July 2016

Retention of Text Material under Cued and Uncued Recall and Open and Closed Book Conditions

Jeffrey S. Nevid
St. John's University, nevidj@stjohns.edu

Yea Seul Pyun
St. John's University, yea.pyun14@stjohns.edu

Brianna Cheney
St. John's University, brianna.cheney12@stjohns.edu

Recommended Citation
Available at: https://doi.org/10.20429/ijsotl.2016.100210
Retention of Text Material under Cued and Uncued Recall and Open and Closed Book Conditions

Abstract
Evidence supports the benefits of effortful processing in strengthening retention of newly learned material. The present study compared two forms of effortful processing, uncued (free) recall and cued recall, under both open and closed book conditions, on both immediate and delayed (one-week) test performance. Participants read a section of a child psychology text and then completed either an uncued recall task in which they typed as much information as they could recall, or a cued recall task, in which they typed answers to study questions. Recall was conducted under open versus closed book conditions. No differences between cued and uncued conditions were obtained, but participants performed better on immediate test performance in the open book condition. No significant effects were found at delayed assessment. The results point to a short-term advantage of effortful review of text materials performed with access to study materials.

Keywords
Retrieval practice, open book, closed book, cued recall, uncued recall

Creative Commons License
This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License.
Retention of Text Material under Cued and Uncued Recall and Open and Closed Book Conditions

Jeffrey S. Nevid, Yea Seul Pyun, and Brianna Cheney
Psychology Department, St. John’s University, Jamaica, NY 11439, USA
(Received 30 July 2015; accepted 17 May 2016)

Evidence supports the benefits of effortful processing in strengthening retention of newly learned material. The present study compared two forms of effortful processing, uncued (free) recall and cued recall, under both open and closed book conditions, on both immediate and delayed (one-week) test performance. Participants read a section of a child psychology text and then completed either an uncued recall task in which they typed as much information as they could recall, or a cued recall task, in which they typed answers to study questions. Recall was conducted under open versus closed book conditions. No differences between cued and uncued conditions were obtained, but participants performed better on immediate test performance in the open book condition. No significant effects were found at delayed assessment. The results point to a short-term advantage of effortful review of text materials performed with access to study materials.

INTRODUCTION
One of the leading advancements in educational research in recent years is the increased recognition of the importance of effortful retrieval in strengthening learning and retention. A substantial body of research demonstrates that practicing retrieval leads to more meaningful learning and more durable retention (Blunt & Karpicke, 2014; Brown, Roediger, & McDaniel, 2014; Roediger & Butler, 2011). The most common example of retrieval practice is the testing effect, which refers to the finding that retrieval of information via testing or quizzing improves later retention to a greater extent than further studying or rereading the material (Carpenter, 2012; McDaniel, Wildman, & Anderson, 2012; Roediger, Agarwal, McDaniel & McDermott, 2011; Roediger & Butler, 2011; Roediger & Karpicke, 2006; Rohrer, Taylor, & Sholar, 2010; for recent reviews of the testing effect see Karpicki & Grimaldi, 2012; Rawson & Dunlosky, 2012; Rohrer & Pashler, 2010; and Rowland, 2014).

The typical paradigm for demonstrating testing effects consists of three phases, an original learning phase, an intervening phase, and a test or assessment phase (Rowland, 2014). The intervening phase allows experimenters to manipulate study only conditions, such as by comparing a restudy or rereading condition with a testing condition. Evidence of a testing effect is shown by increased performance on subsequent assessment of participants exposed to the testing condition. In some cases, a non-intervention control during the intervening phase is used for purposes of comparison.

