Comparison of Study Strategies with Interference in Reading a Passage

Arian Anderson Rabino
De La Salle-College of Saint Benilde, Philippines

Abstract
In 2010, Weinstein, McDermott, and Roediger compared study strategies in reading a passage to determine which among rereading, answering comprehension questions, or generating own questions was most useful for students to remember information. The present study replicated their comparison of study strategies in reading a passage, except for the generate questions condition. Instead, a control condition of studying the passage once was implemented. The author also introduced an additional distracter passage to be read in all conditions. This similar but easier passage served as the interference in recalling information from the previously read passage. The study aimed to investigate the robustness and magnitude of the testing effect by comparing study strategies with interference in reading a passage. After the experiment, the author subjected the participants’ performance and estimates of retained information to between-subjects analyses of variance with study strategy (study/reread/answer comprehension questions) as the independent variable.

Keywords: study strategies, interference in reading

Introduction
In a 2010 study by Weinstein, McDermott, and Roediger, they compared different study strategies in reading several passages to determine which among rereading, answering comprehension questions, or generating own questions was most useful for students to remember information from the passages. Results of their study showed that answering and generating questions were better strategies in facilitating recall of information from the passages than rereading, although the latter strategy was more time-consuming than the former. This finding demonstrated the phenomenon called the testing effect: it happens when students remember information better because they took a test about the material compared with not taking a test (Roediger & Karpicke, 2006b). However, it appears that people are yet to be convinced (or are misinformed) of the effectiveness of testing as majority of students still prefer rereading as a study strategy over taking a test (Karpicke, Butler, & Roediger, 2009). Even after students took both open- and closed-book tests, they still predicted that rereading would allow them to recall more information on a later test (Agarwal, Karpicke, Kang, Roediger, & McDermott, 2008).
Studies on the Testing Effect

Most studies about the testing effect compared learning through a form of repeated studying with repeated testing, only with some variations (Agarwal et al., 2008; Butler, 2010; Butler & Roediger, 2007; Carpenter, 2009; Carpenter & Pashler, 2007; Carpenter, Pashler, & Cepeda, 2009; Chan, McDermott, & Roediger, 2006; Cull, 2000; Johnson & Mayer, 2009; Kang, McDermott, & Roediger, 2007; McDaniel, Anderson, Derbish, & Morrisette, 2007; Roediger & Karpicke, 2006a; Szpunar, McDermott, & Roediger, 2008; Weinstein et al., 2010; Wheeler, Ewers, & Buonanno, 2003). Consequently, these studies compared the amount of retained information of their participants after some time, usually after a period of five minutes, one day, two days, three days, and one week. The longest interval between exposure to a restudy or test condition and test of retained information was 16 weeks (Carpenter et al., 2009). Results of these studies consistently showed that repeated studying produced short-term benefits, whereas repeated testing produced long-term benefits on learning. Students who repeatedly studied recalled more information on the studied material when immediately tested (i.e. after five minutes) compared with students who repeatedly took tests. However, after an interval of days and weeks, students who repeatedly took tests retained more information on the studied material compared with students who repeatedly studied. These results showed that greater forgetting was evident in the repeated studying condition than in the repeated testing condition (Wheeler et al., 2003).

Studies on the testing effect also tried to establish the generalizability and ecological validity of this phenomenon in a wide variety of materials, including Graduate Record Exam items (Karpicke & Roediger, 2007), journal articles (Kang et al., 2007), maps (Carpenter & Pashler, 2007; Rohrer, Taylor, & Sholar, 2010), multimedia tools (i.e. videotaped lectures on art history and narrated animation) (Butler & Roediger, 2007; Johnson & Mayer, 2009), prose passages (Agarwal et al., 2008; Butler, 2010; Chan et al., 2006; McDaniel et al., 2007; Weinstein et al., 2010), Test of English as a Foreign Language preparation items (Roediger & Karpicke, 2006a), U.S. history facts (Carpenter et al, 2009), and word lists (Carpenter, 2009; Chan & McDermott, 2007; Cull, 2000; Szpunar et al., 2008; Toppino & Cohen, 2009; Wheeler et al., 2003). On the other hand, the tests commonly used to measure amount of retention in these studies were either through associated cue-target pairs or analogy, free-recall, recognition (i.e. multiple choice), or recall (i.e. identification). Findings from these studies showed that the testing effect is not limited to memory tests that require simple verbal responses, but also apply to complex tasks that require spatial manipulations and other nonverbal answers (Carpenter & Pashler, 2007). However, among the different test formats, recall tests best produced long-term retention and robust benefits of testing. Recall tests may even be the most effective format.
in enhancing student learning in various educational situations (Butler & Roediger, 2007; Kang et al., 2007; McDaniel et al., 2007). Researchers of the testing effect largely attributed this finding to the greater effort students exert in recollecting information among recall tests, whereas in other test formats such as recognition tests, recollection may also depend on the familiarity of items (Butler & Roediger, 2007; Chan & McDermott, 2007).

