, it may not be that summarizing is futile, but that the benefits are limited to what students write about.

Item Analyses

Given that including an item in a summary was associated with better recall than not including an item, the purpose of these analyses was to test whether the effect was due to including the item, or a result of item selection effects. In other words, were items that were included in summaries items that would have been recalled regardless of whether they were included or not? To address this question, I computed Spearman's Rank Order Correlations between items that were included in summaries and items that were recalled in the restudy condition. Each passage (i.e., sun and Okeechobee) and spaced and massed conditions were analyzed separately. In the massed conditions, what summarizers included did not predict what those in the restudy condition remembered for the sun passage, $r_s(11) = .04$, $p = .912$ nor for the Okeechobee passage, $r_s(11) = .53$, $p = .090$. In the spaced condition, what summarizers included did not predict what restudiers remembered with the sun passage $r_s(11) = .40$, $p = .228$, nor in the Okeechobee passage, $r_s(11) = .58$, $p = .061$.

Given that three out of the four item analyses were $\geq$ .40 (despite being non-significant) suggests that although including something in a summary is associated with increased retention, this effect is not completely due to the causal nature of writing and is driven at least in part by item selection effects. In other words, the item analyses provide evidence that the items that participants included in their summaries were items they would have recalled even if they had not included them, although this cannot explain the entire *inclusion effect* (i.e., enhanced retention for included items).

Difference Scores Analysis

The purpose of this analysis was to further examine the effect of including an item in a summary. First, I calculated the probability of correctly answering a question if it was included in a summary (e.g., .80). I then calculated the proportion of items participants in the restudy condition answered correctly, counting only the items that each participant included in their summary. For example, if Participant 1 included Items 3, 7 and 9, in the summary, and participants in the restudy condition answered these questions at a rate of .40, .50, and .60, respectively, I obtained the average of those three numbers (i.e., .50). I then subtracted the first number (.80) from the second number (.50) to obtain a difference score. This was done for each participant.

The difference score in the massed condition (M = .00, SE = .042) did not reliably differ from zero, $t < 1$, whereas the difference score in the spaced condition did (M = .12, SE = .032), $t(42) = 3.86$, $p < .001$. This suggests that including additional items in a spaced summary benefits retention relative to restudying, but that is not the case for massed summaries.

The major drawback of this analysis is that it is unknown what participants were doing when they were asked to restudy, which is less the case with summarizers who create a product. In other words, even though the analysis provided evidence that including an item in a summary benefits retention relative to restudying, it is unknown if participants restudied when they were asked to. Although participants were monitored for off-task behavior during the experiment, it is not possible to know if a given participant appeared to be restudying when he was staring at the passage blankly without actually studying it. Additionally, it may that participants spent their time studying items that summarizers happened to not include in their summaries. This caveat applies to each of the difference score analyses conducted in the subsequent experiments.

CHAPTER III

EXPERIMENT 2

In direct conflict with the opinion that summarizing is a unique mode of learning, the first experiment found that restudying led to better recall. The failure to find a benefit of summarizing may be due to the short retention interval. The significant correlation between what was included in summaries and subsequently recalled in the spaced condition, along with the same correlation being non-significant in the massed condition, suggests that spacing may contribute to recall, despite its lack of an effect on the final test. Although prior experiments have found benefits of distributing reading on an immediate test (e.g., Glover & Corkill, 1987; Krug, Davis, Glover, 1990), other work has found that the spacing effect is dependent on the delay (Rawson & Kintsch, 2005; Verkoeijen, Rikers, & Ozsoy, 2008). Assuming the latter, the current experiment included an immediate test in addition to a delayed test.

The procedures were similar to the first experiment, except that all variables were manipulated between subjects. Participants read a text, and then restudied it or summarized it in massed or spaced fashion. A short answer test was administered after a 2 min delay, or after a 48 hr delay.

If the null effect of spacing on text memory was due to the test being too soon after exposure to the material, then a Spacing x Retention Interval interaction should be observed, such that the benefits of spacing are observed at the 48 hr test, but not at the 2 min retention interval. Similarly, if the effects of writing on memory are influenced by spacing – as the Experiment 1 finding that the correlation between included idea units and recall is significant only spaced summarizing suggests is the case – then a Task x Retention Interval interaction should be observed such that writing results in better memory than restudying only at the 48 hr test.

Method

Participants and Design

One-hundred and fifty-two UNCG undergraduates participated for course credit, but the 11 people who participated in Session 1 and did not return for Session 2 were eliminated from the analyses. The design was a Spacing (massed vs. spaced) x Task (restudy vs. summary) x Retention Interval (2 m vs. 48 hours). All variables were manipulated between-subjects.

Materials

The text used in this experiment was 264 words long and was about the earth's atmosphere. It was obtained from Wikipedia. The final test consisted of 10 short answer questions. The same word search that was used in the prior experiment served as the distractor for the current experiment.

Procedure

All participants began by reading the text at their own pace. When done, participants in the *spaced* condition completed a word search for 13 min before beginning their assigned task (i.e., *restudy* or *summary*), and participants in the *massed* condition immediately began their assigned task. Those in the summary condition were given 5 min to write a summary of the text with the same instructions as the summary participants in the previous experiment, and participants in the restudy condition were given 5 min to restudy the text with the same instructions as the restudy participants in the previous experiment. Everyone then received 2 min to continue their word search. Next those, those in the *2 min retention interval* condition were given the final test. Those in the *48 hr retention interval* condition were asked to write down what percentage of questions they thought they would remember in 2 days and they were dismissed. They completed the final test questions upon returning to the lab.

Results

Questions Correct

To test the effects of spacing, task, and retention interval on retention, I ran a 2 Spacing (spaced vs. massed) x 2 Task (restudy vs. summary) x Retention Interval (2 min vs. 48 hr) ANOVA on the proportion of final test questions answered correctly. There was no main effect of spacing, $F(1,133) = 2.68$, $MSE = .049$, $p = .104$, $\eta^2 = .020$, but a trend for those in the massed condition ($M = .51$, $SE = .026$) to recall more than those in the spaced condition ($M = .45$, $SE = .027$). There were no recall differences between the restudy condition and the summary condition, $F < 1$. There was a main effect of retention interval, $F(1,133) = 17.14$, $MSE = .049$, $p < .001$, $\eta^2 = .114$, as those in the immediate condition ($M = .56$, $SE = .025$) recalled more than those in the delayed condition ($M = .40$, $SE = .028$). All of the two way interactions were $F < 1$, and the three way interaction was $F(1, 133) = 2.44$, $MSE = .049$, $p = .121$, $\eta^2 = .018$. These results are displayed in Figure 3.

Despite an attempt in the current experiment to highlight the benefits of summarizing and spacing by including a longer retention interval, it produced similar retention for summarizers and restudiers, and a pattern that indicated better retention for massing than spacing. This result will be discussed in more detail in the following sections.

Writing Analyses

Idea unit is defined the same way here as in the previous experiment. Independent samples *t*-test demonstrated that more idea units were included in the massed condition ($M = 3.94$, $SE = .27$) than in the spaced condition ($M = 3.06$, $SE = .26$), $t(68) = 2.37$, $p = .020$. This amounts to including, respectively, 39.4%% and 30.6% of the potential 10 idea units.

As in the first experiment, the number of idea units included in spaced summaries correlated with final recall, $r(35) = .47$, $p = .005$, whereas the same correlation for the massed condition was not significant, $r(35) = .10$, $p = .564$.

To determine the combined effect of including an idea unit in a summary and spacing, I ran a 2 Inclusion (included vs. not included) x 2 Spacing (spaced vs. massed) Repeated Measures ANOVA on the proportion of recalled idea units included vs. not included in summaries. Idea units included in summaries ($M = .68$, $SE = .036$) were more likely to be recalled than non-included idea units ($M = .34$, $SE = .027$), $F(1,64) = 97.87$, *MSE* = .041, $p < .001$, $\eta^2 = .605$. There was no main effect of spacing, $F(1,64) = 2.39$, $p = .127$, $\eta^2 = .036$. There was no Inclusion x Spacing interaction, $F < 1$. These results are displayed in Figure 4.

Replicating the results of Experiment 1, the number of idea units included in summaries correlated with final recall in the spaced condition, but not in the massed condition. Nonetheless, overall, spacing and massing resulted in similar retention. The finding that massed summaries included more idea units than spaced summaries suggests the spacing may have been too long, and participants experienced relative difficulty reorienting to the text and including idea units from it.

Item Analyses

Testing for item selection effects, Spearman's Rank Order correlations demonstrated that what was included in the summary did not predict what was recalled in any of the restudy conditions, but the correlations were relatively large in magnitude. Each of the following results represents the correlation between what was included in a given summary condition (e.g., spaced summary with an immediate test) and the analogous restudy condition (e.g., spaced restudy with an immediate test). For spaced immediate $r_s(10) = .35$, $p = .327$, for massed immediate, $r_s(10) = .40$, $p = .248$, for spaced delay, $r_s(10) = .27$, $p = .454$, and for massed delay, $r_s(10) = .367$, $p = .297$.

Similar to the first experiment, what participants included in summaries did not significantly predict what restudiers remembered, but the correlations were large in magnitude. Again, this creates caution in taking the inclusion effect at face value, as at least part of this effect appears to be due to item selection effects.

Difference Scores Analysis

The difference score in the massed condition (M = .20, SE = .041) was significantly greater than zero, $t(33) = 4.88$, $p < .001$, as was the difference score in the spaced condition, (M = .11, SE = .051), $t(32) = 2.06$, $p = .048$. These results suggest that items included in summaries were later answered at a higher rate by summarizers than the rate at which restudy participants answered those same questions.

Predictions

A Task (restudy vs. summary) x Spacing (spaced vs. massed) ANOVA on predictions of what would be remembered on the final test revealed that participants in the spaced condition did not differ from participants in the massed condition, $F < 1$. There was also no main effect of task, $F < 1$, as the participants in the restudy did not differ from participants in the summary condition. The two variables did not produce a significant interaction, $F(1,53) = 1.21$, $MSE = 580.15$, $p = .276$, $\eta^2 = .022$. These results are summarized in Figure 5.

Predictions did not significantly correlate with recall in the massed restudy condition, $r(13) = .28$, $p = .349$, in the spaced restudy condition, $r(11) = -.09$, $p = .786$, nor in the massed summary condition, $r(15) = .25$, $p = .375$. Predictions did significantly correlate with recall in the spaced summary condition $r(18) = .59$, $p = .010$.

The finding that writing a delayed summary improves predictions of future test performance, as was found here, has been observed in prior research (Anderson & Thiede, 2008; Thiede & Anderson, 2003). The current experiment builds on prior research by also including a restudy a condition. Although it has been shown that writing a delayed summary improves predictions relative to writing an immediate summary or doing nothing, it has not been shown what effect immediate or delayed restudying has. The current experiment provides evidence that summarizing has unique properties that contribute to metamemory.

CHAPTER IV

EXPERIMENT 3

Like the first experiment, the second experiment failed to find a benefit of summarizing relative to restudying. Additionally, no effect of spacing was found. It may be that the spacing effect is less robust with text passages than with word lists, given that spacing is inherent in rereading a text, making it difficult to obtain. Thus, in the third experiment, testing was used rather than spacing. The materials used in the current experiment were used by researchers who obtained a testing effect with them (Kang, McDermott, and Roediger, 2007). The current experiment set out to again compare summarizing and restudying, and to examine if retrieval modifies the effects of summarizing.

Participants in the current experiment all first read a text, and then either wrote a summary of it with the text present, wrote a summary of it with the text absent, recalled as much as they could from it, or restudied the text. A multiple-choice test was given 1-week later. If the effects of testing on memory are replicated here, then those in the text-absent summary and test groups should display better retention than the other two groups. If writing a summary is as effective as it is claimed to be, then writing a text-absent summary should result in better retention than traditional testing, and writing a text-present summary should result in better retention then restudying.

Method

Participants

A total of one-hundred-and-one UNCG undergraduates participated in this experiment for course credit. The 12 people who participated in Session 1 but did not return for Session 2 were eliminated from the analyses.

Materials

The text, which was about 2,500 words, was obtained from *Current Directions in Psychological Science* (Anastasio, Rose, & Chapman, 1999), and the eight multiple-choice questions used as the final test were created by Kang, McDermott, and Roediger (2007). The questions were presented in 10 different randomized orders.

Procedure

All participants began by reading the text and were given 15 min to do so, and told to reread the text if they finished early. After the 15 min, participants were given instructions for the intervening task. Those in the *restudy* condition were given 20 min to restudy the text. Those in the *text-present summary* were given 20 min to write a summary of the text with access to the text the entire time. Those in the *regular test* condition were given 13 min to write a bulleted list of everything they could remember from the text, without access to it. After 13 min, they were given the text for 7 min and instructed to review it and look for things they may have forgotten. Those in the *text-absent summary* were given 13 min to write a summary of the passage, and then returned the text, and given 7 min to review the text and update their summaries if they pleased.

