Educ Psychol Rev (2016) 28:171–196
DOI 10.1007/s10648-014-9290-2

RESEARCH INTO PRACTICE

Does Writing Summaries Improve Memory for Text?

Arie S. Spirgel · Peter F. Delaney

Published online: 25 October 2014

Abstract In five experiments, we consistently found that items included in summaries were better remembered than items omitted from summaries. We did not, however, find evidence that summary writing was better than merely restudying the text. These patterns held with shorter and longer texts, when the text was present or absent during the summary writing, with both short answer and multiple choice criterion tests, with a brief delay prior to the final test or with a several day delay, and regardless of whether the summary was written immediately after reading the text or after a short time away from the text. We additionally found evidence that writing a summary sometimes helped participants estimate how much they learned from the text. However, it seems that students do not write effective summaries because they are quite poor at picking out the important points from the text.

Keywords Memory · Writing to learn · Metamemory · Text memory

The written summary, which is a condensation of a topic's main points (Brown et al. 1983; Hidi and Anderson 1986; Winograd 1984), has received extensive praise for its effects on comprehension and memory. 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). It is thought that summarizing results in deeper understanding than rereading a text because it requires integration of new knowledge with existing knowledge (Wade-Stein and Kintsch 2004). Others have argued that summarizing improves retention and comprehension because writers must consider the entire passage and determine its important elements (Friend 2002; Radmacher and Latosi-Sawin 1995). Some have gone as far as to say there is no better way to promote long-term retention than to write a summary (Westby et al. 2010).

The effective operations that the typical summarizer is thought to use lends credence to these bold claims. Different researchers describe the cognitive operations that summarizers execute using different terminology, but Hidi and Anderson (1986) identified three processes commonly described by proponents of summarizing. First, summarizers supposedly make decisions about what to include and what to delete from an original source. Second,

A. S. Spirgel (✉) · P. F. Delaney (✉)
Department of Psychology, University of North Carolina at Greensboro, PO Box 26170, Greensboro, NC 27402-6170, USA
e-mail: aspirgel@nova.edu
e-mail: p_delane@uncg.edu

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 and McCaleb 1985). Taken together, these transformative processes should allow a summarizer to pick out what is important in the text, understand it, and create a representation of the information that should be memorable.

In fact, the typical summarizer bears little resemblance to the idealized version as described above. A review of the literature on how people summarize leads to a clear conclusion: summary writing is usually a passive and nonreflective process. Summarizing is characterized by the copy-delete strategy, meaning people copy some information from a text and delete other information, but do not make any meaningful alterations (Brown and Day 1983). Sherrard (1986) labeled summarizers in her descriptive study as mechanical, reporting that 82 % of all sentences participants wrote were created by one-to-one source-to-summary mapping and the deletion of source material. Similar summary writing habits were described 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 created sentences whose surface structure was difficult to tie to the source text about 10 % of the time. Garner and McCaleb (1985) replicated this finding, reporting that summaries written by participants in their study scored poorly on a measure of integration and synthesis. Further suggesting that participants infrequently implement rules of summarization, Brown and Day (1983) found that when a text contained no topic sentence, participants invented one in their summaries only about half of the time. Along those lines, Bednall and Kehoe (2011, experiment 2) found that even when participants were given the equivalent of a topic sentence—the definition of the topic they were tasked with summarizing—they include the correct version of it in their summaries only 64 % of the time. These findings converge with a verbal protocol study in which summary writers were less likely than analytic writers to focus on their knowledge of the topic and the task, reflected less on what was being written about, and spent less time planning (Durst 1989). Instead, summarizers were more likely to devote their time to paraphrasing the original text and displayed more concern for surface elements of the task (e.g., spelling, style) rather than searching for meaningful features of the text. All of this is consistent with Penrose's (1992) conclusion that summary writing can be completed without much thought or reflection.

Such conflicting perspectives—on the one hand that writers follow a strict set of rules to achieve their summary goals and on the other that the process is passive—are recapitulated in the literature evaluating the effects of summary writing on learning, which has also produced mixed results. This state of affairs led reviewers of the literature to the conclusion that summary writing has low utility as a memory technique (Dunlosky et al. 2012). Despite this, our perspective is that prior research often lacks rigorous methodology, obscuring any benefits of summarizing that might exist. Thus, the primary goal of the current experiments was to overcome these methodological concerns, which we aimed to achieve via two avenues. First, with the exception of experiment 2, we used a control group that reflects the most typical strategy students utilize when learning texts—rereading (Karpicke et al. 2009)—which we compared to summarizing under a variety of conditions. Second, recognizing that not all students write summaries optimal for learning, the emphasis of our analyses was on the effect on memory of what students *did* write. That is, even if students mainly copy and delete and ignore many of the central ideas of a text in their summaries, what they do in fact write about should be remembered better than what they do not include. Although not a perfect analogy,

this is akin to the generation effect, which is the finding that generating a stimulus results in better memory than reading the stimulus (Bertsch et al. 2007). The following review documents how this consideration has been largely ignored, while also revealing the inconsistent nature of the literature, both methodologically and in terms of the results.

Summarizing as a Memory Strategy: Inconsistent Results

Several studies have compared classes that summarize to traditional instruction. Horton et al. (1985) assigned students in a general chemistry class to summarize eight lectures. Compared to students in the same class who did not write summaries, the treatment group scored significantly higher on a multiple choice exam measuring comprehension of the lectures and problem-solving. Extending these procedures, Radmacher and Latosi-Sawin (1995) had students summarize part of a text and then rank selected essays based on how much revision they needed, as well as justify the rankings; compared to a control class from the prior semester, the writing class scored higher on a final exam. Carroll (2008), who also found a benefit of summarizing, began classes with an explanation of a famous quote related to that day's lecture. 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. Other studies have not found any benefit of writing compared to a baseline control group. Day (1994) randomly assigned students in two classes to summarize lectures or to serve as an attendance group and found no differences on four multiple choice/essay exams between the two groups. Altogether, these results provide tenuous evidence that writing a summary may have some benefits relative to traditional instruction, although widely varying procedures make it difficult to draw conclusions.

Studies that have included a restudying condition—which at least two have done—offer a more reasonable control group. Penrose (1992) had participants read a text and gave them up to 1 h to either write an informative essay about the key concepts or to restudy the key concepts. After each task, participants were asked to recall as much they could from the text. Even though the writing group used 14 more minutes of the 1 h period than the study group, they recalled less of the text. Producing different findings, Davis and Hult (1997) showed participants a video on the development of language that included three 7 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 4 min pauses after each segment; those in the pause group took notes during the video and reviewed the notes during the same 4 min pauses; a control group took notes throughout the video without pause. Although there were no differences among the groups on an immediate multiple choice test, 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. (It was not reported how the summary group compared to the restudy group alone).

One of the most stringent comparison groups included in summarization studies is a generate-question condition. Studies that include generate-question conditions typically have students create and answer questions about a text. For example, King et al. (1984) trained participants 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 h 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. In a similar study, King (1992) trained students how to generate and answer questions about a lecture and also how to write a summary of a lecture. On a multiple choice retention test given immediately after viewing the lecture while implementing their assigned strategy, 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 a note-taking with review group. On a retention 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.