Aside from the long term retention of information, some studies demonstrated the advantage of the testing effect on the transfer of learning in a variety of contexts (Butler, 2010; Rohrer et al., 2010; Weinstein et al., 2010). These researches tried to disprove the criticism that the testing effect can only be explained by the similarity between the learning strategy and manner of assessment, and thus only manifests in situations that require specific and similar responses. If this criticism was true, it violates the goal of education or vocational training which is the transfer of learning in life situations (Butler, 2010). However, these studies successfully established that the benefits of the testing effect are not compromised when transfer is required (Rohrer et al, 2010). This means that the testing effect is not dependent on the similarity between the initial and final tests (Carpenter, 2009).

The Present Study

Although there has been a considerable amount of studies on the testing effect in the past five years mostly by researchers from the Washington University in St. Louis (i.e. by Agarwal, Butler, Chan, Derbish, Kang, Karpicke, McDaniel, McDermott, Roediger, Szpunar, and Weinstein), no study has yet thoroughly investigated the possible influence or effect of interference in studying a passage. But this does not mean that it has not yet been thought of to affect test performance or recall of information. In fact, Roediger and Karpicke (2006b) recognized the possible influence of interference in learning. However, they did not believe that interference has the ability to compromise the benefits of the testing effect. In their words, “the positive effects of testing are often so large that in most circumstances they will overwhelm the relatively modest interference effects,” (p. 203).

A review of studies on interference and task-switching by Kiesel et al (2010) found that interference happens when there is a strong association between the stimuli and task sets. In short, performance is affected by stimuli from similar tasks because the tasks compete with one another. Another study on forgetting in working memory by Portrat, Barrouillet, and Camos (2008) showed that even when time is kept constant between two task conditions, a task that is more challenging and requires more processing time will be more difficult to recall compared with an easier task. These studies, therefore, suggest that interference could influence learning and test performance.
In the present study, the author replicated the comparison of study strategies in reading a passage by Weinstein et al. (2010), except for the generate questions condition. Instead, the author implemented a control condition of studying the passage once. In line with the studies on interference cited, it was also interesting to examine the strength of the testing effect when subjected to an interfering material— which is a similar but easier task—in studying a passage. Hence, the participants read a distracter passage in all study strategy conditions. The study’s introduction of a distracter passage served as the interference in recalling information from the previously read passage.

The following are some of the questions the present study aimed to answer: Will answering a comprehension test remain a better strategy (testing effect) in studying a passage even when presented with an interfering passage? Will rereading be a better study strategy than just studying the passage once? Will the similar interfering passage compete with the original passage during information recall at the final test?

The author hypothesized that answering a comprehension test would yield a difference compared to the study and reread passage conditions, and thus would demonstrate the testing effect. The author also hypothesized that this phenomenon would affect the amount of information retained among the three conditions even with interference. Thus the main aim of the study was to investigate the robustness and magnitude of the testing effect by comparing the different study strategies with interference in reading a passage.

Other than comparing study strategies in Weinstein et al.’s (2010) study, they also measured the perceived retained amount of information of the participants. Results of their study showed that students in the conditions of answer comprehension questions and generate own questions had higher estimates of retained information than students in the reread condition. In the present study, apart from the actual performance on the final tests, the estimates of retained information after doing the conditions were also measured. In addition to the abovementioned hypotheses, the author hypothesized that the estimates of retained information of the participants in the answer comprehension questions condition would be higher than the estimates of participants in the study once and reread conditions.