The purpose of returning the text was so it could serve as feedback, as prior research using these same materials produced a testing effect only when feedback was provided (Kang, McDermott, & Roediger, 2007). The summary and restudy instructions were the same as those provided in the prior experiments.

Next, all participants were then given a questionnaire asking them to indicate, on a scale from 1 (Strongly Disagree) to 7 (Strongly Agree), how interesting they found the material they read and how much they enjoyed participating in the experiment. The questionnaire also asked that students choose a number between 1 and 100, corresponding to what percentage of questions they thought would remember in one week. All participants were asked to return to the lab after a week, and completed the multiple-choice test then.

Results

Final Test

To test for the effects of condition on retention, a one-way ANOVA on the proportion of final test questions answered correctly, with condition as the between subjects variable was ran, producing a main effect of condition, $F(3,85) = 2.92$, $MSE = .034$, $p = .039$, $\eta^2 = .093$. Follow-up tests revealed that the text-present summary group answered more questions correctly than the text-absent summary group, $t(44) = 2.91$, $p = .006$, as did the test group, $t(45) = 2.02$, $p = .050$. The restudy group did not differ from the text-present summary group, $t(40) = 1.21$, $p = .233$, the regular test group, $t < 1$, nor the text-absent summary group, $t(40) = 1.46$, $p = .153$, nor were any other comparisons significant. These data are summarized in Figure 6 below.

Unlike much prior research, the current experiment did not produce a testing effect. Experiments that have produced a testing effect with text passages have included at least one of two features missing from the current experiment: they either used short text passages (e.g., Storm, Bjork, & Storm, 2010; Karpicke & Roediger, 2010; Roediger & Karpicke, 2006b) or used longer texts but the content of the intervening tests was the same as the content on the final test, or similar to it (e.g., Butler, 2010; Chan, 2010; Kang, McDermott, & Roediger, 2007). In an attempt to measure practical benefits of testing with text passages, the current experiment met neither of those criteria, providing a plausible explanation for why a testing effect was not observed. In the case of a short text, on an intervening test, it is plausible to successfully retrieve most of a text. In fact, Roediger and Karpicke (2006b, Experiment 1) found that on an intermediate free recall test, participants were able to recall 70% of texts that were, on average, 266 words. In the current experiment, in which the text was about 2,500 words, on the intervening test, participants recalled only 23% of answers to what were given as later questions. This explanation is consistent with McDaniel, Howard, and Einstein's (2009) findings, in which they found that testing produced superior retention to note-taking when they used short passages, but the two were equivalent when the passages were longer and more complex.

Research that has obtained a testing effect with longer texts gives participants intervening tests with cues (e.g., short answer or multiple-choice tests) that are sufficient to boost successful retrievals relative to the free recall test used in the current experiment. Although this method provides evidence that this type of testing improves memory relative to reading the same statement with the answer provided, it does not reflect the type of retrieval practice students engage in when they do not know the questions that they will be asked on the final test.

Also of note is that participants in the test group remembered more on the final test than participants in the text-absent summary group, suggesting a negative effect of writing this type of summary. Both these groups included the same amount on the intermediate test, yet the test group remembered more on the final test. There were two major differences in procedures between these two groups. First, during the intervening task period, participants in the test condition were told to write everything they could remember from the text in list form, and those in the text-absent summary condition were told to integrate what they could remember into summary form. Second, during the feedback phase, the test group was told that they could review the text and their output, whereas in addition to these instructions, the text-absent summary group was told they could continue updating their summaries.

If the differences in retention were a result of the instructions administered during the feedback phase, it suggests that participants in the test condition used the feedback time to review the text, whereas participants in the text-absent summary group used the time to continue working on their summaries, without necessarily making use of the text. If successful retrievals did not persist during this period, it makes sense that the text-absent summary group scored poorly relative to the test group. If the instruction to list everything (test group) as opposed to create a summary (text-absent summary group) caused the retention differences, it suggests that writing a summary without access to the text is a particularly poor way of writing a summary.

Writing Analyses

Idea unit is defined the same way here as in the previous experiments. A one-way ANOVA revealed that the number of idea units differed across conditions, $F(2,64) = 18.27$, $MSE = 1.08$, $p < .001$, $\eta^2 = .368$. Follow-up independent-samples *t*-tests showed that the text-present summary group included more idea units ($M = 3.48$, $SE = .27$) than the test group ($M = 1.88$, $SE = .19$), $t(45) = 4.91$, $p < .001$, and the text-absent summary group ($M = 1.78$, $SE = .18$), $t(44) = 5.31$, $p < .001$, but the latter two did not differ, $t < 1$. This amounts to including, respectively, 43.5%, 23.5%, and 22.2% of the potential 8 idea units.

As discussed in the previous section, the difficulty of the intervening test for the retrieval groups limited the number of items that they could recall. Although text-absent summary participants were given the opportunity to update their summaries when returned the text, the fact their inclusion rate mirrored the test group suggests that they did not take the opportunity to do so.

To examine the effects of including an item in the summary and test conditions, a 2 Inclusion (included vs. not included) x 3 Condition (text-present summary vs. text-absent summary vs. test) Repeated Measures ANOVA was conducted on the probability of recalling an item. There was a main effect of inclusion, $F(1,61) = 139.41$, $MSE = .048$, $p < .001$, $\eta^2 = .668$, such that the probability of remembering an item that was included in a protocol ($M = .89$, $SE = .031$) was greater than the probability of remembering an item that was not included ($M = .46$, $SE = .028$). The main effect of condition was not significant, $F(2, 64) = 1.67$, $MSE = .073$, $p = .197$, $\eta^2 = .049$. There was no Inclusion x Condition interaction, $F < 1$. Figure 7 below summarizes these data.

Replicating the prior experiments, including an idea unit in a summary was associated with a large increase in the probability of later remembering that idea unit.

Item Analyses

Spearman's Rank Order correlations were conducted to test for item selection effects. The correlation between what was included in the text-present summary condition and what was remembered in the restudy condition was $r_s(8) = .78$, $p = .022$. What was included in the test condition was also significantly correlated with what was remembered in the restudy condition, $r_s(8) = .94$, $p = .001$, and what was included in the text absent summary condition correlated with what was recalled in the restudy condition $r_s(8) = .90$, $p = .003$.

These item analyses follow the same general pattern of the prior experiments, but in the current experiment the correlations were significant. These data convincingly demonstrate that including an item in a protocol is predictive of what a participant who does not write anything will remember. Thus, these data support the idea that inclusion effects can be at least part attributed to item selection effects.

Difference Scores Analysis.

The difference score in the text-present summary condition (M = .18, SE = .025) was significantly greater than zero, $t(22) = 7.12$, $p < .001$. The difference score in the regular test condition (M = .082, SE = .060) was not different than zero, $t(22) = 1.37$, p = .19, nor was in the text-absent summary condition, (M = .082, SE = .060), $t < 1$.

These results diverge considerably from the item analyses. Although the item analyses suggest that writers included the easiest items in their summary, thus making it appear as if there is a benefit to including an item in a summary, the difference scores analysis suggests that there is in fact a benefit to including items in text-present summaries. It may be that although item selection effects contributed to the inclusion effect, inclusion still results in a benefit beyond what is accounted for by item selection effects.

Questionnaires

The groups did not differ in their interest in the topic, $F(3,84) = 1.53$, MSE = 2.05, $p = .213$, $\eta^2 = .051$, how much they enjoyed participating, $F(3,84) = 1.76$, MSE = 1.64, $p = .161$, $\eta^2 = .059$ nor in the percentage of questions they predicted they would remember on the final test, $F < 1$. These data are summarized in Figures 8 - 10.

Overall, people's recall predictions were not related to actual memory, $r(88) = .026$, $p = .811$. Examining this correlation by condition reveals that the only exception to this null occurred in the text-present summary condition, in which the two variables were significantly related, $r(23) = .42$, $p = .047$. For the restudy group the correlation was $r(19) = -.10$, $p = .678$, for the text-absent summary group $r(23) = -.23$, $p = .287$, and for the test group, $r(23) = .057$, $p = .796$.

The finding that text-present summary participants were the most accurate in their prediction of future test performance replicates Experiment 2. It also extends the findings by showing that writing a text-absent summary does not create the same benefit.

CHAPTER V

EXPERIMENT 4

The third experiment resulted in another failure to find memory benefits of summarizing relative to restudying. The most consistent finding across the first three experiments was that, relative to not including a given item, including something in a summary resulted in a striking increase in the probability of later remembering that item. The difference score analyses provided evidence that writing a text-present summary facilitates retention more so than restudying, but as discussed, this finding comes with caveats. And although the item analyses hinted that these results were due to item selection effects, these results did not rule out the possibility that there is in fact a benefit of including something in a summary relative to restudying it. The goal of the current experiment was to experimentally, rather than statistically, determine the value of including an item in a summary as opposed to restudying it. The prior experiments have not allowed me to know what participants in the restudy condition are focusing on. The current experiment set out to resolve this limitation by rather than simply having participants restudy a text, asking them to underline the parts of the text that they would include in a summary if they were to write one. More than comparing overall memory on the final test, this condition allowed me to compare the probability of remembering an item that was included in a summary to the probability of remembering an item that was underlined but not actually written. This helps determine the value of including writing an item, relative to simply reading it.

Consistent with the goal of the current experiment, underlining is not a helpful technique for improving overall text retention (e.g., Johnson, 1988; Nist & Hogrebe, 1987; Silvers & Kreiner, 1997). The primary goal of the current experiment is to have underlining elucidate what people focus on when they restudy a text, and compare retention of the parts of the text that receive special attention (in this case, underlined items) to items that are included in summaries. Thus, it is useful that underlining does not appear to provide a benefit beyond reading without underlining.

If the results of the prior experiments are replicated, then there should be no differences in overall retention between the underline and summary groups. If the inclusion effect is more than an artifact of item selection effects, then including an item in a summary vs. not should offer a larger benefit than the difference between underlining an item vs. not.

Method

Participants

Forty-six UNCG undergraduates participated in this experiment for course credit.

Materials

The text used in this experiment was an article from *Current Directions in Psychological Science* (Treiman, 2000) about literacy. The article was about 2,500 words long. Kang, McDermott, and Roediger (2007), who originally used these materials, also created eight multiple-choice questions that were used here as the final test. Two-digit math problems that required multiplying two numbers were included as a distractor. The same 3-item questionnaire as the previous experiment was used.

Procedure

After signing the consent form, participants were given up to 20 min to read the text. Participants in the *summary* condition were given the same instructions as text-present summary participants in the previous experiment. Participants in the *underline* condition were given the same instructions as participants in the summary condition (except that "summary" was replaced with "underline"). In addition, participants in the underline condition were asked to underline the parts of the text that they would include in a summary and told to put corresponding numbers next to where in the summary each underlined portion of the text would appear (e.g., they were instructed to put a "1" next to the underlined part of the text that would serve as the first sentence). Summarizers and underliners were given 12 min, and told to continue looking over the materials if they finished before time was up. They were then given the same questionnaire that was administered in the prior experiment, which was followed by the math distractor activity for 5 min. Lastly, participants were given the 8 multiple-choice questions, debriefed, and dismissed.

Results

Writing Analyses

Idea unit was defined here the same way as in the prior experiments. (For underliners, idea unit was defined as underlining a part of the text that is the answer to a final test question). Underliners underlined more idea units (*M* = 2.30, *SE* = .28) that would appear on the final test than summarizers wrote (*M* = 1.04, *SE* = .21), $F(1,43) = 12.39$, *MSE* = 1.44, $p = .001$, $\eta^2 = .288$. This amounts to including, respectively, 28.8% and 13.0% of the potential 8 idea units.

To examine the effect of including an item on retention by condition, a 2 Inclusion (included vs. not included) x 2 Condition (underline vs. summary) Repeated Measures ANOVA was conducted on the proportion of items remembered on the final test. The results are displayed in Figure 11. This analysis tests the prediction that including an item in a summary should boost retention more than underlining an item, which is an ineffective memory strategy. There was a main effect of inclusion $F(1,34) = 8.26$, *MSE* = .046, $p = .007$, $\eta^2 = .195$, such that the probability of remembering an included item (*M* = .81, *SE* = .053) was higher than the probability of remembering a non-included item (*M* = .68, *SE* = .029). No significant main effect of condition was observed, $F(1,34) = 2.51$, $p = .122$, $\eta^2 = .069$. There was also an Inclusion x Condition interaction, $F(1,34) = 5.89$, $p = .021$, $\eta^2 = .148$, such that the probability of remembering an included item (*M* = .94, *SE* = .043) was higher in the summary condition than the probability of remembering a non-included item (*M* = .67, *SE* = .047), $t(15) = 4.23$, $p =$

.001, but in the underline condition, the probability of remembering an underlined item (M = .70, SE = .084) did not differ from the probability of remembering a non-underlined item (M = .68, SE = .048), $t < 1$.