In terms of its effect on memory, this review demonstrates that students who write a summary do not consistently outperform students assigned to traditional instruction, restudying, or generating questions. In addition to the inconsistent methods used in the writing and memory literature, we contend that the capricious results in the field are at least in part due to the variation in students' ability to identify and subsequently write about the important elements of a text. Like the generation effect—in which participants who produce an item remember it better than reading it—any benefits of writing may be limited to what participants write about. At least two experiments have addressed this possibility.

Dyer et al. (1979) had students read a text, after which some wrote a summary of the text. Either 30 min or 1 week later, participants answered questions about the text that could either be answered in one or two words and required direct factual knowledge, or questions that could be answered in 2 to 15 words and required the combination of two or more elements from the text. Even though participants who wrote a summary did not differ overall on the final test from participants who instead engaged in a spatial relations task, conditional probabilities indicated that items that were included in summaries were more likely to be correctly answered on the final test. This indicates that despite any overall differences, including an item in a summary is associated with increased retention for that item. Another study also provided support for this idea. In the relevant conditions in an experiment by Bednall and Kehoe (2011, experiment 2), participants read about six logical fallacies and either wrote a summary of each one or wrote a summary and provided an explanation generation. Although there was no overall benefit of summary writing relative to the other groups, in both writing conditions, the number of correct definitions participants included in their summaries was positively correlated with scores on all three final tests: a recognition test, a cued recall test, and an application test. These results further support the hypothesis that as the amount of relevant information in students' summaries increases, so should how much they remember from the text.

Rationale for the Current Studies

In the current experiments, we were primarily interested in whether including an item in a summary is associated with an increased likelihood in remembering that item, as well as in drawing firmer conclusions on the general benefits of (or lack of) summarizing relative to other memory strategies. The current experiments employed restudy control conditions—which are effective given that rereading texts offers little educational benefit (Callender and McDaniel 2009)—and a variety of different text lengths, retention intervals between study and the final test, and types of final test. Taken together, the five experiments provide a much stronger test of the benefits of summarizing on memory than previous research.

The various texts we selected are of varying length and topic, and most of them have been used in previous studies examining memory in college students. Hence, we know that they are age- and ability-appropriate for typical college students. Likewise, some of the questions were taken from earlier published articles that obtained memory advantages for the conditions they predicted they would (Kang et al. 2007; Roediger and Karpicke 2006b). When the questions were written by us, we tried to use questions similar to earlier studies on memory for text. All of the questions we used could be answered with literal overlapping from the text (i.e., inference or transformation were not required). An overview of the experiments is shown in Table 1, including the length of the text, the retention interval, and the type of questions on the final test. Experiment 1 compared writing a summary, restudying text, and doing nothing. Experiment 2 compared underlining and summarizing. Experiment 3 tested whether increasing the delay between reading the text and writing the summary would lead to more benefits of summary writing. Experiment 4 included both a delay between reading the text and writing as well as including a longer retention interval prior to the final test. In experiment 5, we again gave the final test after a long retention interval and varied whether people wrote their summaries with the text in front of them (as in our earlier studies) or whether they had to write the summary from memory. To preview the results, although participants' inclusion of items in summaries generally improved retention for those items, we were unable to find any retention advantage of summary writing over restudying, despite using a variety of different types of texts and manipulations. The apparently paradoxical result occurred because people included relatively few of the important points, yet summary writing helped them recall what they had written.

In addition to the main hypothesis that memory would be improved by including something in a summary relative to not including it, we took the opportunity to test a few other claims that might have a bearing on summary writing's effectiveness as a memory strategy. First, some authors have suggested that writing a summary would serve as a kind of feedback to students that would allow them to better gauge how well they would do on a future test (Anderson and Thiede 2008; Thiede and Anderson 2003). Except in experiment 3, we asked participants to predict how well they would do on the final test and examined how effectively they were able

Table 1 Overview of experiments

Experiment	Conditions	No. of words in text	RI	Type of test	Final test results
1	Text-present, restudy, do nothing	2500	Immediate	Short answer	Summary > control; restudy = control
2	Underline vs. summary	2500	Immediate	Multiple choice	Summary = underline
3	2 (immediate vs. delay) × 2 (restudy vs. summary)	256–258	Immediate	Short answer	Immediate = delay; restudy = summary
4	2 (immediate vs. delay) × 2 (restudy vs. write) × 2 (2 m vs. 48 h)	264	2 m or 48 h	Short answer	Immediate = delay; restudy = summary; 2 m > 48 h
5	3 (restudy vs. text-present summary vs. text-absent summary)	2500	7 days	Multiple choice	Text-present > text-absent; restudy = text-present; re-study = text-absent

to predict their final score. Second, we have often heard educators suggest that even if summary writing does not enhance memory directly, it makes students more interested and engaged in the topic, which would increase their future motivation to study. We thought this claim to be important enough to start including some rudimentary tests in the form of questionnaires about enjoyment and interest in experiments 2 and 5.

Experiment 1

Our first experiment was a simple test of whether writing a summary is better than doing nothing. We instructed participants to read a psychology article and then to write a summary of the article, restudy the article, or work on something else. A final short answer test was then administered.

Method

Participants A total of 68 undergraduate psychology students participated for course credit. They were tested in groups of four people or fewer with an experimenter present. No one participated in any of our experiments more than once.

Materials The text was an article about imagination inflation from *Current Directions in Psychological Science* (Garry and Polascek 2000) with approximately 2500 words. A word search that contained one-word movie titles was used as the distractor task. The final test was created by having the first author, another psychology graduate student, and a senior undergraduate student independently select the ten most important idea units in the text. The eight-item short answer test was created from the six items everyone agreed on, plus two items that two out of three raters agreed on (see Appendix).

Procedure Participants were randomly assigned to one of the conditions. Participants had 15 min to read the article. Participants who finished early were instructed to reread the text until time elapsed. Then, participants 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 conditions had 20 min to work and were told to keep reviewing the materials if they finished early. Participants in the *control group* worked on a word search for 20 min instead of summarizing or rewriting.

Participants were then asked to indicate what percentage of questions they thought they would answer correctly on the final test and received 5 min to work on the word search. They then received a paper with the final test questions and could use as much of the remaining hour as needed to complete the questions.

Results and Discussion

Participants who scored 2.5 SD above or below the mean for their group on the final test were removed from the analyses. Two participants met this criterion, both from the control

condition. In each experiment, we conducted *inclusion analyses* to determine if including an item in a summary results in increased retention relative to omitted items.

Inclusion Analyses For the following analyses and for all subsequent experiments, *idea unit* was defined as including text in a summary that could be used to answer a question on the final test. For example, for experiment 4, if a participant wrote "the Earth's atmosphere is made up of a bunch chemicals," they would not get credit because it does not include the information needed to answer the question from the final test, "The Earth's atmosphere is a layer of what?" If instead, the participant replaced "chemicals" with "gases," s/he would get credit. We did not give credit to participants who by chance included lures from the multiple choice tests in their summaries, because we excepted that when participants did this it would reduce the likelihood of correctly answering that question (Fazio et al. 2010).