Rationale of the replication. The present study replicated Weinstein et al.’s (2010) comparison of study strategies in reading a passage. Their study served as the backbone of the present research because it is one of the most recent contributions to the growing literature about the testing effect. Moreover, their study has sensible implications to students’ study habits, educators’ teaching styles, and scholastic undertakings, in general. The author also chose to replicate Weinstein et al.’s study on passages instead of
other materials because of its practicality and convenience. Passages are part of the educational tradition that students must read and study at all levels of their education, from grade school to high school, to college, and sometimes, even extends until they land a job. The difference between the present study and original study was the exclusion of the generate questions condition. The author removed this condition because Weinstein et al. found that it had no difference with the answer comprehension questions condition, and that generating own questions consumed time the most. In spite of the present study’s no deliberate manipulation of interference, it also marked a slight deviation to the original study by Weinstein et al. The introduction of an interfering material aimed to investigate the robustness and magnitude of the testing effect, and how it could affect the participants’ final test performance and recall of information. Finally, the present replication study tried to determine the generalizability and benefits of the testing effect to Filipino students.

Relevance of the study. More than contributing to the body of knowledge about the testing effect, the author’s primary motivation behind the study was to encourage people to maximize testing in various forms of learning situations in the Philippines. By doing so, Filipino educators and students could take advantage of the phenomenon, and apply it in their scholastic undertakings. On the one hand, educators would be more interested to construct diagnostic tests with the goal of helping students retain information in the long-term memory. On the other hand, students would be more focused in studying lessons as they anticipate a diagnostic test with the goal of obtaining better marks in examinations and retaining more information over time.

Method

Participants

Eighty seven students from De La Salle-College of Saint Benilde, Manila, participated in the study for extra course credit. All students were freshmen taking a General Psychology class. Their age range was 15 to 20 years with a mean age of 16.98 years ($SD = 0.93$). Of the 87 participants, 42 were male, 44 were female, and one student with undeclared sex. Participants were composed of 80 Filipino, two Chinese, one Dutch, one Filipino-Japanese, and three undeclared nationality. The author grouped the participants according to their sections, and sampled by convenience.
Materials

The study used two passages adapting Wikipedia pages on the Taj Mahal and the Harry Potter book series. The passage on the Taj Mahal was the same passage used by the study of Weinstein et al (2010), but was reduced from 575 words down to approximately 432 words. An excerpt of the Taj Mahal passage is the following:

The Taj Mahal is a monument located in Agra, India, that was constructed in the twenty-two years between 1631-1653 by a workforce of twenty-two thousand. It is one of the finest examples of Mughal architectural style. The Emperor Shah Jahan used his vast wealth to commission the construction of the Taj Mahal as a mausoleum for his favorite wife, Mumtaz Mahal, who died during childbirth.

Meanwhile, the Harry Potter passage served as the distracter passage and had approximately 422 words. An excerpt of the Harry Potter passage is the following:

Harry Potter is a series of seven fantasy novels written by British author J. K. Rowling. The books chronicle the adventures of the adolescent wizard Harry Potter and his best friends Ron Weasley and Hermione Granger, all of whom are students at Hogwarts School of Witchcraft and Wizardry.

For the six-item final test, the same questions devised by Weinstein et al were used with two questions obtained from each paragraph. All questions could be answered by a single word or a short phrase.

In addition to the final test, participants took another six-item comprehension test about the Taj Mahal. Participants assigned in the answer comprehension questions took the comprehension test, while all participants from all conditions (study, reread, answer comprehension questions) took the final test after the experiment. The same comprehension test devised by Weinstein et al (2010) was used in the study. All questions could be answered by a single word or a short phrase. Some of the questions in the final test had the same questions found in the comprehension test but were worded differently.

The author obtained the participants’ age, sex, program, school, year level, nationality, assigned study strategy, preferred study strategy, and estimate of retained information after doing the assigned conditions using a questionnaire. Responses on the estimate of retention ranged from 0 (“not remember anything at all”) to 100 (“remember everything perfectly”).

Design

The study used a between-subjects design with study strategy (study, reread, answer comprehension questions) as the independent variable. The
use of a fishbowl technique randomly determined the order of conditions and assignment of conditions to the class sections.

**Procedure**

Students participated in the study voluntarily. They also completed the study as a section with the experimenter and their General Psychology professor present in the room during the session. The author told the participants that they would study a three-paragraph passage for five minutes for a later test. They were also told that they would do an assigned study strategy, answer a questionnaire, and read an additional material for another five minutes.