The observed interaction confirms the prediction that writing an item in a summary creates retention benefits beyond reading a given item. Still, this result must be interpreted with caution, because summarizers did not write as much as participants underlined. Even though summarizers retained a higher proportion of what they did write, they remembered more of less, resulting in the same final test accuracy.

Final Test

A one-way ANOVA was conducted on the proportion of questions answered correctly and showed that the underline condition (M = .70, SE = .044) and the summary condition (M = .66, SE = .031), $F < 1$, did not differ.

Item Analyses

Testing for item selection effects, Spearman's Rank Correlation showed that what was included in summaries, did not correlate with what was recalled in the underline condition, $r_s(8) = -.10$, p = .807. This finding is consistent with the interpretation of the inclusion data, which imply that including an item in a summary provides a unique contribution to retention.

Difference Scores Analysis

For this experiment, the underline condition was treated as the restudy condition. That is, I calculated the probability of writers correctly remembering items they included in their summaries, and subtracted the probability associated with underliners remembering only those items included in summaries. The difference score reflects the difference between the first and second proportion. The difference score for the writing condition was significantly greater than zero (M = .25, SE = .050), $t(15) = 4.97$, $p < .001$, suggesting that when limiting the comparison to included items, including items in summaries benefited retention more-so than spending that same time underlining.

Questionnaire

There were no differences in interest in the topic between the underline condition (M = 5.13, SE = .25) and the summary group (M = 5.0, SE = .25), $F < 1$. There were no differences in how much participants in the underline condition (M = 4.91, SE = .21) and the summary condition (M = 5.1, SE = .28) enjoyed participating in the experiment, $F < 1$. There were no differences in predictions between the underline condition (M = 69.35, SE = 3.21) and the summary condition (M = 66.65, SE = 2.94), $F < 1$.

Predictions did not correlate with recall in the underline condition $r(23) = -.19$, $p = .391$, nor in the summary condition $r(23) = -.14$, $p = .539$. In contrast to the prior two experiments, writing a summary failed to result in a significant correlation between what participants predicted they would remember and what they actually did. Unlike the prior experiments that produced this effect, the current experiment used a short retention interval.

CHAPTER VI

EXPERIMENT 5

Without fail, each experiment produced an inclusion effect (i.e., better retention for items included in summaries than those not). Still, it is unknown if this effect is causal; that is, does the act of writing produce the inclusion effect, or are these items that people would have remembered regardless of whether they were written or not? To address this question, two summary conditions were included in the current experiment. All participants began by reading a text. Then, in one summary condition, participants were given the same instructions as in the prior experiments. In another, participants were encouraged to summarize at least ten things they thought would be on a final test. The purpose of the latter summary condition was to determine if by encouraging participants to write more they would also remember more. Other participants restudied the text, and a fourth group, which served as a baseline control, engaged in an unrelated distractor after first reading the text. Everyone was then given a final short answer test. The baseline control group was included to determine if writing a summary improves retention relative to doing nothing, a question that has not been addressed in the prior experiments.

If writing about something that will appear as a question on a later test does improve memory, then participants in the new instruction group should remember more than participants in the other groups. This prediction is based on the assumption that giving the new instructions will result in participants including more idea units in their summaries than those given the regular instructions. If the results of the prior experiments are replicated, then the participants in the regular instructions group should recall the same amount as those in the restudy group. Assuming that writing and restudying offer a benefit beyond doing nothing, the regular instructions and restudy group should remember more than those in the control group.

Crucial to the expectation of the current experiment, prior research has found that altering instructions influences summary content and learning outcomes. For example, Christopherson (1981) trained and instructed college students either to include semantic roles in a summary, to include main points, or participants received no explicit instruction. It was correctly predicted that participants receiving semantic roles training would writer higher quality summaries, with higher scores indicative of including more main ideas in fewer words. Bean and Steenwyk (1984) compared 6th graders who were trained to summarize text either by identifying the global structure of the text or by identifying the parts of the text that contribute to its gist. Compared to students who received no training, trained students scored higher on a comprehension test and wrote higher quality summaries. Other studies using more extensive training procedures have also found an effect on summary content (e.g., Friend, 2001; Rinehart, Stahl, & Erickson, 1986; Taylor & Beach, 1984).

Although these strategies for improving summaries were effective, they do not accomplish the current goal of having participants write more of what they expect would be on the final test. Thus, the current instructions deviate from the existing summary training instructions.

Method

Participants

Eighty UNCG undergraduates participated in this experiment for course credit.

Materials

The text used in this experiment was an article from *Current Directions in Psychological Science* (Garry & Polaschek, 2000) about imagination inflation and is approximately 2,500 words. The same word search used in the prior experiments was used as the distractor task. The final test consisted of 8-short answer questions that I created. The questions were created after another graduate student, an undergraduate student and I independently selected what we thought were the 10 most important idea units of the text. I created questions from the 6 items we had 100% agreement on, and the two items that two out of three us of agreed on.

Procedure

Participants were given 15 min to read the article. Participants who finished early were instructed to reread the text. Next, participants in the *regular summary* and *restudy* conditions were given the same instructions as the respective conditions in the previous experiments. In the *new instruction summary* condition, participants were given the same instructions as in the regular instruction condition, except that they were told to select ideas that they predicted would be on the final test. They were instructed to include at least 10 such ideas in their summaries. Participants in all three conditions were given 20 min and instructed to continue reviewing the materials if they finished before time. Those in the *control* group spent the 20 min working on a word search.

Participants were then given the word search, asked to indicate what percentage of questions they thought they would answer correctly on the final test, and 5 min to work on the word search. Those in the control group continued on their word search. Participants were then given the final test questions and could use as much of the remaining hour as needed to complete the questions.

Results

For the purposes of these analyses, participants who scored 2.5 or more standard deviations above the mean for their group on the final test were removed from the analyses. Two participants met this criterion.

Final Test

This analysis tests the effect of condition on retention. The ANOVA on the proportion of final test questions answered correctly, with condition as the between-subjects variable, showed that condition did not have a significant effect on retention, $F(3,86) = 2.19$, $MSE = .033$, $p = .095$. $\eta^2 = .071$ (see Figure 12). However, planned independent samples *t*-tests did reveal that new instructions group ($M = .32$, $SE = .035$) scored higher on the final test than the control group ($M = .20$, $SE = .024$), $t(44) = 2.9$, $p = .006$, as did the regular instructions group ($M = .29$, $SE = .035$), $t(43) = 2.27$, $p = .028$. Participants in the restudy condition scored marginally higher ($M = .31$, $SE = .054$), $t(45)$

= 1.93, p = .060. These results show that although writing resulted in better memory than the control group, instructing participants to write more did not have an effect on memory beyond administering the regular instructions.

Writing Analyses

Idea unit is defined here the same way as in the previous experiments. If the instruction to write more were successful, participants in the new instructions group should have included more idea units than the regular instructions group. Independent samples t-test did not reveal any differences in the number of included idea units between the regular instructions group (M = 2.00, SE = .29) and the new instructions group (M = 2.40, SE = .28), $t(41) = 1.00$, demonstrating that summaries were not reliably influenced by the instruction. This amounts to including, respectively, 25% and 30% of the potential 8 idea units.

To examine the effects of including an item in a summary and condition on retention, a 2 Inclusion (included vs. not included) x 2 Condition (new instructions vs. regular instructions) Repeated Measures ANOVA was conducted on the proportion of questions recalled for items included in summaries and those that were not. There was a main effect of inclusion, $F(1,36) = 51.04$, MSE = .078, $p < .001$, $\eta^2 = .586$, as included items (M = .63, SE = .053) were more often remembered than non-included items (M = .17, SE = .031). There was no significant main effect of condition, $F < 1$, nor did the two variables produce an interaction, $F < 1$. These results are displayed in Figure 13. This finding replicates the results of each of the prior experiments, in which items included in summaries were retained at a higher rate than omitted items.

Given that participants in the two summary conditions did not differ in the number of idea units included suggests that students experience difficulty identifying the important information from a text. Even when explicitly instructed to include the most important points from the text in their summaries, participants generally failed to do so, demonstrating low agreement with the students who chose the questions.

Item Analyses

Although what was included in the writing conditions did not predict was recalled in the restudy or control group using traditional significance levels, each of the correlations was large in magnitude. The correlation between what was included in regular instruction summaries and what was recalled in the restudy condition was $r_s(8) = .49$, $p = .21$, and what was recalled in the control condition, $r_s(8) = .67$, $p = .069$. The correlation between what was included in new instruction summaries and was recalled in the restudy condition was, $r_s(8) = .56$, $p = .152$, and what was recalled in the control condition, $r_s(8) = .69$, $p = .059$. These results provide evidence that what was included in summaries at least in part predicted what was recalled by non-summary writers. These findings are suggestive that the inclusion effect is at least partly due to item selection effects.

Difference Scores Analysis

The difference score in the regular instructions condition (M = .22, SE = .084) was significantly greater than zero, $t(14) = 2.58$, $p = .022$, whereas the difference score in the new instructions condition, (M = .15, SE = .085) was not, $t(16) = 1.72$, $p = .11$. In other words, in the regular instructions condition, item inclusion benefited retention, whereas the effect was not significant in the new instructions condition. This difference may be due to the fact that participants in the latter condition included numerically more items in their summaries, and it may that those additional items were more difficult, thus reducing the benefit of including more items (i.e., diminishing returns).

Predictions

The predictions across the groups did not differ, $F < 1$. These results are displayed in Figure 14.

Predictions did not correlate with final test recall in any of the conditions, and overall was $r(84) = .12$, $p = .272$. Like the previous experiment (which also did not include a delayed test), but unlike Experiments 2 and 3 (which included a delayed test), writing a summary did not improve metamemory accuracy.

CHAPTER VII

EXPERIMENT 6

Each experiment thus far has failed to find a benefit of summarizing over restudying. Experiment 5 attempted to encourage participants to summarize more text that they thought they would be tested on as a way to directly measure writing's effect on memory. This goal was stymied, as participants who were instructed to write about the most important ideas wrote the same amount as participants who were not given the instruction. This failure of instruction may have been because "important idea" was not validly operationalized. The goal of the current experiment was to more objectively define the important parts of the text. This was accomplished by using a text that contained seductive details.

A seductive detail is a part of a text that piques the reader's interest, but is not relevant to the central message (Garner, Gillingham, & White, 1989). For example, a text that describes the process of how lightning is created might relay a seductive detail about the susceptibility of golfers and people in open fields to strikes (Harp & Mayer, 1997). Although interesting, knowing that golfers are at a higher risk to be struck does not help the reader understand the process of how lightning is formed. The drawback of seductive details is that they reduce overall text comprehension relative to text that does not contain seductive details (e.g., Lehman, Schraw, McCrudden, & Hartley, 2007; Mayer, Griffith, Jurkowitz, & Rothman, 2008; Sanchez & Wiley, 2007). If the function of writing a summary is to capture the important details of a text, then writing one should help inoculate against the detrimental effects of seductive details. An examination of the three hypotheses developed by Harp and Mayer (1998) of how seductive details negatively influence comprehension reveals that if summary writing achieves its function, then summary writers' retention of important ideas should be less impaired by seductive details than those who restudy a text.

According to the distraction hypothesis, seductive details steal attention away from main ideas, thus impairing recall of the latter. If this hypothesis is correct, summarizing should minimize the detrimental effects of seductive details, because one goal of a summary is to identify the important points of a text and devote attention to them (Friend, 2002; Radmacher & Latosi-Sawin, 1995). With the disruption hypothesis, readers experience difficulty creating a coherent mental model because seductive details interfere with the chain of events. Because a successful summary should contain a coherent representation of the chain of events in a text, it should help overcome the breaks in information created by seductive details. Lastly, the diversion hypothesis suggests that seductive details harm recall because readers form mental models of the seductive details, rather than of the main ideas. Again, if summarizing does in fact require people to consider the entire text and select the most important ideas, then writing a summary should help overcome the detrimental effects of seductive details.

Method

Participants

Fifty-seven UNCG undergraduates participated in this experiment for course credit.

Materials

The text used in this experiment was a 967 word-article about how lightning is formed and was originally used by Harp and Mayer (1998) and adapted by Lehman, Schraw, McCrudden, and Hartley (2007) to increase referential clarity. The text contained 9 idea units central to how lightning is formed (e.g., raindrops and ice crystals fall), and 12 seductive details tangential to the process of lightning (e.g., lightning struck a football player during practice). The same word search distractor that was used in prior experiments was included.