For this experiment, both authors independently scored each statement, with 91.4 % agreement (139/152). For this and each subsequent experiment, when scoring disagreements occurred, we defaulted to the first author's scoring. In the current experiment, summary writers included 25 % of the idea units from the final test in their summaries—a rather low proportion.

To examine the effects of including an item in a summary on retention, a paired samples *t* test was conducted on the proportion of questions recalled for items included in summaries and those that were not. Included items (M=.64, SE=.078) were more often remembered than non-included items (M=.16, SE=.036), $t(17)=5.82$, $p<.001$, $d=1.89$.

Final Test This analysis tested the effect of condition on retention. The ANOVA on the proportion of final test questions answered correctly, with condition as the between-subjects variable, was only marginally significant, $F(2, 65)=2.41$, MSE=.035, $p=.098$, $\eta_p^2=.069$. However, post hoc independent samples *t* tests suggested that the summary group (M=.29, SE=.035) scored higher on the final test than the control group (M=.20, SE=.024), $t(44)=2.27$, $p=.028$, $d=.67$. Participants in the restudy condition scored marginally higher than the control group (M=.31, SE=.054), $t(45)=1.93$, $p=.060$, $d=.56$. The summary and restudy groups did not differ, $t<1$. Thus, although writing was better than doing nothing, it was no better than the restudy group.

On the one hand, summarizing did not improve memory relative to restudying. The writing data, however, show that including an idea unit in a summary was 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. This *inclusion effect* was obtained in each of the subsequent experiments as well.

Predicted Recall The predictions across the groups did not differ, $F(2, 61)=1.11$, MS=391.7, $p=.337$. The percent of items participants predicted they would recall was 63 (SD=17) for the summary group, 65 (SD=25) for the restudy group, and 71 (SD=15) in the control group. Predictions did not correlate with final test recall in any of the conditions, and overall was $r(62)=.10$, $p=.439$.

Writing apparently did not benefit metacognitive accuracy in this experiment. To preview, subsequent experiments replicate the finding that writing summaries after a delay improves metacognitive accuracy.

Although there were no differences on the final test between the summary and restudy groups on the final test, including an item a summary produced robust effects on memory. This does not lead to the automatic conclusion that writing an item results in better retention than restudying that same item, as it is unknown if participants were actively restudying when they

were supposed to be. Rather, some participants may have been daydreaming when they appeared to be restudying, or they may have had a tendency to restudy items that summarizers tended to ignore.

Experiment 2

Experiment 2 selected a different control group: underlining. It seemed clear from experiment 1 that information that was included in the summary was recalled better than information that was not included in the summary. Asking participants to underline what they would include in a summary would allow us 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 writing an item, relative to simply reading and underlining it. 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.

Importantly in terms of our goals for the current experiment, despite some exceptions (e.g., Fass and Schumacher 1978; Johnson 1988), underlining has typically not been found to provide a text retention benefit beyond reading without underlining (e.g., Nist and Hogrebe 1987; Silvers and Kreiner 1997), and a recent review of the topic found it to be of little educational utility (Dunlosky et al. 2012). The primary goal of the current experiment was 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 were included in summaries.

Method

Participants A total of 46 undergraduate psychology students participated in this experiment for course credit. They were tested individually with an experimenter present.

Materials The text was an approximately 2500 word article about literacy from *Current Directions in Psychological Science* (Treiman 2000). Kang et al. (2007), who originally used these materials, also created eight multiple choice questions that were used here as the final test. Two-digit by two-digit multiplication problems were included as a distractor. A short questionnaire was created that asked participants to indicate how interesting they found the material and how much they enjoyed participating in the experiment on a 1 (strongly disagree) to 7 (strongly agree) scale. It also asked that students provide an aggregate judgment of learning (JOL) by indicating what percentage of questions they thought they would get right on the final test.

Procedure Participants were randomly assigned to one of the conditions and were given up to 20 min to read the text. Participants in the *summary* condition were given the same instructions as the summary participants in the previous experiment. Participants in the *underline* condition were given the same instructions, 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).

After the 20 min was up, summarizers and underliners were given 12 min to complete their respective tasks and were told to continue looking over the materials if they finished before time was up. They were then given the questionnaire on interest, enjoyment, and aggregate JOL, which was followed by the math distractor activity for 5 min. Lastly, participants were given the eight multiple choice questions, debriefed, and dismissed.

Results and Discussion

Inclusion Analyses Idea unit for the summary group was defined here the same way as in experiment 1. The first author scored all of the statements, and the second author scored a subset of them; for that subset, we agreed on 87.5 % of them (91/104).

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), $t(43)=3.52$, $p<.001$, d=1.06. This amounts to including, respectively, 28.8 and 13.0 % of the eight potential idea units. Our next question was whether including an item in a summary would benefit memory more than underlining an item. A 2 inclusion (included vs. not included) × 2 condition (underline vs. summary) mixed ANOVA was conducted on the proportion of items remembered on the final test. Ten participants who did not include any of the final items were not analyzed.

The results are displayed in Fig. 1. There was a main effect of inclusion, $F(1, 34)=8.26$, MSE=.046, p=.007, η_p^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, $F(1, 34)=2.51$, MSE=.087, p=.122, η_p^2=.069. However, there was an inclusion × condition interaction, $F(1, 34)=5.89$, MSE=.046, p=.021, η_p^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, d=1.50, 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), $t<1$.

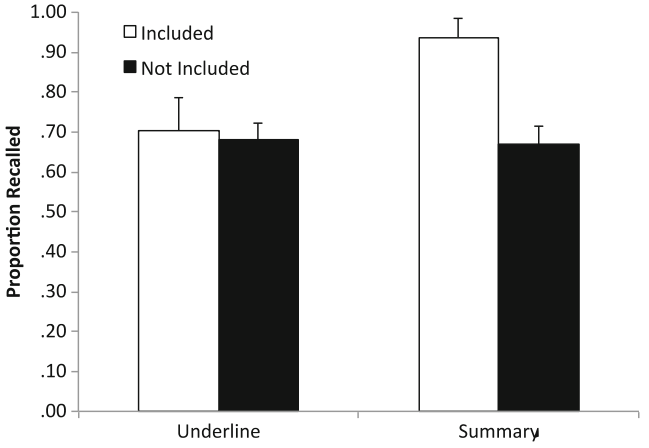

Fig. 1 Probability of answering a question correctly as a function of whether or not its answer was included during summarizing or underlining, experiment 2. *Error bars* represent ±SE

After we observed in experiment 1 that including an item in a summary was associated with a large increase in later remembering that item, our goal for the current experiment was to determine if that increase was due to the act of summarizing, or the extra attention allotted to these items. To answer this question, we measured whether including an item in a summary had a bigger effect on retention than underlining an item; although it did, this finding must be interpreted with caution, as participants in the underline condition underlined more than those in the summary group wrote. Thus, the proportion of included idea units that were later recalled was higher in the summary group because they included fewer idea units than the underline group.