Retrieval practice involving more effortful processing, such as with recall tasks, tends to produce greater retention than recognition tasks such as multiple-choice tests (Roediger & Butler, 2011). A recent meta-analysis supported retrieval effort theories of the testing effect, which attributes the benefits of testing to the effort, intensity, and depth of retrieval processes (Rowland, 2014). Retrieval practice involving either uncued (free recall) or cued recall tasks, though not significantly different from each other in their effects, generally produced more robust learning benefits than less effortful recognition tests. That said, recognition tasks also demonstrated reliable evidence of the testing effect. Although retrieval practice often involves a testing intervention, it can also take other forms, including answering a set of study questions (Roediger & Butler, 2011) or writing down as many ideas or concepts as one can recall from reading text materials, or creating concept maps based on the prior reading (Blunt & Karpicke, 2014; Rohrer & Pashler, 2010).

The testing effect challenges the common study practice of merely rereading text material in preparation for exams and points to the advantages of active retrieval strategies for strengthening recall and retention of text material. However, the typical paradigm for evaluating testing effects involves the study of brief passages or memorization of paired-associate lists without access to the study materials during the retrieval phase. Moreover, rereading or restudying conditions may not involve the same level of effortful processing as active retrieval practice. By contrast, most students study with their texts and study materials readily available and typically study sections of a text at a time, rather than just a few paragraphs. However, a limitation of naturally occurring study conditions is that students may passively re-read text material or review notes or highlighted text passes without engaging in the effortful retrieval processes that may be needed to produce stronger and more durable retention. Moreover, students who employ more active study strategies, such as answering questions posed as learning objectives, typically complete these tasks while having access to their reading materials. Further research is needed to examine the robustness of retrieval effects while controlling for effortful processing, especially for longer text passages that typify student study sessions.

A recent study compared a closed book memory retrieval condition with an open book restudying condition, with both conditions employing equivalent effortful retrieval formats (paragraph-style free recall or concept mapping) (Blunt & Karpicke, 2014). The results showed superior performance of a closed book memory retrieval task at a one-week assessment, with no differences between the two retrieval formats. However, the learning materials in this study consisted of two brief reading passages totaling less than 500 words and students had two reading exposures to the text materials, with each followed by a retrieval exercise. The present study sought to compare effortful retrieval tasks under open and closed book conditions based on longer text passages and single reading and practice periods to simulate more typical studying conditions.

The present study compared retrieval practice with books closed with the equivalent effortful task with books open. Moreover,
study materials consisted of a text section of about 1,700 words read during a 15-minute reading phase. We also compared the relative effectiveness of retrieval tasks involving open book vs. free recall task (writing down as much material from the reading materials as one can recall) with a cued recall task (answering study questions based on key content in the reading material). The use of study questions to cue recall may provide additional retrieval cues to strengthen memory retrieval effects relative to unassisted free recall. Moreover, we examined the effects of the experimental variables on performance following retrieval practice and a delayed assessment about a week later to ascertain longer-term effects of retrieval practice.

**METHOD**

**Subjects**
A total of 137 undergraduate students (23 males, 104 females) in introductory psychology courses at a large metropolitan northeastern university participated in the study in partial fulfillment of a course requirement. Ninety-two percent of the participants were between the ages of 18 and 21 (M = 19.24, SD = 1.95). The study sample comprised 75 freshmen, 29 sophomores, 16 juniors, 14 seniors, and 3 who failed to specify their college level. Self-identified ethnicities were as follows: Non-Hispanic White, 26.3%; Black or African American, 21.2%; Hispanic or Latino, 15.3%; Asian or Pacific Islander, 25.5%; and other, 11.7%. Participants were recruited through an electronic sign-up system.

To provide an incentive for effort, participants received raffle tickets with chances to win $50 or $25 gift cards depending on their performance on their quiz performance at immediate and delayed assessment. All participants received one raffle ticket after completing both parts of the study as additional raffle tickets based on their best performance on the two quizzes according to the following schedule: 5 additional raffle tickets for correctly answering at least 50% of the questions, 10 additional tickets for correctly answering at least 60% of the questions, 15 additional tickets for correctly answering at least 70% of the questions, or 20 additional tickets for correctly answering at least 80% of the questions.