Procedure

After signing the consent form, participants were given 7 min to read the text. Participants were then assigned to either the *summary* condition or the *restudy* condition. Participants in the summary and restudy conditions were given the same instructions as in the previous experiments (summary participants were given the traditional, text-present summary instructions). Both groups were given 8 min to complete their task. After working on a word-search for 2 min and indicating their interest in the passage using the provided scale and what percentage of the text they predicted they would remember on a final test, participants were given 8 min to recall as much as they could from the text, and were then debriefed and dismissed.

Results

Final Test

The first analysis tested the hypothesis that writing a summary helps people to pick the important information out of a text. If this hypothesis were correct, then an interaction should have been observed such that restudiers remember more of the seductive details, but summarizers remember more of the main ideas. A 2 Condition (summary vs. restudy) x 2 Item Type (seductive vs. main idea) Repeated Measures ANOVA was conducted on the proportion of main ideas and seductive details recalled. The results are displayed in Figure 15. The results revealed a main effect of item type, such that seductive details (M = .41, SE = .021) were remembered at a higher rate than main ideas (M = .13, SE = .024), $F(1,55) = 59.92$, MSE = .037, $p < .001$, $\eta^2 = .520$. There was also a main effect of condition, $F(1,55) = 10.44$, MSE = .018, $p < .002$, $\eta^2 = .160$, such that participants in the restudy group (M = .31, SE = .018) recalled a higher proportion of items than participants in the summary group (M = .23, SE = .018). Although the interaction was not significant, $F(1,55) = 1.32$, MSE = .049, $p = .255$, $\eta^2 = .023$, participants in the restudy group recalled more seductive details (M = .47, SE = .030) than those in the summary group (M = .35, SE = .025), $t(55) = 3.16$, $p = .003$, whereas the two groups did not differ in recall of main ideas, $t < 1$.

The prediction that writing a summary helps pick out the important information from a text was not supported. Instead, replicating Experiment 1, restudying resulted in an overall benefit to retention relative to summarizing.

Writing Analyses

Idea unit is defined as including one of the predefined seductive details or main ideas in a summary. A paired samples *t*-test showed that the proportion of total seductive items included in summaries (M = .19, SE = .03) did not differ from the proportion of total main idea items in included in summaries (M = .18, SE = .04), $t < 1$.

To test for the effects of item type and inclusion on retention, an Item Type (main vs. seductive) x Inclusion (included vs. not included) Repeated Measures ANOVA was conducted on the proportion of recalled included and non-included main idea and seductive details. There was a main effect of inclusion, $F(1,12) = 8.34$, $MSE = .137$, $p = .014$, $\eta^2 = .410$, such that included items (M = .43, SE = .082) were recalled at a higher rate than non-included items, (M = .14, SE = .030). There was also a main effect of item type, $F(1,12) = 15.00$, $MSE = .063$, $p = .002$, $\eta^2 = .565$, as seductive items (M = .42, SE = .049) were remembered at a higher rate than main idea items (M = .15, SE = .050). There was no Inclusion x Item Type interaction, $F(1,12) = 1.47$, $MSE = .095$, $p = .228$, $\eta^2 = .109$. These results are displayed in Figure 15.

Item Analyses

Spearman’s Rank Order correlations revealed that seductive items that were included in summaries predicted seductive items that were recalled in the restudy condition, $r_s(12) = .75$, $p = .005$. Main ideas that were included in summaries did not predict main ideas that were recalled in the restudy condition, $r_s(9) = .23$, $p = .552$. The significant correlation for seductive items indicates that items included in summaries were items that participants were at a high likelihood to remember had they not been included in summaries. For main ideas, the correlation was not significant, but given that overall recall was low for these item types, the analysis may have suffered from restriction of range.

Difference Scores Analysis

For seductive items, the difference score was not reliably different than zero, ($M = .068$, $SE = .086$), $t < 1$, nor was it for main idea items ($M = .075$, $SE = .082$), $t < 1$. These results are consistent with the final test data, as both suggest that restudying was better or just as good as summarizing in terms of its effect on retention.

Interest

There was no difference in interest between the restudy ($M = 4.75$, SE .25) and the summary group (M= 4.59, $SE = .26$), $t < 1$.

Predictions

There was no difference in predictions between the restudy ($M = 69.93$, $SE = 4.49$) and the summary group ($M = 60.29$, $SE = 4.25$), $t(54) = 1.56$, $MSE = 535.40$, $p = .13$.

Predictions did not correlate with total recall in the summary group, $r(28) = -.011$, $p = .96$, nor in the restudy group, $r(28) = -.012$, $p = .95$.

CHAPTER VIII

GENERAL DISCUSSION

The most robust finding across the experiments was that the probability of later remembering an item that was included in a summary was greater than the probability of remembering an item was not included in a summary (the *inclusion effect*). At face value, this finding indicates that writing something causes increased retention. This effect, however, was at least partly a result of item selection effects as the items that participants included in summaries tended to correlate with what participants remembered in restudy conditions. Although these correlations were not always significant, they were generally large in magnitude, but underpowered given that the number of items included in these analyses was limited to the number of questions on the final test. With future research, this problem can be remedied by including tests with more questions. Then it would be possible to better untangle whether writing something improves retention, or if things that are written would be remembered regardless of whether they are included or not.

In addition, the difference score analyses provided preliminary evidence across the experiments that writing a summary does produce a benefit to retention, but such a benefit is limited to what is included. Thus, unlike testing, which facilitates retention of related but non-tested material (Chan, 2010), the benefits of writing a summary appear to be limited to what is written.

Contrary to the belief that writing a summary optimizes text retention, in six experiments, writing a summary did not improve retention relative to restudying, and, if anything, restudying was a more effective strategy for remembering text (Experiments 1 and 6). This general pattern persisted across variations in spacing, retrieval difficulty, retention interval, and text-length. The following section discusses an explanation for these results.

Summary Writing: No Better Than Restudying

Table 2.
Summary of Final Test Results

Experiment #	Results
1	Restudy > Summary
2	Restudy = Summary
3	Text-Present Summary = Test = Restudy; Text-Present Summary = Test > Text-Absent Summary
4	Restudy = Underline
5	Regular Instructions Summary = New Instruction Summary = Restudy; Regular Instructions Summary = New Instruction Summary > Control
6	Restudy > Summary

The results of the current experiments provide consistent evidence that if one's goal is to remember information from a text, writing a summary is no better than and in some cases inferior to restudying a text. This conclusion is in direct conflict with the overwhelming support summary writing has received. This raises the question, why was writing a summary so ineffective in the current experiments? The most apparent possibility is that writing a summary is a passive activity resulting in few attempts to understand the material being written about. An overview of how people summarize suggests this is an accurate description.

If summarizers tended to "integrate new material with what they already know" (Wade-Stein & Kitsch, 2004, p. 334) and "connect ideas in the passage" (Friend, 2002, p. 40) to what they write, it might be expected that writing has a larger effect on memory than it did here. But summarizers' writing behavior is the antithesis of the gist captured by these quotes. The following review demonstrates that summarizers avoid the very processes that are thought to improve retention.

Different researchers describe the cognitive operations that summarizers execute using different terminology, but Hidi and Anderson (1986) identified three processes common to most descriptions of summarizing. The rules are derived from Kintsch and van Dijk's (1978) model of text production, which include deletion, generalization, and construction. First, summarizers make decisions about what to include and what to delete from an original source. Second, summarizers substitute superordinates for lower level items. For example, a source text that includes the list *rigatoni*, *ziti*, and *spaghetti*, might be substituted by *pasta* in a summary. Third, to achieve summarization goals, writers must alter the surface structure of the source by integrating ideas within in it. Others have pointed out that if a topic sentence is missing, an ideal summary will contain a constructed one (e.g., Garner & McCaleb, 1985). Research suggests that writers implement the rules at a superficial level.

Sherrard (1986) labeled summarizers in her descriptive study as mechanical, reporting that 82% of all sentences participants wrote were created by one-to-one mapping and the deletion of source material. Suggesting that participants in Sherrard's small sample were not unique, Brown and Day (1983) found that when a text contains no topic sentence, participants invented one in their summaries only about half of the time. Garner and McCaleb (1985) replicated this finding, determining that summaries written by participants in their study scored poorly on a measure of integration and synthesis. Similar summary writing habits were reported by Winograd (1984), who found that participants combined two sentences from the source text about 60% of the time, reproduced sentences from the source text about 25% of the time, and create sentences whose surface structure is difficult to tie to the source text about 10% of the time. These findings converge with a verbal protocol study investigating monitoring differences between analytic and summary writers (Durst, 1989). Summary writers were less likely to focus on their knowledge of the topic and the task, reflected less on what was being written about, and spent less time planning. Instead, summarizers were more likely to devote their time to paraphrasing the original text and displayed more concern for surface elements (e.g., spelling, style) of the task rather than searching for meaningful features of the text. All of this is consistent with Penrose's (1992) conclusion that writing can be completed without much thought or reflection. Using the language of Bereiter and Scardamalia (1987), summary writing seems to encourage knowledge-telling, which is writing whatever comes to mind without implementing any restructuring.

Finally, beyond simply copying text directly into summaries, the limited benefit of summary writing may have been affected by participants' general inability to predict what they would be tested on. The number of items included in summaries that appeared on the final test was relatively low in each of the experiments. A manipulation in Experiment 5 designed to encourage participants to identify the important elements did not influence their success to do so. Summary writers demonstrated further struggles in identifying the important elements of a text in Experiment 6, in which they included the same proportion of irrelevant seductive items as main points. The positive effects of summarizing might increase with an increased ability to identify main points. The difference scores analyses were particularly illustrative of this hypothesis. In several of the experiments, the difference scores analyses demonstrated a relatively large increase in retention as a result of including items in a summary. Still, the ability to identify the main points of a text of a text might be enough to increase retention without needing to summarize. In other words, an increasing ability to identify main points might result in increased benefits from summarizing, but also be associated with increased retention from restudying.

Potential Benefits of Summarizing

The results of the current experiments provided no evidence that writing a summary offers measurable improvements in memory relative to restudying. Besides whatever writing practice is gained from creating a summary, that does not mean that there are no benefits to summarizing. The current section explores those benefits.

Metamemory

In each of the experiments that included a delayed test (Experiments 2 and 3), after creating some types of summaries, people's predictions of how well they would do on the final test significantly correlated with how well they actually did. This finding is with precedent.

Thiede and Anderson (2003) had participants summarize texts immediately after reading, after a delay, or not all. They found that people's predictions of what they would remember on a final test were more accurate in the delayed summary group than either of the two groups. They replicated this finding in later work (Anderson & Thiede, 2008), and attributed the result to participants in the delayed summary condition basing their judgments on the text's gist rather than on specific details. Judgments based on the former are thought to be more diagnostic of later comprehension, because comprehension requires a global understanding of a text.

There are consistencies and discrepancies with Anderson and Thiede's findings and the current findings. Consistent with Anderson and Thiede'swork, the current findings yielded a significant correlation between recall predictions and recall only in the delayed writing condition (i.e., spaced), but not in the immediate writing condition (i.e., massed) or restudy conditions (Experiment 2). In Experiment 3, recall predictions again correlated with recall, but only in the text-present summary group. Although this is not a true delayed-writing condition, the long nature of the text creates a situation that makes the subsequent writing functionally delayed.

Anderson and Thiede (2008) point out that their gist-based hypothesis (described above) applies only to text comprehension, not memory. Given that the questions in the current experiments were memory-based, a more suitable explanation for the current findings may be the accessibility hypothesis, which states that predictions are based on the amount of information that can be accessed from memory (Baker & Dunlosky, 2006; Benjamin, Bjork, & Schwartz, 1998). Based on this hypothesis, writing a summary should result in more accurate estimates of learning than restudying, because writing a summary provides participants a more explicit opportunity to assess what they know. Across each experiment, items included in summaries were remembered at a higher rate than items not included in summaries; summary writers may have sensed the special status of included items relative to non-included items, and used them as a guide for predicting their later memory.

Although the results of Experiments 2 and 3 suggest a metacognitive benefit of writing, the failure to replicate this finding in Experiments 4 – 6 demonstrates a boundary condition. Methodologically, the major difference between the first set of experiments and the second set is that the first had a delayed test, and the second set had an immediate test. Although evidence suggests that people do not account for forgetting across time when learning lists of words (Kornell & Bjork, 2009), when learning texts, participants might make adjustments, thus resulting in summarizers improving accuracy when the test is delayed.

Product Over Process

Another potential benefit of summarizing, and one that was not tested here, is that it is not the process of writing a summary that aids memory, but the product that it creates (Emig, 1977). Unlike restudying a text, writing a summary results in a product. An experiment on note-taking found that reviewing notes improved recall beyond the improvement gained from generating the notes (Kiewra, DuBois, Christian, McShane, Meyerhoffer, & Roskelley, 1991). In a similar vein, Reder and Anderson (1980) concluded that reading a summary helps remember a text just as well if not better than reading the entire text, because having extra details in a text may makes it more difficult to understand. The product of a summary might allow participants to review elements of a text that they would have otherwise forgotten.