Final Test An independent *t* test comparing the proportion of questions answered correctly found no difference between the underline condition (M=.70, SE=.044) and the summary condition (M=.66, SE=.031), $t<1$. This finding is at odds with the belief that summarizing is a valuable strategy for memorizing a text, particularly when compared to underlining, which is considered a relatively inert study strategy (Silvers and Kreiner 1997). Subsequent experiments will include manipulations designed to maximize the benefit of summarizing.

Interest, Enjoyment, and Predictions After not observing any memory differences between the summary and restudy group in experiment 1, we gathered interest and enjoyment assessments here to see if there was a secondary effect of summarizing that could ultimately lead to better retention, for example by keeping students on task longer. We found no evidence for this possibility, nor did we in any of the subsequent experiments in which we gathered these data. There were no differences in interest in the topic between the underline condition (M=5.13, SE=.25) and the summary group (M=5.00, SE=.25), $t<1$. There were no differences in how much participants in the underline condition (M=4.91, SE=.21) and the summary condition (M=5.10, SE=.28) enjoyed participating in the experiment, $t<1$. No differences were detected between the predictions in the underline condition (M=69.35, SE=3.21) and the summary condition (M=66.65, SE=2.94), $t<1$.

Predictions neither correlated with recall in the underline condition, $r(23)=-.19$, $p=.391$, nor the summary condition, $r(23)=-.14$, $p=.539$. This finding is consistent with experiment 1, as well as other research indicating that a significant relationship between predictions and actual performance is observed when summaries are written after a delay, but not when written immediately (as they were here).

Experiment 3

For the third experiment, we asked whether increasing the gap between reading the text and writing about it would help. Making a test more difficult often enhances the effect of the test (e.g., Carpenter and DeLosh 2006; Pyc and Rawson 2009), so we thought that perhaps giving people time to forget some of the text they had read before summarizing it would enhance the value of summarizing relative to restudying. Our reasoning was that if participants are retrieving parts of the text while writing, the delay might enhance the impact of the summary writing through more difficult retrieval. Another idea was that perhaps delaying the summary writing would allow participants time to understand the text better, and therefore, they would be better able to pick out main points after the delay. Alternatively, one might find that participants just look back at the text again and show no benefit beyond rereading.

Method

Participants A total of 88 undergraduate psychology students participated for course credit. They were tested in groups of up to four people with an experimenter present.

Materials Two text passages, one about the sun (256 words) and another about Lake Okeechobee (258 words), were used. 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). In addition, we created 11 short answer questions testing memory for each text. In pilot data, participants correctly answered 55 % of each set of questions (see Appendix). The same word search that was used in experiment 1 was again used as a distractor.

Design Whether participants restudied or wrote summaries was varied between subjects. Whether the restudying or summarizing was completed immediately or delayed was varied within subjects, with one of the passages being used for each condition. The presentation order of the passage and whether participants underwent massed or spaced procedures first was counterbalanced.

Procedure Participants were randomly assigned to one of the conditions. They began by reading one of the two text passages at their own pace. In the *delayed condition*, they worked on the word search for 8 min before writing or restudying. In the *immediate condition*, they immediately began writing or restudying the passage.

The *restudy group* was 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." The *summary group* was 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, participants from the immediate condition subsequently completed the delayed condition with the other text, and vice versa.

Results and Discussion

Inclusion Analyses The first author scored all of the statements, and the second author scored a subset of them; for that subset, we agreed on 87.7 % of them (251/286).

A paired samples t test showed that the immediate condition did not result in including a different number of idea units (M=5.07, SE=.37) than the delayed condition (M=4.84, SE=.35), t<1. This amounts to including, respectively, 46.1 and 44.0 % of the potential 11 idea units.

Next, we 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) × 2 delay (immediate vs. delayed) mixed 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 is associated with an increased probability of later recalling that item and whether that probability varies by delay. These data are summarized in Fig. 2. There was a main effect of inclusion, $F(1, 38)=74.32$, MSE=.052, $p<.001$, $\eta_p^2=.662$,

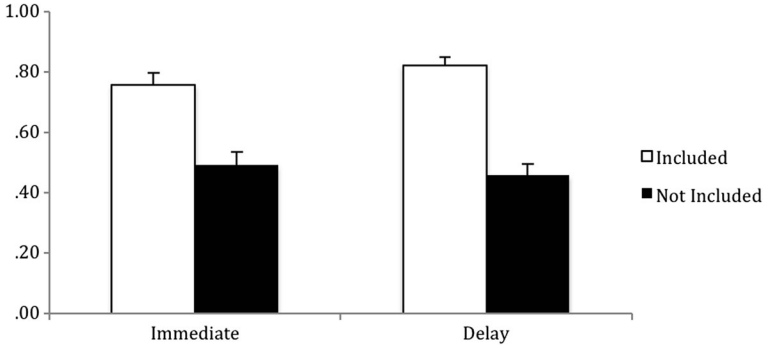

Fig. 2 Probability of answering a question correctly as a function of whether or not its answer was included during summarizing, experiment 3. *Error bars* represent ±SE

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 delay, $F<1$, nor a delay × inclusion interaction, $F(1, 38)=2.23$, MSE=.041, p=.144, η_p^2=.055. Replicating the prior experiments, participants were more likely to remember items that they previously included in their summaries than items they did not.

Final Test A 2 delay (immediate vs. delayed) × 2 task (summary vs. restudy) ANOVA on the proportion of questions answered correctly was conducted, with repeated measures on the former factor. Figure 3 summarizes these results. There was neither a main effect of delay, $F<1$, nor a delay × task interaction, $F<1$. There was a marginal main effect of task, $F(1, 86)=3.90$, MSE=7.56, p=.052, η_p^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).

The final test results resulted in the third failure to find any benefit of summarizing over restudying. Rather, participants who restudied tended to answer more final test questions correctly than those who summarized, a finding that directly opposes claims favoring summary writing. This result is consistent with Penrose's (1992) finding, who also reported that after a short retention interval students who restudy recall more than those who write; however, Davis and Hult (1997) found that summarizing leads to better retention than restudying, but only

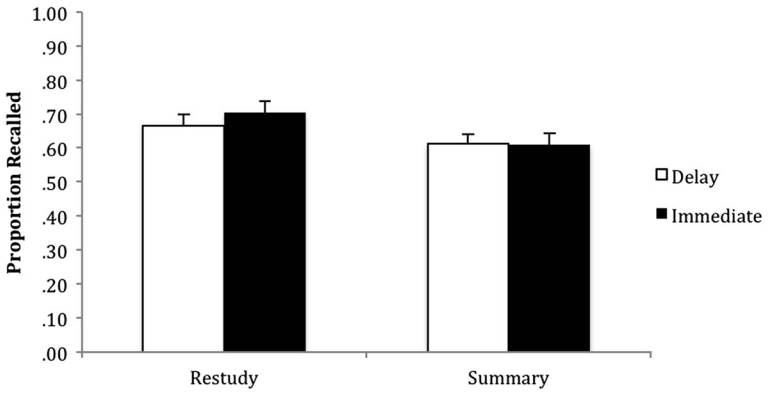

Fig. 3 Proportion of final test questions answered correctly, experiment 3. *Error bars* represent ±SE

after an extended retention interval. The possibility that the benefits of writing only emerge after a delayed retention interval will be investigated in the next experiment.