Depending on the assigned condition, participants read the passage about the Taj Mahal for five minutes; did the assigned study strategy (reread, answer comprehension questions) for five minutes except for the control condition; answered the questionnaire for five minutes; read the Harry Potter passage for another five minutes; and finally, answered the final test about the Taj Mahal for five minutes. The answer comprehension questions condition was an open-book test. After the experiment, the author debriefed all the participants about the purpose and nature of the study. Sessions took an average of 30 minutes.

**Results**

Presented below are the results of the comparison of study strategies as a function of estimates of retained information (how much information participants thought they would remember on the final test); and actual test performance (how many test questions were the participants able to answer correctly). These two dependent variables were subjected to between-subjects analyses of variance (ANOVAs), with study strategy (study/reread/answer comprehension questions) as the independent variable. Results were not significant at $p > .05$.

**Estimates of Retained Information**

Participants estimated the amount of information they would remember from the Taj Mahal passage on a scale of 0 ("not remember anything at all") to 100 ("remember everything perfectly") after doing the assigned study strategy (study/reread/answer comprehension questions). Participants were somehow convinced that they would do better on the final test (i.e. retained more information from the passage) after doing the assigned study strategy of answering comprehension questions ($M_{estimates} = 68.52; SD = 14.06$) than after reading the passage twice ($M_{estimates} = 65.71; SD = 13.40$) or studying the passage once ($M_{estimates} = 66.78; SD = 13.51$).
However, there was no significant difference in the estimates of retained information among the three study strategies, $F(2, 84) = 0.28, p > .05$.

![Bar chart showing estimates of retained information among the three study strategies.](image)

**Figure 1.** Mean estimates of retained information on a scale of 0 ("not remember anything at all") to 100 ("remember everything perfectly") among the three study strategies.

**Actual Performance**

Performance was measured in terms of the questions answered correctly in the six-item final test. Each answer was scored as either correct or incorrect, and no half-points were awarded. Questions could be answered by a single word or a short phrase. Points were not deducted for incorrect spellings or grammatically incorrect phrases. For questions that required short phrases, as long as the main idea was included in the answer, it was considered as correct. For example, in the question “In 1983, the Taj Mahal was declared as what site?”, the correct answer is “UNESCO World Heritage Site.” Participants would get a point for answers such as “UNESCO,” “Heritage Site,” “UNESCO Heritage Site,” “UNESCO Site,” and “World Heritage Site.”
Figure 2. Mean scores of actual performance on a six-item test among the three study strategies.

In terms of actual performance, participants somehow performed better after reading the passage twice ($M_{performance} = 3.57; SD = 1.37$) than studying the passage once ($M_{performance} = 3.12; SD = 1.41$) and answering the comprehension questions ($M_{performance} = 3.08; SD = 1.53$). However, there was no significant difference in the actual performance of the participants among the three study strategies, $F(2, 84) = 1.03, p > .05$.

Discussion

The main aim of the study was to investigate the robustness and magnitude of the testing effect by comparing the different study strategies with interference in reading a passage. The author hypothesized that answering a comprehension test would yield a difference between the study and reread passage conditions, and thus would demonstrate the testing effect. The author also hypothesized that this phenomenon would affect the amount of information retained among the three conditions even with interference. However, the results yielded no significant difference among the study strategy conditions in terms of the participants’ actual performance. In addition, the author hypothesized that the estimates of retained information would reflect the participants’ actual performance. In this case, the results of the estimates of retained information indeed reflected
the actual performance of the participants, which led to a non-significant difference. Nevertheless, the results of the study were still in line with previous researches on the testing effect.

**Actual Performance**

The trend of the means in the actual performance was consistent with the findings of Roediger and Karpicke’s (2006a), Toppino and Cohen’s (2009), and Wheeler et al’s (2003) studies. In Roediger and Karpicke’s experiment 1, after five minutes, participants who studied the passage twice recalled more compared with those in the study once and initial test conditions. Similarly, in Wheeler et al’s experiment 1, after five minutes, participants in the repeated study condition recalled more than those in the repeated test condition. Most importantly, in Toppino and Cohen’s experiment 2, after five minutes, additional studying yielded a small but non-significant recall advantage over testing.

A probable explanation for the non-significant difference of actual performance among the study strategies was that testing effect had not yet manifested its magnitude. As what was demonstrated in various studies by advocates of the testing effect (i.e. by Agarwal, Butler, Chan, Derbish, Kang, Karpicke, McDaniel, McDermott, Roediger, Szpunar, and Weinstein), it was only after an interval of days and weeks that the initial testing, and the testing effect, had manifested its benefits on the long-term retention of information. Consequently, when retention intervals are extended in days or weeks forgetting becomes more evident in the study condition, whereas higher recall is observed in the test condition (Wheeler et al., 2003). Hence, in the present study, the immediate recall test after five minutes was not enough to demonstrate the testing effect.