An additional benefit of creating a product is that it helps instructors to assess what students do and do not know (Westby, Culatta, Lawrence, & Hall-Kenyon, 2010). Assuming students write about information that they understand or have a high likelihood of later remembering (as they did in each of the current experiments), an instructor can adjust the content of his or her teaching to focus on material that students tend to omit from summaries.

Integration Across Texts

The current experiments found that summarizing a text does not result in better retention than restudying. The case may be different when people summarize across multiple texts, which they did not do in the current experiments. By necessity, synthesizing multiple sources into one document requires a shift away from copying and shift towards the construction of a new text. Several experiments have demonstrated this principle.

Mateos and Sole (2009) found that older high school and college students' summaries from multiple documents included more synthesized products than non-synthesized ones. In contrast to what research on summarizing from a single source has found (e.g., Winograd, 1984), Segev-Miller (2007) reported that students in her study implemented a variety of conceptual, rhetorical, and linguistic transformations. Similarly, Wiley and Voss (1999) found that compared to students who wrote from a single source, students who wrote from multiple sources included more transformed sentences and connectives, as well as fewer sentences borrowed from the original material. When synthesizing documents, students must establish new connections, and in doing so, they create thematic chunks (Spivey, 1997). The habits that students exhibit while writing from multiple sources suggest that they create a synthesized organizational structure that did not previously exist (Segev-Miller, 2004). Having a coherent structure facilitates retention (Maury & Teisserenc, 2005; McNamara & Kintsch; 1996 McNamara, Kintsch, Songer, & Kintsch 1996), thus, writing from multiple sources might create conditions in which summarizing improves retention more than restudying does.

Conclusion

Summarizing has repeatedly been lauded as a method for improving text-retention relative to other learning strategies. In direct conflict with such praise, the current experiments demonstrate that summarizing is no better than restudying. The burden of discovering a theory that explains how summarizing improves text retention relative to restudying should be on those who still believe that it does. The burden, however, may fall to those who take a skeptical approach to the effects of writing on memory, as some think that “because of its ideological status, rhetoric and composition will not loosen their grip on writing as a mode of learning” (Ackerman, 1993, p. 361). It is not clear what it will take for advocates to loosen their grip, but the current set of experiments delivers a blow to the “movement that has produced no definitive study that confirms the relationship between writing to learn and learning to write, and no aggregate of studies [that] provides a compelling case for emphasizing writing as a unique tool for learning” (Ochsner & Fowler, 2004, p. 122).

REFERENCES

Ackerman, J.A. (1993). The promise of writing to learn. *Written Communication, 10,* 334-370.

Anastasio, P. A., Rose, K. C., & Chapman, J. (1999). Can the media create public opinion? A social-identity approach. *Current Directions in Psychological Science, 8,* 152-155.

Anderson, M.C.M & Thiede, K.W. (2008). Why do delayed summaries improve metacomprehension accuracy? *Acta Psychologica, 128,* 110-118.

Bahrick, H.P., Bahrick, L.E., Bahrick, A.S., & Bahrick, P.E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science, 4,* 316-321.

Baker, J. M. C., & Dunlosky, J. (2006). Does momentary accessibility influence metacomprehension judgments? The influence of study judgment lags on accessibility effects. *Psychonomic Bulletin and Review, 13,* 60–65.

Bangert-Drowns, R.L., Hurley, M.M., Wilkinson, B. (2004). The Effects of School-Based Writing-to-Learn Interventions on Academic Achievement: A Meta-Analysis. *Review of Educational Research, 74,* 29-58.

Bazerman, C., Little, J., Bethel, L., Chavkin, T., Fouguetter, D., & Garufis, J. (2005). *Reference guide to writing across the curriculum.* Retrieved from http://wac.colostate.edu/books/bazerman_wac/

Bean, T.W., & Steenwyk, F.L. (1984). The effect of three forms of summarization instruction on sixth graders' summary writing and comprehension. *Journal of Reading Behavior, 16,* 297-306.

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General, 127,* 55-68.

Benton, S.L., Kiewra, K.A., Whitfill, J.M., & Dennison, R. (1993). Encoding and External- Storage Effects on Writing Processes. *Journal of Educational Psychology, 2,* 267-280.

Bereiter, C., & Scardamalia, M. (1987). The psychology of written composition. Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, A.L., & Day, J.D. (1983). Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior, 22,* 1-14.

Brown, A.L., Day, J.D., & Jones, J.D. (1983). The developmental of plans for summarizing texts. *Child Development, 54*, 968-979.

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1118-1133.

Callender, A.A., & McDaniel, M.A. (2009). The limited benefits of rereading educational texts. *Contemporary Education, 34,* 30-41.

Carpenter, S. K., & DeLosh, E. L. (2005). Application of the testing and spacing effects to name-learning. *Applied Cognitive Psychology, 19,* 619-636.

Carpenter, S. K., & DeLosh, E. L. (2006). Impoverished cue support enhances subsequent retention: Support for the elaborative retrieval explanation of the testing effect. *Memory & Cognition, 34,* 268-276.

Carpenter, S. K., & Pashler, H. (2007). Testing beyond words: Using tests to enhance visuospatial map learning. *Psychonomic Bulletin & Review, 14,* 474-478.

Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of U. S. history facts. *Applied Cognitive Psychology, 23,* 760-771.

Carroll, D.W. (2008). Brief Report: A simple stimulus for student writing and learning in the introductory psychology course. *North American Journal of Psychology, 10,* 159-164.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132,* 354-380.

Chan, J. C. K. (2010). Long-term effects of testing on the recall of nontested materials. *Memory, 18*, 49-57.

Christopherson, S.L. (1981). Effects of knowledge of semantic roles on summarizing written prose. *Contemporary Educational Psychology, 6*, 59-65.

Davis, M., & Hult, R.E. (1997). Effects of writing summaries as a generative learning activity during note taking. *Teaching of Psychology, 24,* 47-49.

Day, S. (1994). Learning in large sociology classes: Journals and attendance. *Teaching Sociology, 22,* 151-165.

Delaney, P. F., & Knowles, M. E. (2005). Encoding strategy changes and spacing in free recall. *Journal of Memory and Language, 52,* 120-130.

Delaney, P. F., & Verkoeijen, P. P. J. L. (2009). Rehearsal strategies can enlarge or diminish the spacing effect: Pure versus mixed lists and encoding strategy. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *35,* 1148-1161.

Delaney, P. F., Verkoeijen, P. P. J. L., & Spirgel, A. S. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. *Psychology of Learning and Motivation, 53*, 63-147.

Dempster, F.N. (1996) Distributing and managing the conditions of encoding and practice. In E.L Bjork & R.A Bjork (Eds.). *Memory: Handbook of perception and cognition* (10, 317-44) New York: Academic Press.

Durst, R.K. (1989). Monitoring processes in analytic and summary writing. *Written Communication, 6*, 340-363.

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

Emig, J. (1977). Writing as a Mode of Learning. *College Composition*, *28,* 122-128.

Foos, P. (1995). The effect of variation in text summarization opportunities on test performance. *Journal of Experimental Education*, *63,* 89-95.

Friend, R. (2001). Effects of strategy instruction on summary writing of college student. *Contemporary Educational Psychology*, *26,* 3-24.

Friend, R. (2002). Summing it up: Teaching summary writing to enhance science learning. *The Science Teacher, 69,* 40-43.

Garner, R., & McCaleb, J.L. (1985). Effects of text manipulations on quality of written summaries. *Contemporary Educational Psychology, 10*, 139-149.

Garner, R., Gillingham, M.G., & White, C.S. (1989). Effects of "seductive details" on macroprocessing and microprocessing on adults and children. *Cognition & Instruction, 6,* 41-57.

Garry, M., & Polascek, D.L.L. (2000). Imagination and Memory. *Current Directions in Psychological Science, 9,* 6-10.

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

Harp, S.F., & Mayer, R.E. (1997). The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology, 1,* 92-102.

Harp, S.F., & Mayer, R.E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology, 90,* 414-434.

Hidi, S., & Anderson, V. (1986). Written Summaries: Task demands, cognitive operations, and implications for instruction. *Review of Educational Research, 56,* 473-493.

Hintzman, D. L. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola symposium* (pp. 77-99). Erlbaum: Hillsdale, NJ.

Horton, P.B., Fronk, R.H., & Walton, R.W. (1985). The effect of writing assignments on achievement in college general chemistry. *Journal of Research in Science Teaching, 22,* 535-541.

Johnson, L.L. (1988). Effects of underlining sentences on passage and sentence retention. *Literacy and Instruction, 28,* 18-32.

Kang, S.H.K., McDermott, K.B. & Roediger, H.L. (2007). Test format and corrective feedback modulate the effect of testing on memory retention. *The European Journal of Cognitive Psychology, 19*, 528-558.

Karpicke, J.D., & Roediger, H.L. (2008). The critical importance of retrieval for learning. *Science, 319*, 966-968.

Karpicke, J. D., & Roediger, H. L. (2010). Is expanding retrieval a superior method for learning text materials? *Memory & Cognition, 38,* 116-124.

Karpicke, J.D., & Roediger, H.L. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language*, *57*, 151-162.

Karpicke, J.D., Butler, A.C., & Roediger, H.L. (2009). Metacognitive strategies in student learning: Do students practice retrieval when they study on their own? *Memory, 17,* 471-474.

Kiewra, K.A., DuBois, N.F., Christian, D., McShane, A., Meyerhoffer, M., & Roskelley, D. (1991). *Journal of Educational Psychology, 83,* 240-245.

King, A. (1992). Comparison of self-questioning, summarizing, and notetaking-review as strategies for learning lectures. *American Educational Research Journal, 29,* 303-323.

King, J.R., Biggs, S., & Lipsky, S. (1984). Students self-questioning and summarizing as reading study strategies. *Journal of Reading Behavior*, *16,* 205-218.

Kintsch, W., & van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363-394.

Kirby, J.R., & Pedwell, D. (1991). Students' approaches to summarisation. *Educational Psychology, 11*, 297-307.

Kornell, N., & Bjork, R.A. (2009). A stability bias in human memory: Overestimating remembering and underestimating learning. *Journal of Experimental Psychology: General, 138*, 449-468.

Krug, D., Davis, B., & Glover, J.A. (1990). Massed versus distributed repeated reading: A of forgetting helping recall? *Journal of Educational Psychology, 75,* 40-49.

Kuo, T., & Hirshman, E. (1996). Investigations of the testing effect. *American Journal of Psychology, 109,* 451-464.

Lehman, S., Schraw, G., McCrudden, M.T., & Hartley K. (2007). Processing and recall of seductive details in scientific text. *Contemporary Educational Psychology, 32,* 569-587.

Mateos, M., & Sole, I. (2009). Synthesizing information from various texts: A study of procedures and products at different educational levels. *European Journal of Psychology of Education, 4,* 435-451.

Maury, P., & Teisserenc, A. (2005). The role of connectives in science text comprehension and memory. *Language and Cognitive Processes, 20,* 489-512.

Mayer, R.E., Griffith, E., Jurkowitz, I.T.N., & Rothman, D. (2008). Increased interestingness of extraneous details in a multimedia science presentation leads to decreased learning. *Journal of Experimental Psychology: Applied, 14,* 329-339.

McDaniel, M.A., Howard, D., & Einstein, G.O. (2009). The read-recite-review strategy: Effective and portable. *Psychological Science, 20,* 516-522.

McLeod, S. & Shirley, S. (1987). National Survey of Writing Across the Curriculum Program. Retrieved from http://wac.colostate.edu/books/mcleod_programs/appendix.pdf

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

McNamara, D.S., Kintsch, E., Songer, N.B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*, 1-43

Nist, S., & Hogrebe, M.C. (1987). The role of underlining and annotating in remembering textual information. *Reading Research and Instruction, 27,* 12-25.

Ochsner, R., & Fowler, J. (2004). Playing devil's advocate: Evaluating the literature of the WAC/WID movement. *Review of Educational Research, 74,* 117-140.

Palmquist, M., Melzer, D., Bernhardt, S.A., Brooks, K., Jory, J., Zawacki, T.M.,…De Pew, K.E. The WAC Clearinghouse. Retrieved from http://wac.colostate.edu/intro/pop3a.cfm.

Pashler, H., Rohrer, D., Cepeda, N. J., & Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin & Review, 14*, 187-193

Penrose, A.M. (1992). To Write or Not to Write: Effects of Task and Task Interpretation on Learning through Writing. *Written Communication, 9,* 465-500.

Radmacher, S.A., & Latosi-Sawin, E. (1995). Summary Writing: A tool to improve student comprehension and writing in psychology. *Teaching of Psychology, 22,* 113-115.