Experiment 4

The first three experiments produced reliable evidence that including a given item in a summary benefits retention compared to not including that item; however, we were yet to find a benefit of summarizing compared to restudying. The failure to find a benefit of summarizing relative to restudying in experiment 3 may have been due to the short retention interval. Testing effects, for example, usually are much larger on a delayed test than on an immediate one (e.g., Congleton and Rajaram 2012; Roediger et al. 2011; Toppino and Cohen 2009; Wheeler et al. 2003). Therefore, the current experiment extended the retention interval out to 48 h. As in the previous experiment, people were either asked to write or restudy immediately after finishing reading the text, or there was a short break between reading the text and the writing or restudying task.

Method

Participants and Design A total of 152 undergraduate psychology students 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 task (restudy vs. summary) × delay (immediate vs. delayed) × retention interval (2 m vs. 48 h). All variables were manipulated between subjects. Participants were tested in groups of up to four people with an experimenter present.

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 ten short answer questions that we created. In pilot testing, participants answered on average 58 % of the questions correctly (see Appendix). The same word search that was used in earlier experiments served as the distractor for the current experiment.

Procedure Participants were randomly assigned to one of the conditions. All participants read the text at their own pace. Some participants then received the restudy or summary condition instructions immediately (*immediate condition*), and others did so after 13 min of the word search (*delay condition*). Participants 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 in the 2 min retention interval condition were given the final test. Those in the 48 h 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 and Discussion

Inclusion Analyses The first author scored all of the statements, and the second author scored a subset of them; for that subset, we agreed on 90 % of them (126/140).

An independent samples t test demonstrated that more idea units were included in the immediate condition (M=3.94, SE=.27) than in the delay condition (M=3.06, SE=.26), $t(68)$=2.37, p=.020, d=.57. This amounts to including, respectively, 39.4 and 30.6 % of the potential ten idea units.

To determine the combined effect of including an idea unit in a summary and delay, a 2 inclusion (included vs. not included) × 2 delay (immediate vs. delayed) ANOVA was conducted on the proportion of recalled idea units included vs. not included in summaries. These results are displayed in Fig. 4. 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, η_p^2=.605. There was neither a main effect of delay, $F(1, 64)$= 2.39, p=.127, η_p^2=.036, nor an interaction, F<1.

Despite including a longer retention interval, we did not find different levels of retention as a function of delay. The finding that immediate summaries included more idea units than delayed summaries suggests the delay may have been too long, and participants experienced relative difficulty reorienting to the text and including idea units from it.

Final Test To test the effects of delay, task, and retention interval on retention, we ran a 2 task (restudy vs. summary) × 2 delay (immediate vs. delay) × 2 retention interval (2 min vs. 48 h) ANOVA on the proportion of final test questions answered correctly. These results are displayed in Fig. 5. There was no main effect of delay, $F(1, 133)$=2.68, MSE=.049, p=.104, η_p^2=.020, but a trend for those in the immediate condition (M=.51, SE=.026) to recall more than those in the delay 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, η_p^2=.114, as those in the 2 min condition (M=.56, SE=.025) recalled more than those in the 48 h 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, η_p^2=.018.

Despite an attempt in the current experiment to highlight the benefits of delaying summaries by including a longer retention interval, it produced similar retention for summarizers and

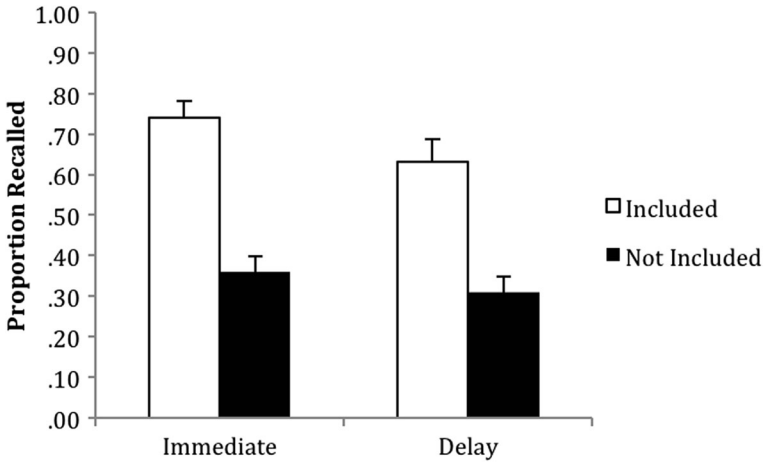

Fig. 4 Probability of answering a question correctly as a function of whether or not its answer was included during summarizing, experiment 4. *Error bars* represent ±SE

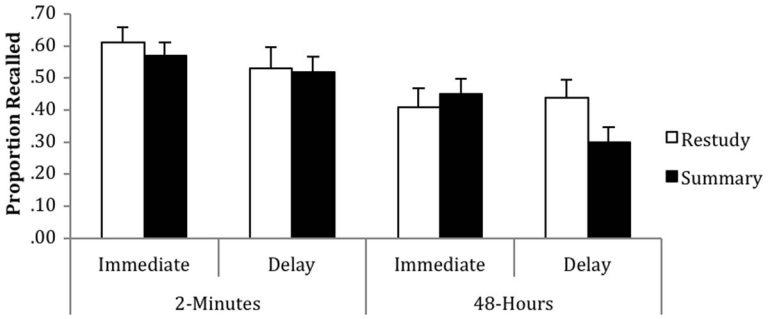

Fig. 5 Probability of answering a question correctly as a function of delay and task, experiment 4. *Error bars* represent ±SE

restudiers. Furthermore, there was no impact of giving people a delay between reading the text and the summary writing or restudy exercise.

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

Experiment 5

The studies presented so far all allowed people to write their summaries with the texts in front of them. However, this might facilitate a copy-and-delete summary writing strategy that results in minimal processing of the text. Perhaps if people were asked to write the summaries without access to the text, they would be forced to retrieve the most important information from the text and consequently show a larger benefit of summary writing. Essentially, we hoped to obtain an effect similar to the testing effect, by which a task that encourages retrieval of relevant material improves learning relative to restudy (for reviews, see Delaney et al. 2010; Roediger et al. 2011). Experiment 5 therefore forced some participants to write their summaries without the text, followed by immediate feedback in the form of being able to view the text again and update the summary. As the testing effect is often detected only after a few days, we gave the final test after a 7 day delay. Other experiments that have included test-absent conditions have done so only as part of investigating some other phenomenon; thus, they lack an adequate design or detail in their results to draw adequate conclusions on the effect on memory of writing without access to the text (Anderson and Thiede 2008; Dyer et al. 1979; Kirby and Pedwell 1991; Thiede and Anderson 2003). At least conceptually, prior research demonstrates that writing without access to the text should be an effective text retention strategy; experiments have shown that when reading a text, less frequent stopping to summarize results in better retention than more frequent stopping, presumably because the former requires greater retrieval effort (Foos 1995; Spurlin et al. 1988).