Another likely explanation for the non-significant difference in the actual performance of the participants was the influence of interference (i.e. the passage about the Harry Potter book series) on the recall of information from the Taj Mahal passage. As Kiesel et al (2010) summarized in their review of studies on interference, any task is potentially an interfering material if its task sets have similar features in the original task that is supposed to be recalled or retrieved. In this case, the passage about the Harry Potter series almost contained the same amount of word number as the passage about the Taj Mahal. Moreover, the study by Portrat et al (2008) showed that a more difficult task would be harder to retrieve or recall after an interfering material. In this case, the subject matter of the passage about the Taj Mahal was essentially meant to be informative, whereas the passage about the Harry Potter book series could be entertaining. Thus, the Taj Mahal passage was more difficult to read than the Harry Potter passage.

Likewise, recall could be affected by the type of reading material a particular task entails. Linderholm and van den Brock in their 2002 study
tested the assumption that readers vary in the type of cognitive processes used when reading expository texts as part of studying compared with reading for entertainment. They found that “readers tailor the types of cognitive processes and strategies they use depending on the reason for reading, and the pattern of cognitive processes that these readers engage in during reading has an effect on the retention of information,” (p. 778).

**Estimates of Retained Information**

The trend of the means in the estimates of retained information followed that of the predictions made by the participants in Weinstein et al’s (2010) study. In their study, results showed that students in the conditions of answer comprehension questions and generate own questions had higher estimates of retained information than students in the reread condition. In this study, the same pattern of favoring the condition of taking a test (i.e. answer comprehension questions) was observed. This suggests that participants felt that they would remember more information on the final test after answering the comprehension questions. This also indicates that students are aware of the benefits of testing on information recall and long-term retention.

However, a likely reason for the non-significant difference of estimated retention among the study strategies was the overestimation of the perceived benefits of studying once and rereading in learning among students (Karpicke et al, 2009). Although students are aware of the benefits of testing, students tend to prefer rereading over other study strategies because it yields short-term benefits, and because students “base their predictions of future performance on what produces rapid short-term gains,” (Roediger & Karpicke, 2006a, p. 254). The same preference was found in the study. Of the 87 participants, 60 students indicated in their questionnaire of favoring the study strategy of rereading than any other study strategy.

**Limitations and Recommendations**

The most obvious limitation of the present study was the lack of subsequent recall tests among the study strategies after the immediate five-minute recall test. Due to time constraints, the author was only able to test the classes in one session, thereby failing to establish and produce the benefits of the testing effect. As what was demonstrated in the literature of the testing effect, it usually takes days or weeks before the advantage of the testing effect over any other study strategy is manifested. Future researchers must make it a point to do subsequent recall intervals if the testing effect is the focus of the study.

Second, similar to Carpenter et al’s (2009) study limitations, the participants of the present study were not individually randomized to the
study strategy conditions, but instead were assigned to groups according to their class section. There is always that possibility that some other extraneous variables, like individual differences, might have had influenced the composition of the groups. Future researchers should take this into consideration; perhaps a randomized block design is preferable.

Third, the administration of the questionnaire before the introduction of the interference passage in all study strategy conditions might have confounded the experiment. Instead of just the purported five-minute interference caused by distracter passage, the addition of the five minute interval during the filling up of the questionnaire might have increased the possible influence or effect of the interference. Hence, instead of having a planned five-minute interference period, a ten-minute interference period could have actually taken effect. Administering the questionnaire before or after the experiment might be advisable in future researches.

Finally, the reduction from the original eight-item comprehension tests and final recall tests used in Weinstein et al’s (2010) study down to six-item tests might have made the tests too short. The reduction of the items might have influenced the allowable correct number of responses in the final recall test, thereby providing fewer chances for the participants to answer items correctly. Future researchers might consider retaining the original items and passages used by Weinstein et al in their study.

References


Further correspondence can be sent to Mr. Arian Anderson Rabino at the Center for Learning and Performance Assessment, De La Salle-College of Saint Benilde, 2544 Taft Ave. Manila, Philippines, e-mail: rabinor@dls-csb.edu.ph