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

Reder, L.M. & Anderson, J.R. (1980). A comparison of texts and their summaries: Memorial consequences. *Journal of Verbal Learning and Verbal Behavior, 19*, 121-134.

Rinehart, S.D., Stahl, S.A., & Erickson, L. (1986). Some effects of summarization training on reading and studying. *Reading Research Quarterly, 21,* 422-438.

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

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

Rogers, B. (2001). TOEFL CBT Success. Princeton, NJ: Peterson's.

Rohrer, D., & Taylor, K. (2006). The effects of overlearning and distributed practice on the retention of mathematics knowledge. *Applied Cognitive Psychology, 20,* 1209-1224.

Sanchez, C. & Wiley, J. (2006). Effects of working memory capacity on learning from illustrated text. *Memory & Cognition, 34,* 344-355.

Segev-Miller, R. (2004). Writing from sources: The effect of explicit instruction on college students' processes and products. *L1-Educational Studies in Language and Literature, 4,* 5-33.

Segev-Miller, R. (2007) Cognitive processes in discourse synthesis: the case of intertextual processing strategies, in: L. Van Waes, M. Torrance & D. Galbraith (Eds) Writing and cognition: research and applications, Studies in Writing, 20, 231–250.

Sherrard, C. (1986). Summary writing: A topographical study. *Written Communication, 3,* 324-343.

Silvers, V. & Kreiner, D. (1997). The effects of pre-existing inappropriate highlighting on reading comprehension. *Reading Research and Instruction, 36,* 217-223.

Spivey, N. N. (1997).*The constructivist metaphor: Reading, writing, and the making of meaning*. San Diego, CA: Academic Press.

Spurlin, J. E., Dansereau, D. F., O'Donnell, A. M., & Brooks, L. W. (1988). Text processing: Effects of summarization frequency on text recall. *Journal of Experimental Education, 56,* 199–202.

Stein, B.L., & Kirby, J. (1992). The effects of text absent and text present condition on summarization and recall of text. *Journal of Literacy Research, 24,* 217-232.

Storm, B.C., Bjork, R.A. & Storm, J.C. (2010). Optimizing retrieval as a learning event: When and why expanding retrieval practice enhances long-term retention. *Memory & Cognition,38,*244-253.

Taylor, B.T., & Beach, R.W. (1984). The effects of text structure instruction on middle-grade students' comprehension and production of expository text. *Reading Research Quarterly, 19,* 134-146.

Thiede, K.W., & Anderson, M.C.M (2003). Summarizing can improve metacomprehension accuracy. *Contemporary Educational Psychology, 28,* 129-160.

Treiman, R. (2000). The foundations of literacy. *Current Directions in Psychological Science,* 9, 89-92.

Verkoeijen, P. P. J. L., Delaney, P. F., Bouwmeester, S., Coppens, L. C., & Spirgel, A.S. (In Preparation). No testing effect in category lists.

Verkoeijen, P.J.L., Rikers, R.M.J.P, & Ozsoy (2008). Distributed rereading can hurt the spacing effect in text memory. *Applied Cognitive Psychology, 22,* 685-695.

Wade-Stein, D., Kintsch, E. (2004). Summary Street: Interactive Computer Support for writing. *Cognition & Instruction, 22,* 333-362.

Weinstein, Y., McDermott, K.B., & Roediger, H.L. (2010). A comparison of study strategies for passages: Re-reading, answering questions, and generating questions. *Journal of Experimental Psychology: Applied*, *16,* 308-316.

Westby, C., Culatta, B., Lawrence, B., & Hall-Kenyon, K. (2010). Summarizing expository texts. *Topics in Language Disorders, 30,* 275-287.

Wiley, J. & Voss, J. F. (1999) Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology, 91*, 301-311.

Winograd, P.N. (1984). Difficulties in summarizing texts. *Reading Research Quarterly, 19,* 404-425.

Wormeli, R. (2004). *Summarization in any subject: 50 techniques to improve student learning.* Alexander, VA: Association for Supervision and Curriculum.

APPENDIX A

FIGURES

Figure 1.
Experiment 1 proportion of final test questions correct as a function of task and spacing. Error bars represent $\pm$ *SE*.

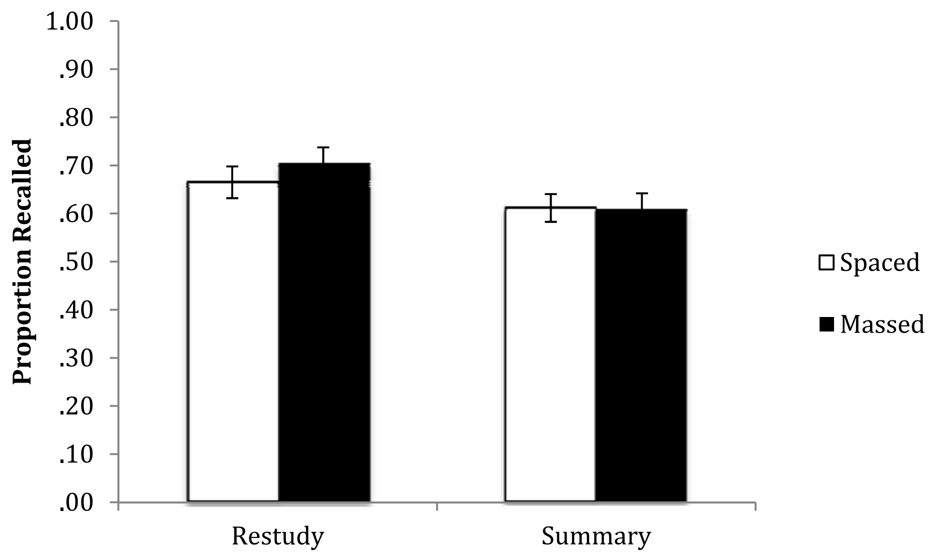


Figure 2.
Experiment 1 probability of recalling an item if it was included in a summary or not, as a function of spacing. Error bars represent ± *SE*.

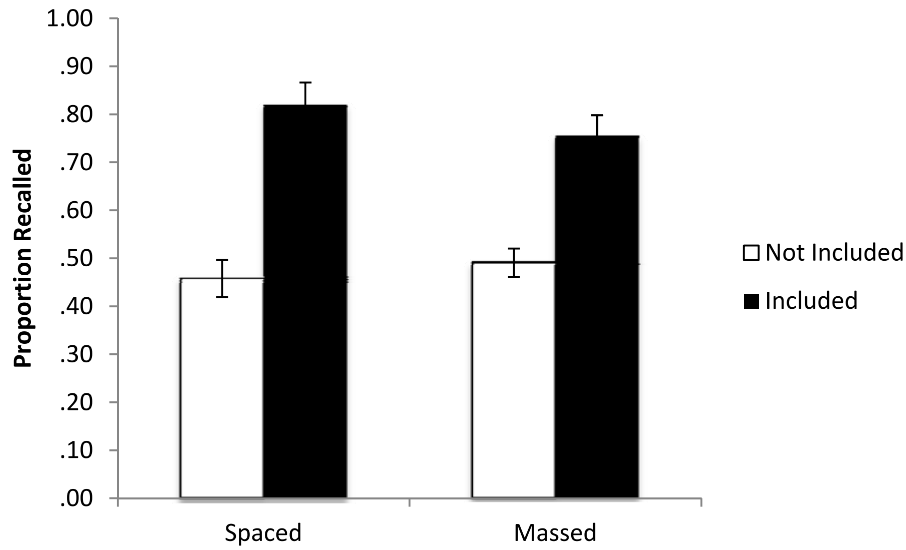


Figure 3.
Experiment 2 proportion of final test questions correct as a function of task, spacing, and retention interval. Error bars represent ± *SE*.

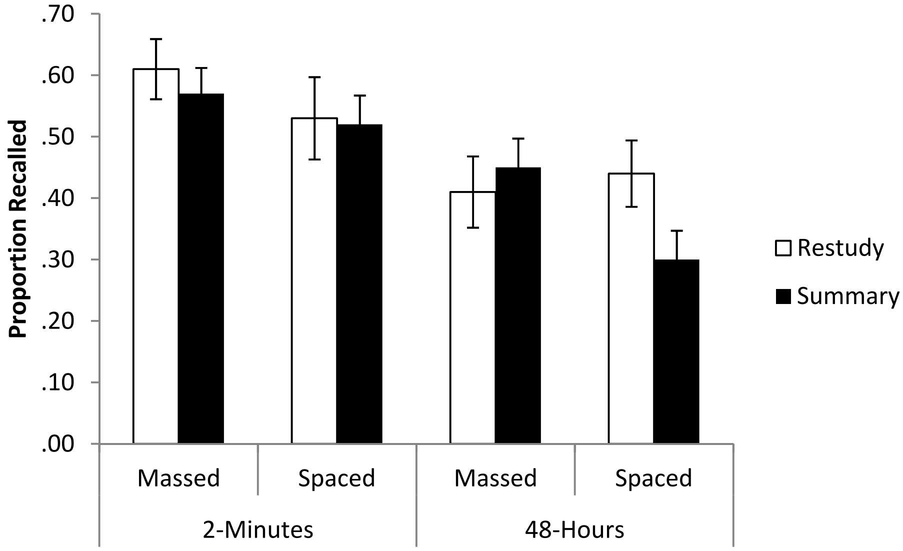


Figure 4.
Experiment 2 probability of recalling an item if it was included in a summary or not, as a function of spacing. Error bars represent ± *SE*.

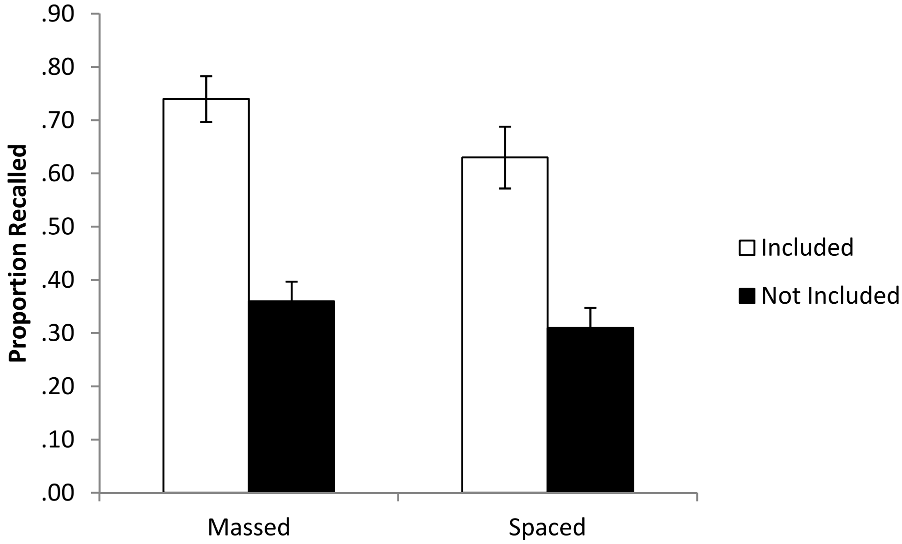


Figure 5.
Experiment 2 predictions of final test memory as a function of spacing and task. Error bars represent ± *SE*.

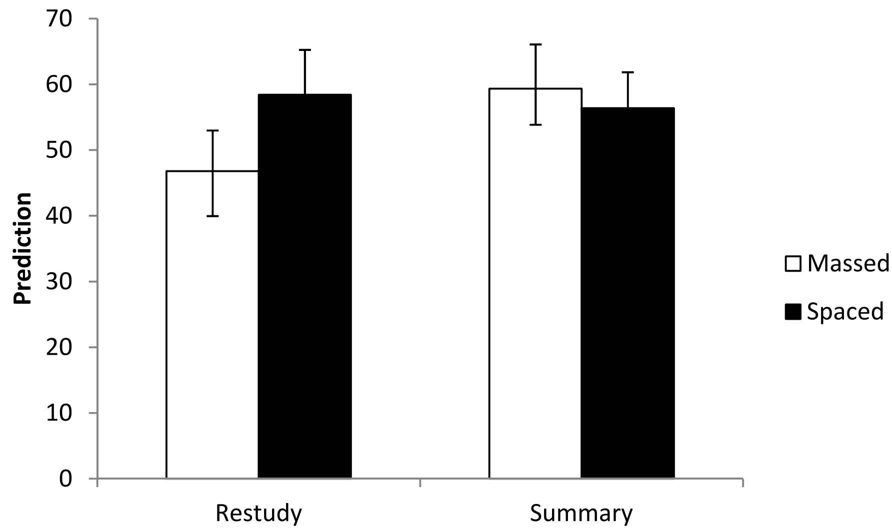


Figure 6.
Experiment 3 proportion of final test questions correct as a function of condition. Error bars represent ± *SE*.