Method

Participants A total of 77 undergraduate psychology students participated for course credit. The 12 people who participated in session 1 but did not return for session 2 were eliminated

from the analyses. Participants were tested in groups of up to four people with an experimenter present.

Materials The text, which was about 2500 words, was obtained from *Current Directions in Psychological Science* (Anastasio et al. 1999), and the eight multiple choice questions used as the final test were created by Kang et al. (2007). The questions were presented in ten different randomized orders.

Procedure Participants were randomly assigned to one of the conditions. 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. Participants in the *restudy* condition were given 20 min to restudy the text. Those in the *text-present summary* condition were given 20 min to write a summary of the text with access to the text the entire time. Participants in the *text-absent summary* condition were given 13 min to write a summary of the passage, and then returned the text, and received 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 et al. 2007). The summary and restudy instructions were the same as those provided in the prior experiments.

All participants then completed 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 0 and 100 corresponding to what percentage of questions they thought would remember in a week. All participants were asked to return to the lab after a week and completed the multiple choice test then.

Results and Discussion

Inclusion Analyses The first author scored all of the statements, and the second author scored a subset of them; for that subset, we agreed on 93.2 % of them (97/104).

The text-present summary group included more idea units (M=3.48, SE=.27) than the text-absent summary group (M=1.78, SE=.18), $t(44)=5.31$, $p<.001$, $d=1.58$. This amounts to including, respectively, 43.5 and 22.2 % of the potential eight idea units.

To examine the effects of including an item in the summary and test conditions, a 2 inclusion (included vs. not included) × 2 condition (text-present summary vs. text-absent summary) mixed ANOVA was conducted on the probability of recalling an item. There was a main effect of inclusion, $F(1,40)=115.68$, MSE=.040, $p<.001$, $\eta_p^2=.743$, such that the probability of remembering an item that was included in a protocol (M=.91, SE=.030) was greater than the probability of remembering an item that was not included (M=.44, SE=.040). The main effect of condition was not significant, $F(1, 40)=1.40$, MSE=.063, $p=.243$, $\eta_p^2=.034$. There was no inclusion × condition interaction, $F<1$. Figure 6 summarizes these data.

The inclusion data replicate the previous four experiments: including an item in a summary is associated with an increase in the probability of later remembering that item. In the current experiment, this effect did not differ between conditions, providing further support for the point that the relatively low overall retention rate for participants in the text-absent summary group can be partly attributed to the low number of items they included in their summaries.

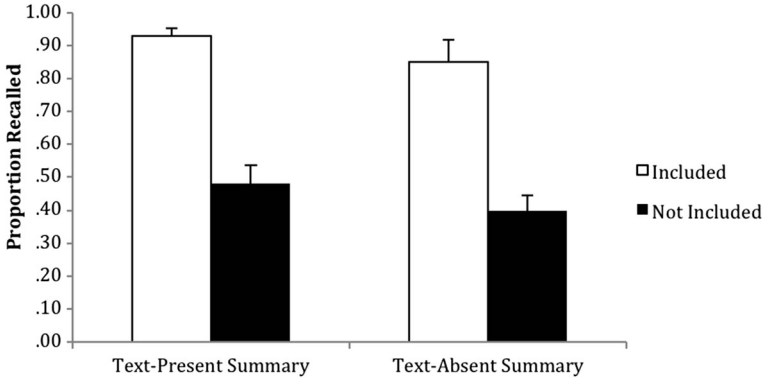

Fig. 6 Probability of answering a question correctly as a function of whether or not its answer was included during summarizing, experiment 5. *Error bars* represent ±SE

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 run, producing a main effect of condition, $F(2, 62)=4.06$, MSE=.035, $p=.022$, $\eta_p^2=.116$. Follow-up tests revealed that the text-present summary group ($M=.67$, SD=.18) answered more questions correctly than the text-absent summary group ($M=.51$, SD=.19), $t(44)=2.91$, $p=.006$, $d=.86$. The restudy group ($M=.60$, SD=.20) did not differ from the text-present summary group, $t(40)=1.21$, $p=.233$, $d=.37$, nor the text-absent summary group, $t(40)=1.46$, $p=.153$, $d=.45$.

We included a text-absent summary group in the current experiment as a way to counteract students' use of the copy-delete strategy. Instead, participants who wrote summaries without access to the text performed poorly on the final test, remembering less than those in the text-present group. As corroborated by the writing data, which demonstrated that the text-present summary group included almost twice as many idea units in their summary as the text-absent summary group, this suggests that participants experienced difficulty recalling information from the text to include in their summaries. Text-absent-like conditions typically result in better final test performance than test-present ones, but only when the text-absent conditions are easy enough to produce a sufficient number of successful retrievals (Kang et al. 2007, experiment 1). This experiment also failed to obtain any evidence that summary writing is a better method for remembering text than restudying, as participants in these groups performed identically.

Interest, Enjoyment, and Predictions The groups did not reliably differ in their interest in the topic, $F(2, 73)=1.38$, MSE=1.97, $p=.257$, $\eta_p^2=.037$. The means and standard deviations were as follows: restudy ($M=4.52$, SD=1.73), text-present ($M=5.04$, SD=.96), and text-absent ($M=5.16$, SD=1.49). There were also no differences in how much they enjoyed participating, $F(2, 73)=1.93$, MSE=1.50, $p=.153$, $\eta_p^2=.050$. The means and standard deviations were as follows: restudy ($M=4.26$, SD=1.32), text-present ($M=4.93$, SD=.94), and text-absent ($M=4.72$, SD=1.40). The percentage of questions they predicted they would remember on the final test also did not differ, $F(2, 73)=1.10$, MSE=432.77, $p=.340$, $\eta_p^2=.029$. The means and standard deviations were as follows: restudy ($M=59.2$, SD=22.7), text-present ($M=66.9$, SD=15.0), and text-absent ($M=60.0$, SD=24.3).

People's recall predictions were significantly related to memory only in the text-present summary condition, $r(23)=.42$, $p=.047$. For the restudy group, the correlation was $r(19)=$

−.10, p=.678, and for the text-absent summary group $r(23)$=−.23, p=.287. This result replicates the experiment 4 finding that final test predictions made after writing a delayed text-present summary (or after the natural delay created by a long text) are more predictive of final test performance than those made after restudying; this experiment also extends previous findings and demonstrates that the same effect observed here with text-present summaries does not apply to text-absent summaries.

General Discussion

Based on limited evidence, many have claimed that writing a summary captures some of the most effective processes for remembering texts. Thus, our primary goals for the current set of experiments were to explore whether including something has a beneficial effect on memory, and whether summary writing is a helpful memory strategy compared to restudying. The straightforward result is that participants were more likely to remember items they included in their summaries, but summary writing is not particularly helpful relative to restudying. Despite our repeated attempts to find evidence of the latter, our results indicated that at best summary writing was better than doing nothing (experiment 1), but it was often no different than restudying (experiments 1, 4, and 5), and in one case, it was worse (experiment 3). There was also no difference in retention between underlining—a relatively ineffective learning strategy—and summary writing (experiment 2). Even when some of the most effective manipulations for improving retention were introduced, summarizing was still no better than restudying (experiments 3–5).

Why Doesn't Summary Writing Improve Memory for Text?

Although the final test results of each of our experiments suggest that writing is not a useful way to remember text, a more precise conclusion is that writing is not useful when people do not know what to write about. When people write an optimal summary by including the important elements of a text, summarizing may be an effective strategy for remembering text. In each experiment, we found that people were more likely to remember items that they included in their summaries than items that they did not. Other research has also found that including items in a summary is associated with better retention of those items (e.g., Bednall and Kehoe 2011, experiment 2; Dyer et al. 1979). Thus, although summary writing, like retelling, may encourage selective rehearsal (Marsh 2007) and does not appear to facilitate retention of non-included items the way testing facilitates retention of nontested items (e.g., Chan 2009; Chan et al. 2006), a potential upside of summary writing is that it effectively aids in the retention of items that are included in summaries, even if participants mostly copy and delete.

Although we found improved memory for what people actually wrote about, to benefit from this in a practical, educational sense, summarizers must be able to identify the important elements of a text, something participants in our experiments struggled with. They seemingly lacked the ability to discriminate important from less important points, even though these texts were often used in earlier memory studies with college students. Tallying across the seven total conditions in experiments 1 through 5 in which participants wrote with access to the original text, their summaries included on average only 35.5 % of idea units that they would later be asked about on the final tests (23.43 out of 66).

Training students to write better summaries might then lead to improved memory for text. Studies with school children have managed to produce dramatic improvements in the qualities

of summary writing by providing students with teacher-led examples and with helpful feedback (approximately d=.8 in studies reviewed by Rosenshine and Meister 1994). Furthermore, improving summary writing has been a central effort of research in the "Summary Street" program, which used latent semantic analysis to provide automatic rapid feedback on student summaries (Wade-Stein and Kintsch 2004). Summary Street provides students with content coverage feedback as well as helps them to see where they are writing wordy or redundant information. This sort of feedback enabled sixth-grade students to enhance their summary writing skill, and summaries after practicing with it had more even content coverage of entire articles than early summaries, and the students seemed to maintain their gains even when feedback was removed in a subsequent session. Perhaps with such training, summary writing might make a comeback as a memory strategy. However, given the dismal lack of ability of college students to distinguish key concepts from salient but useless details, it seems unlikely to be very useful at this time. A recent review of the literature made a similar point, concluding that although summarizing can be an effective learning strategy for skilled summarizers, many learners lack the skills to benefit from it (Dunlosky et al. 2012).

A final point is that our results might at first blush seem to be inconsistent with recent research on the testing effect, which usually results in better memory for tested items than for restudied items, at least after a day or more of delay (for reviews, see Delaney et al. 2010; Roediger et al. 2011). However, our analysis comparing items that were in the summaries to the corresponding items when they were restudied suggested that indeed, included items benefitted from summary writing. Furthermore, it is not clear that the copy-delete strategy that was likely used by most of our participants is the same as receiving a memory test on the key points. Tests need to target the relevant material in order to be effective, and students are apparently poor at picking out that relevant material by themselves.

The Effect of Summary Writing on Interest, Enjoyment, and Metamemory

In our experiments, summary writing and restudying did not affect people's interest in the topic or their enjoyment in the procedures in different ways. We did obtain evidence that writing has a unique effect on people's metamemory such that under certain circumstances, it helps students predict how well they will perform on the final test. When participants wrote text-present summaries after a delay, or after a functional delay created by a long text, and were given a test after a long retention interval, there was a significant relationship between how well they predicted they would do on their final test and how well they actually did. Although it is an imperfect comparison given that we examined interindividual correlations and they examined intraindividual correlations, Anderson and Thiede (Anderson and Thiede 2008; Thiede and Anderson 2003) also found that writing delayed summaries, but not immediate ones, resulted in a reliable association between predictions and test performance. In their experiments, they only included an immediate test, whereas in our experiments, we only observed this association when the final tests were delayed. According to Thiede and Anderson (2003), predictions made after delayed summaries are based on the gist of the text rather than specific details, and because gist memories are more accessible after a long retention interval than specific memories (Reyna and Brainerd 1995), they should improve participants' sensitivity to performance on a delayed test.

Summarizing's effect on metamemory is a ripe avenue for future research. We did not obtain any evidence that summary writing is a more useful method than restudying, but it may offer a downstream effect not measured by our tests. For example, writing a delayed summary might help students pick out the important points from a text, allowing them to study more productively by biasing their time to items that are most likely to be included on a test.

Conclusion

Summary writing is consistently cited as an optimal text retention strategy, but in five experiments, we did not find evidence to support this hypothesis. It is possible that under the right conditions, writing a summary is as effective as it is claimed to be, but despite our best efforts, our data do not support this perspective. It is not clear what it will take for those who support writing as a learning tool to alter their perspective, because as Ackerman (1993, p. 361) wrote, due to "its ideological status, rhetoric and composition will not loosen their grip on writing as a mode of learning." This certainly appears to be the case, as the writing-to-learn movement has been going strong for approximately 50 years, yet 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 and Fowler 2004, p. 122). We hope that these experiments provoke summary writing supporters to run experiments designed to demonstrate its efficacy, as opposed to simply commending it without evidence.

Writing is a cornerstone of education and an important skill in its own right (Emig 1977). We do not mean to suggest that writing is useless for learning, or that writing is not something that should be taught. Indeed, there is a rich psychology of writing literature that explores how people write and ways of improving writing (e.g., Bereiter and Scardamalia 1993; Kellogg 1994; Kellogg and Raulerson 2007), and there is evidence that essay writing can facilitate comprehension of text (Newell 1984) and learning in school (for a meta-analysis, see Bangert-Drowns et al. 2004). So, even though learning to write effective summaries is useful in many contexts, the current evidence suggests that summary writing is not the most effective way to capitalize on writing for learning. Other kinds of writing, such as essays that promote knowledge transformation, are seemingly more helpful in promoting learning.

Likewise, there may yet be conditions under which summary writing is helpful. For example, we suspect that summarizing may be more useful when one has to bring together information from disparate sources into a single, coherent summary, as this would tend to promote the generation of new inferences that help to create a more coherent form of text (e.g., Kintsch 1998; Wiley and Voss 1999). However, we suggest that educators move away from the use of summaries of a single article, which takes a long time and does not promote learning any more than rereading does.

Appendix: Materials

Text for Experiment 1: Garry and Polascek (2000)

Questions for Experiment 1

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 _____.

A. passage of time
B. pretending to be someone else
C. vividness of memory
D. familiarity

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

A. the lost-in-the mall procedure
B. imagination inflation
C. hypnosis
D. false memory

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

A. regression towards the mean
B. a meaningless effect size
C. over confidence
D. age effects

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

A. questioning
B. decreasing
C. maintaining
D. increasing

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

A. Familiarity
B. Source confusion
C. Individual differences
D. A Heisenberg like suggestion

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?