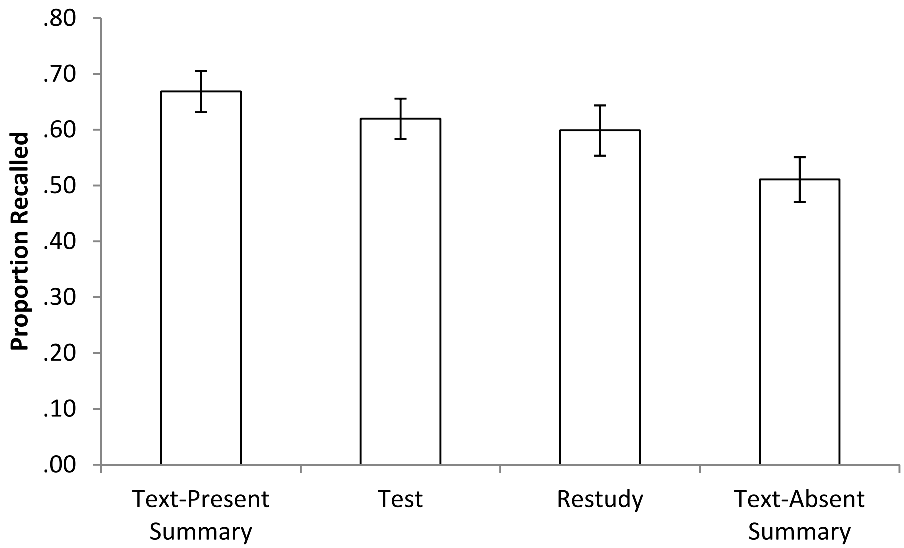


Figure 7.
Experiment 3 probability of recalling an item if it is included in a summary or not, as a function of whether it was included during a task or not, and condition. Error bars represent ± *SE*.

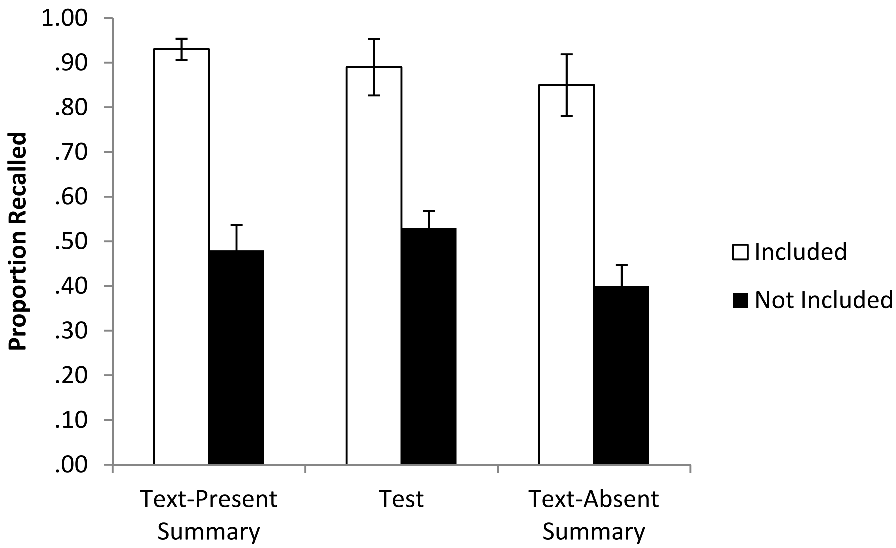


Figure 8.
Experiment 3 interest as a function of condition. Error bars represent ± *SE*.

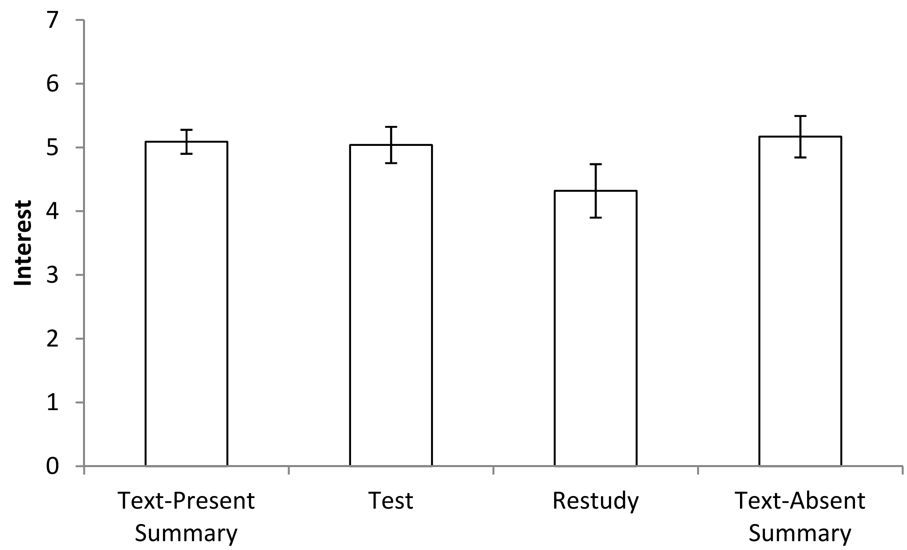


Figure 9.
Experiment 3 enjoyment as a function of condition. Error bars represent ± *SE*.

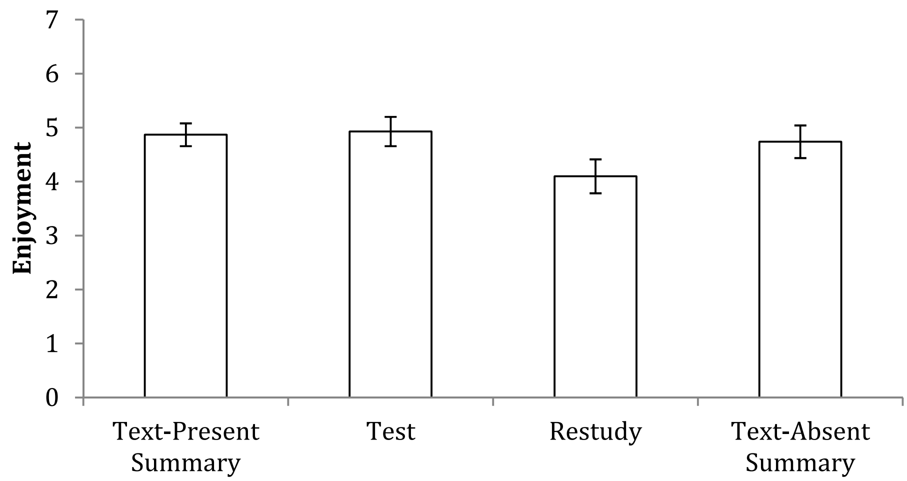


Figure 10.
Experiment 3 predictions of final test memory as a function of condition. Error bars represent ± *SE*.

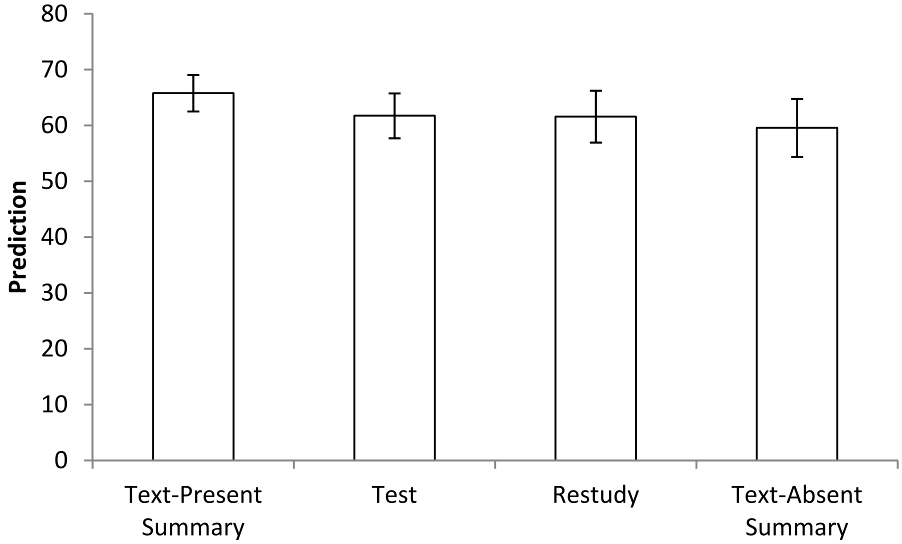


Figure 11.
Experiment 4 probabilities of recalling an item as a function of whether it was included during a task or not and condition. Error bars represent ± *SE*.

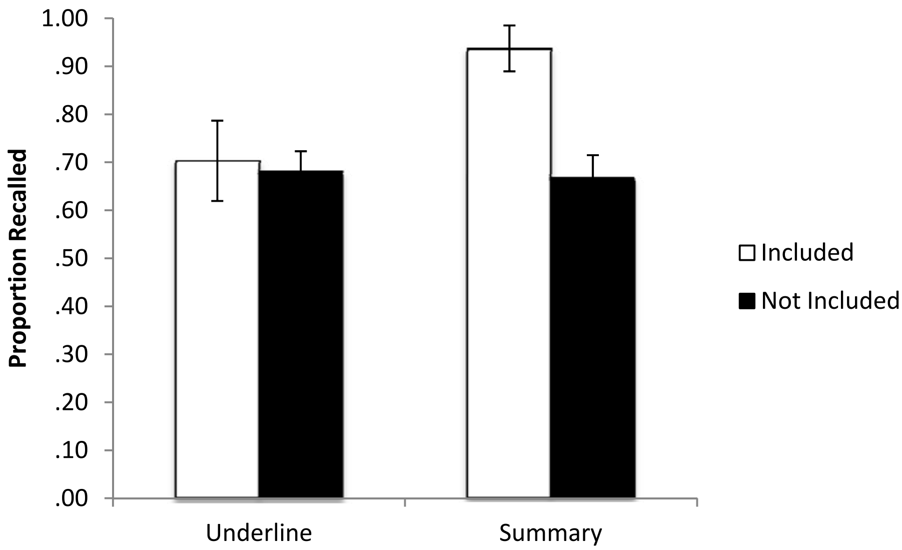


Figure 12.
Experiment 5 proportion of final test questions correct as a function of condition. Error bars represent ± *SE*.

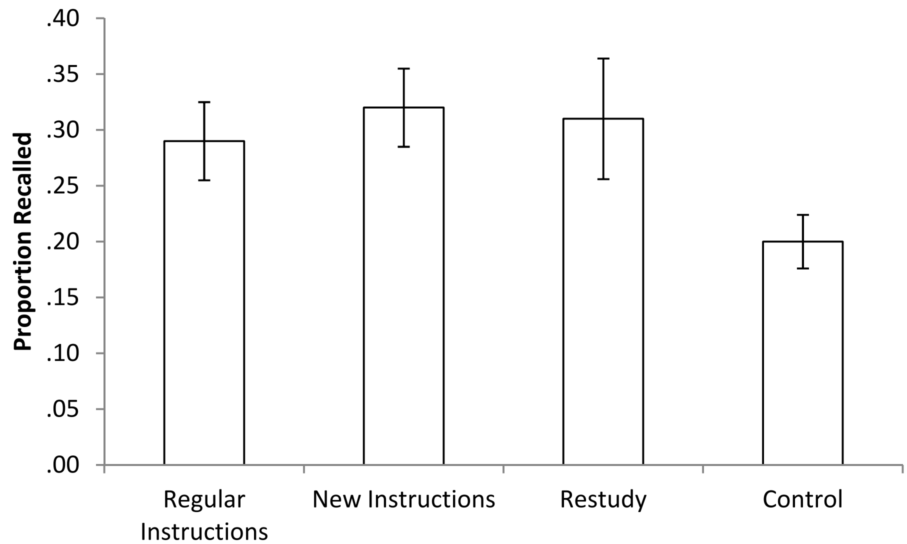


Figure 13.
Experiment 5 probability of recalling an item if it was included in a summary or not, as a function of inclusion, and item type. Error bars represent ± *SE*.

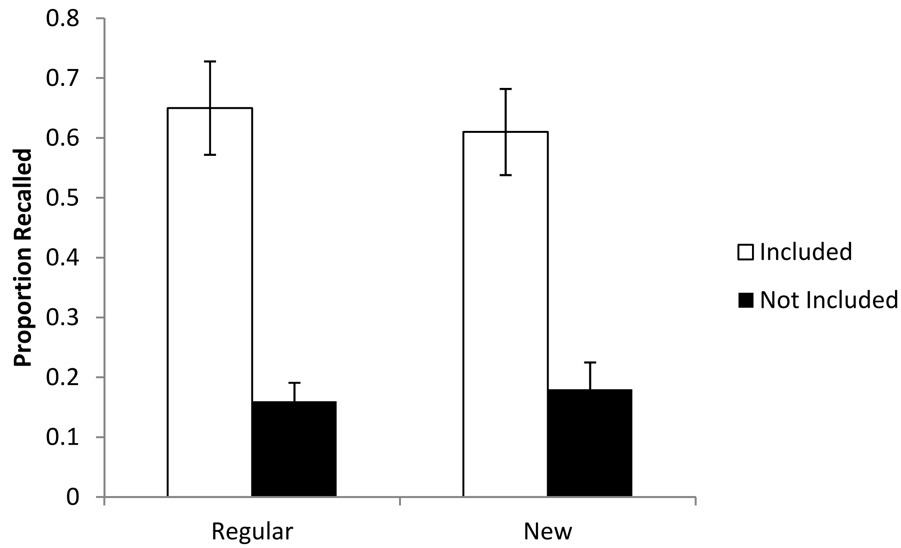


Figure 14
Experiment 5 predictions of final test memory as a function of condition. Error bars represent ± *SE*.

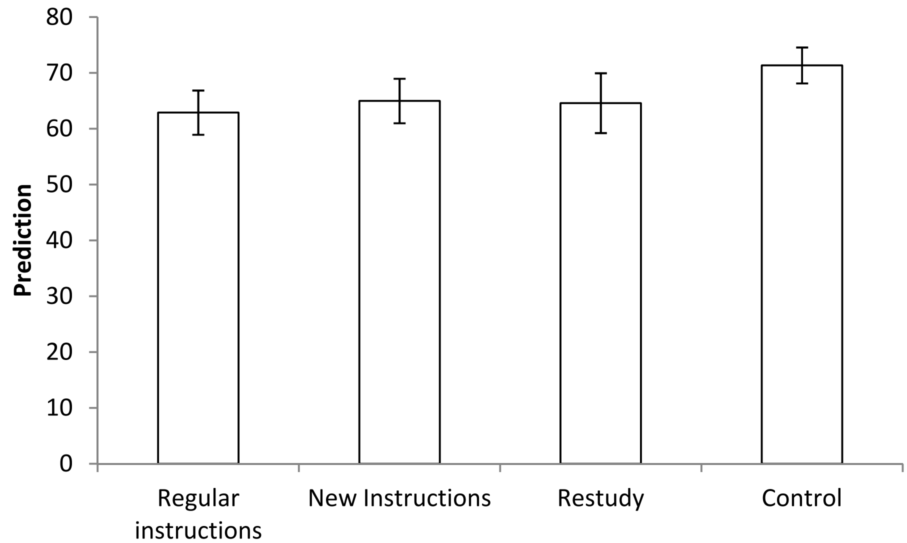


Figure 15.
Experiment 6 proportion of main ideas and seductive details recalled as a function of condition. Error bars represent ± *SE*.

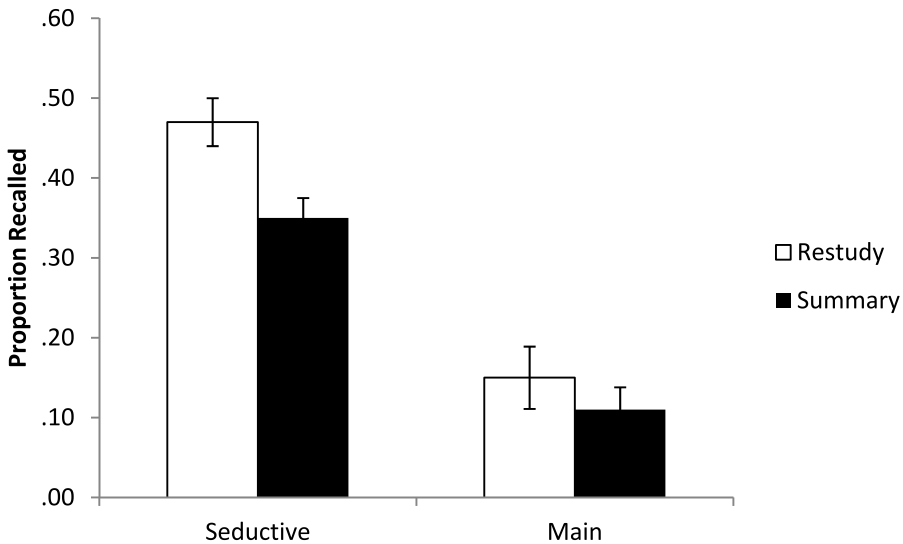
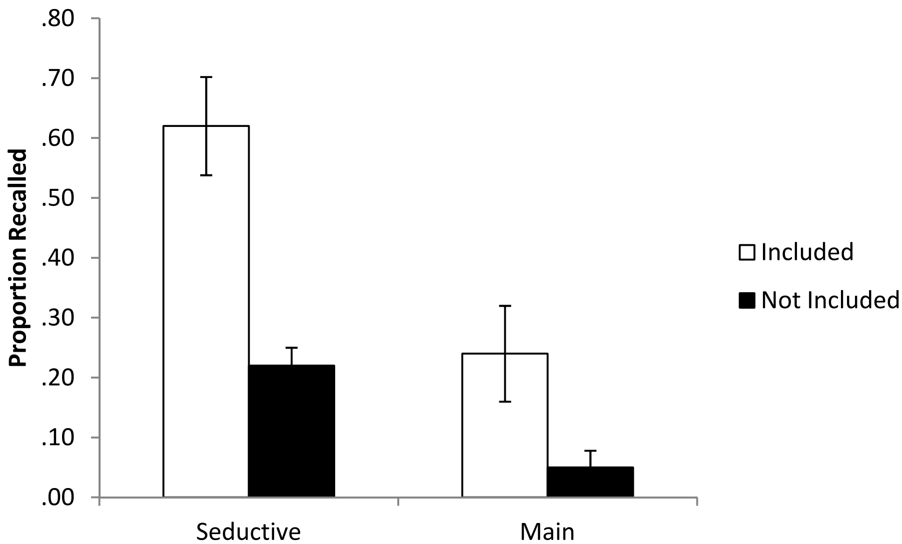


Figure 16.
Experiment 6 probability of recalling an item if it was included in a summary or not, as a function of item type, and inclusion. Error bars represent ± *SE*.



APPENDIX B

MATERIALS

Passage and Questions for Experiment 1

SUN:

The Sun today is a yellow dwarf star. It is fueled by thermonuclear reactions near its center that convert hydrogen to helium. The Sun has existed in its present state for about 4 billion, 600 million years and is thousands of times larger than the Earth.

By studying other stars, astronomers can predict what the rest of the Sun's life will be like. About 5 billion years from now, the core of the Sun will shrink and become hotter. The surface temperature will fall. The higher temperature of the center will increase the rate of thermonuclear reactions. The outer regions of the Sun will expand approximately 35 million miles, which is about the distance to Mercury. The Sun will then be a red giant star. Temperatures on the Earth will become too hot for life to exist.

Once the Sun has used up its thermonuclear energy as a red giant, it will begin to shrink. After it shrinks to the size of the Earth, it will become a white dwarf star. The Sun may throw off huge amounts of gases in violent eruptions called nova explosions as it changes from a red giant to a white dwarf.

After billions of years as a white dwarf, the Sun will have used up all its fuel and will have lost its heat. Such a star is called a black dwarf. After the sun has become a black dwarf, the Earth will be dark and cold. If any atmosphere remains there it will have frozen onto the Earth's surface.

1. What type of star is the sun today?
2. What fuels the sun?
3. How long has the sun been it is present state?
4. What can astronomers study to predict the rest of the sun's life?
5. What two major changes will happen to the core of the sun in about 5 billion years from now?
6. In about 5 billions years, about how many miles will the outer regions of the sun expand?
7. What type of star will the sun be in about 5 billion years?
8. What type of star will the sun be after it shrinks to the size of Earth?
9. What are the violent eruptions called that the sun may throw off when it changes what type of star it is?
10. What is a star that used up all of its heat called?

11. When the sun has used up all of its heat, what will happen to any atmosphere that remains?

OKEECHOBEE:

Lake Okeechobee locally referred to as The Lake or The Big O, is a freshwater lake in the U.S. state of Florida. It is the second-largest freshwater lake wholly within the continental United States (after Lake Michigan) and the largest in the southern United States. Okeechobee covers 730 square miles (1,890 km²), approximately half the size of the state of Rhode Island.

Okeechobee is thought to have been formed out of the ocean about 6,000 years ago when the waters receded. At its capacity, the lake holds 1 trillion gallons of water and is the headwaters of the Everglades.

The name Okeechobee comes from the Hitchiti words oki (water) and chubi (big), and means "big water". It was previously called Macaco and Mayaimi, the latter the origin of the name of the city Miami, Florida by way of the Miami River. The floor of the lake is a limestone basin, and the lake varies in depth from 1 to 13 feet (0.3 to 4 m). Its water is somewhat murky from nutrient-enriched runoff from surrounding farmlands. The surface is above sea level. The lake is enclosed by a 20-foot (6 m) high dike built by the U.S. Army Corps of Engineers after a hurricane in 1928 breached the old dike, flooding surrounding communities and claiming thousands of lives. There are several inflows, including Taylor Creek and the Kissimmee River, and several small outlets, such as the Miami River, the New River on the east, and the Caloosahatchee Rover (via the Caloosahatchee Canal and Lake Hicpochee) on the southwest.

1. What is one of the two names that Lake Okeechobee is locally called by?
2. Where does Lake Okeechobee rank in terms of largest lakes in the United States?
3. How many square miles does Lake Okeechobee cover?
4. Lake Okeechobee is about half the size of what?
5. About how long ago was Lake Okeechobee thought to have been formed?
6. At its capacity, how many gallons of water does Lake Okeechobee hold?
7. Lake Okeechobee is the headwaters of what?
8. What is the English translation of Okeechobee?
9. The floor of Lake Okeechobee is what type of basin?
10. Who built the dike that encloses Lake Okeechobee?
11. What year was the dike built?

Passage and Questions for Experiment 2

The atmosphere of Earth is a layer of gases surrounding the planet earth that is retained by Earth's gravity. The atmosphere protects life on Earth by absorbing ultraviolet solar radiation, warming the surface through heat retention (greenhouse effect), and reducing temperature extremes between day and night. Dry air contains roughly (by volume) 78% nitrogen, 21% oxygen, 0.93% argon, 0.038% carbon dioxide, and small amounts of other gases. Air also contains a variable amount of water vapor, on average around 1%. The atmosphere has a mass of about five quintillion kg, three quarters of which is within about 11 km of the surface. The atmosphere becomes thinner and thinner with increasing altitude, with no definite boundary between the atmosphere and outer space. An altitude of 120 km (75 mi) is where atmospheric effects become noticeable during atmospheric reentry of spacecraft. The Karman line, at 100 km (62 mi), also is often regarded as the boundary between atmosphere and outer space.

Air is mainly composed of nitrogen, oxygen, and argon, which together constitute the major gases of the atmosphere. The remaining gases are often referred to as trace gases, among which are the greenhouse gases such as water vapor, carbon dioxide, methane, nitrous oxide, and ozone. Filtered air includes traces amount of many other chemical compounds. Many natural substances may be present in tiny amounts in an unfiltered air sample, including dust, pollen, and spores, sea spray, volcanic ash, and meteoroids. Various industrial pollutants also may be present, such as chlorine (elementary or in compounds), fluorine compounds, element mercury, and sulfur compounds such as sulfur dioxide.

1. The Earth's atmosphere is a layer of what?
2. What feature of Earth retains Earth's atmosphere?
3. The atmosphere protects Earth by absorbing what?
4. The warming of the Earth's surface through heat retention is called what?
5. Dry air contains roughly 0.93% of what?
6. On average, water vapor makes up what percentage of air?
7. What fraction of the atmosphere is within 11 km of the surface?
8. What happens to the atmosphere with increasing altitude?
9. What is the boundary between atmosphere and outer space often called?
10. Non-major gases in the atmosphere are often referred to as what?

Passage and Questions for Experiments 3 and 4

See Kang, McDermott, & Roediger, 2007

Passage and Questions for Experiments 5

Article: Garry & Polaschek, 2000
Questions:

There are two main mechanisms that have been proposed to account for the boost in confidence of having experienced an imagined counterfactual event. One is source confusion, the other is ______________.

When a person is more confident they experienced imagined counterfactual events than nonimagined counterfactuals, it is known as_____.

Garry et al. (1996) refuted the notion that the imagination-inflation effect is merely the phenomenon of_______.

Given the size of the imagination inflation effect, the authors suggest that future research look toward _______ the size of the imagination inflation effect.

_____predict(s) greater imagination inflation for long-ago imagined events.

According to Heaps and Nash (1999), which of the following factors predicts people's tendency to become more confident that they have actually experienced an event after imagining it?

Why do the findings of memory-related effects of imagination have clinical implications?

Loftus (1993) was the first systematic study to show what?

Passage for Experiment 6

See Lehman, Schraw, McCrudden, and Hartley (2007)