A. Their susceptibility to influence of an authoritative person
B. The vividness of their mental imagery
C. Their predisposition to hypnotic suggestion
D. Their arousal to emotional stimuli

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

A. Because patients might be imagining their disorder/illness
B. Because various psychotherapy techniques involve imagined performance.
C. Because the therapist may find it difficult to distinguish reality from imagination
D. Because imagination of events can lead to people to believe they experienced anxiety-provoking events

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

A. Coherent, detailed events could be implanted by researchers
B. Emotional events tend to be particularly salient and memorable
C. Counterfactual thoughts can affect people's judgment of outcomes

Text for Experiment 2: Treiman (2000)

Questions for Experiment 2: See Kang et al. (2007)

Passage and Questions for Experiment 3

See Rogers (2001) for sun text.

What type of star is the sun today?

What fuels the sun?

How long has the sun been in its present state?

What can astronomers study to predict the rest of the sun's life?

What two major changes will happen to the core of the sun in about 5 billion years from now?

In about 5 billion years, about how many miles will the outer regions of the sun expand?

What type of star will the sun be in about 5 billion years?

What type of star will the sun be after it shrinks to the size of Earth?

What are the violent eruptions called that the sun may throw off when it changes what type of star it is?

What is a star that used up all of its heat called?

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 km^2), 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-feet (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.

What is one of the two names that Lake Okeechobee is locally called by?

Where does Lake Okeechobee rank in terms of largest lakes in the United States?

How many square miles does Lake Okeechobee cover?

Lake Okeechobee is about half the size of what?

About how long ago was Lake Okeechobee thought to have been formed?

At its capacity, how many gallons of water does Lake Okeechobee hold?

Lake Okeechobee is the headwaters of what?

What is the English translation of Okeechobee?

The floor of Lake Okeechobee is what type of basin?

Who built the dike that encloses Lake Okeechobee?
What year was the dike built?

Passage and Questions for Experiment 4

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.

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

Text for Experiment 5: Anastasio et al. (1999)

Questions for Experiment 5: see Kang et al. (2007)

Acknowledgments This research was conducted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of North Carolina Greensboro, and portions were previously presented at the 2011 Psychonomic Society Meeting. We thank dissertation committee members Robert Guttentag, Lili Sahakyan, Paul Silvia, and Dayna Touron for their guidance and feedback. For their assistance with data collection and scoring, were are thankful to Carmelina Almanzar, Vanessa Alvarado, Jessica Anderson, Brittany Briceno, Josh Butler, Beatrice Campbell, Courtney Canter, Ann Marie Elmayan, Jenna Haddock, Sarah Harner, Hannah Hendricks, Latasha Holden, Brittany Horton, Tiffany Lam, Joanne Lamb, Torcia Lee, La'Quanda McGeachy, Ashleigh Moore, Jacob Negley, Anna Parisi, Gepernard "Tania" Polidor, Wyatt Smith, Ashkan Tajeddini, Roxy Toledo, and Luz Toribio.

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

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.

Bednall, T. C., & Kehoe, E. J. (2011). Effects of self-regulatory instructional aids on self-directed study. *Instructional Science, 39*, 205–226.

Bereiter, C., & Scardamalia, M. (1993). The psychology of written composition. Hillsdale: Erlbaum.

Bertsch, S., Pesta, B. J., Wiscott, R., & McDaniel, M. A. (2007). The generation effect: a meta-analytic review. *Memory & Cognition, 35*, 201–210.

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.

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

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.

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.

Chan, J. C. K. (2009). When does retrieval induce forgetting and when does it induce facilitation? Implications for retrieval inhibition, testing effect, and text processing. *Journal of Memory and Language, 61*, 153–170.

Chan, J. C. K., McDermott, K. B., & Roediger, H. L. (2006). Retrieval-induced facilitation: initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology. General, 135*, 553–571.

Congleton, A., & Rajaram, S. (2012). The origin of the interaction between learning method and delay in the testing effect: the roles of processing and conceptual retrieval organization. *Memory & Cognition, 40*, 528–539.

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., 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.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2012). Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*, 4–58.

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.

Fass, W., & Schumacher, G. M. (1978). Effects of motivation, subject activity, and readability on the retention of prose materials. *Journal of Educational Psychology, 70*, 803–807.

Fazio, L. K., Agarwal, P. K., Marsh, E. J., & Roediger, H. L. (2010). Memorial consequences of multiple-choice testing on immediate and delayed tests. *Memory & Cognition, 38*, 407–418.

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

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.

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

Hidi, S., & Anderson, V. (1986). Written summaries: task demands, cognitive operations, and implications for instruction. *Review of Educational Research, 56*, 473–493.
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. *European Journal of Cognitive Psychology, 19*, 528–558.
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.
Kellogg, R. T. (1994). *The psychology of writing*. New York: Oxford University Press.
Kellogg, R. T., & Raulerson, B. A. (2007). Improving the writing skills of college students. *Psychonomic Bulletin & Review, 14*, 237–242.
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. (1998). *Comprehension: a paradigm for cognition*. Cambridge: Cambridge University Press.
Kirby, J. R., & Pedwell, D. (1991). Students' approaches to summarization. *Educational Psychology, 11*, 297–307.
Marsh, E. J. (2007). Retelling is not the same as recalling: implications for memory. *Current Direction in Psychological Science, 16*, 16–20.
Newell, G. E. (1984). Learning from writing in two content areas: a case study/protocol analysis. *Research in the Teaching of English, 18*, 265–287.
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.
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.
Pyc, M., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language, 60*, 437–447.
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.
Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: an interim synthesis. *Learning and Individual Differences, 7*, 1–75.
Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: taking memory tests improves long-term retention. *Psychological Science, 17*, 249–255.
Roediger, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. *Psychology of Learning and Motivation, 55*, 1–36.
Rogers, B. (2001). *TOEFL CBT success*. Princeton: Peterson's.
Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: a review of the research. *Review of Educational Research, 64*, 479–530.
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.
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.
Thiede, K. W., & Anderson, M. C. M. (2003). Summarizing can improve metacomprehension accuracy. *Contemporary Educational Psychology, 28*, 129–160.
Toppino, T. C., & Cohen, M. S. (2009). The testing effect and the retention interval. *Experimental Psychology, 56*, 252–257.
Treiman, R. (2000). The foundations of literacy. *Current Directions in Psychological Science, 9*, 89–92.
Wade-Stein, D., & Kintsch, E. (2004). Summary street: interactive computer support for writing. *Cognition & Instruction, 22*, 333–362.
Westby, C., Culatta, B., Lawrence, B., & Hall-Kenyon, K. (2010). Summarizing expository texts. *Topics in Language Disorders, 30*, 275–287.
Wheeler, M. A., Ewers, M., & Buonanno, J. F. (2003). Different rates of forgetting following study versus test trials. *Memory, 11*, 571–580.

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: Association for Supervision and Curriculum.