**Procedure**
The experiment consisted of five phases: (1) a reading phase; (2) a retrieval phase; (3) a distractor phase; (4) an immediate assessment phase, and (5) a delayed assessment phase. During the 15-minute reading phase, participants were presented with a text passage of approximately 1,700 words drawn from a child psychology textbook (Feldman, 2014). During the retrieval phase, participants practiced a retrieval exercise for 15 minutes based on one of four study conditions to which they were randomly assigned: (1) cued recall with book open; (2) cued recall with book closed; (3) uncued recall with open book; and (4) uncued recall with closed book. In the free or uncued recall condition, participants were instructed to use a computer keyboard and mouse interaction as they could recall about the text passage they had just read. In the cued recall condition, they were presented with three study questions and asked to type any information they recalled in the text passage. In the open book task, students completed the retrieval task while they had access to the original reading materials. In the closed book task, students completed the identical retrieval task, but without access to the reading materials. After the retrieval phase, participants completed a demographic questionnaire as a distractor task, which was then followed immediately by a 20-item multiple-choice quiz assessing basic knowledge of the text content. The twenty-item multiple-choice quizzes assessed basic content of acquisition of text material, including concepts related to Piaget's stages of moral development (incipient cooperation, preconventional morality, conventional morality, and autonomous cooperation), critiques of Piaget's model, prosocial behavior, social learning theory, abstract modeling, and reciprocity. The questions were scaled to lower middle level Bloom taxonomy assessing skills of remembering, understanding, and applying concepts (Anderson & Krathwohl, 2001).

Two comparable versions of the quizzes were constructed to assess knowledge of the same content from different perspectives. The design included one version at immediate assessment and an alternate version at delayed testing. Participants were randomized to the order of the alternate versions administered at the two testing occasions. In constructing alternate forms assessing the same concepts, we rephrased questions using “feature-to-concept” and “concept-to-feature” formats (Hannin, Lozano, Frias, Picollo-Hernandez, & Furman, 2010). In the delayed assessment phase, participants returned to the laboratory approximately a week after initial assessment to complete the alternate version of the quiz. Sample quiz items are shown in the Appendix. Alternate versions of the quizzes were used to control for testing effects, such that participants were not tested on the same questions. Use of alternate forms of assessment increases ecological validity, as students often take practice quizzes when preparing for actual exams that include a different but related set of items than those included on actual exams. Regardless of the assigned study condition, all participants completed the delayed assessment quiz without any additional cues or prompting.

**RESULTS**
One hundred thirty two participants completed both parts of the study. Those who failed to complete either the immediate assessment and so were dropped from the analysis of delayed retention effects. All participants completed every question in immediate and delayed assessments. Preliminary analysis showed no significant differences between the two alternate forms of the quiz, t(131) = .88, p = .38. Moreover, the two versions showed a moderately strong relationship, r = .60, p < .001, even though they were administered approximately a week apart. Means and standard deviations for student performance on the multiple choice quizzes at both immediate and delayed (one week) assessment intervals are shown in Table 1. Not surprisingly, we also found poorer retention over time when comparing immediate (M = 74.24, SD = 14.94) and delayed (M = 66.82, SD = 16.93) assessments, t(131) = 6.66, p < .001.

Results of hierarchical multiple regression analyses testing retention relative interaction effects of experimental variables on performance at immediate and delayed testing intervals are shown in Tables 2 and 3, respectively. For each analysis, we first entered word count in the analyses to control for differences in retrieval production. Word count was measured by the number of words students typed into the computer during the retrieval phase (M = 262.90, SD = 95.19). Word count emerged as a significant predictor of quiz performance at delayed testing, f(1, 135) = 21.00, p < .001. Delayed assessment was added to the model (see Table 3). Further, the interaction term was included in the model (see Model 2 in Table 3). The interaction effect failed to significantly contribute to prediction of immediate test performance, f(1, 132) = .48, p = .48. No significant effects at the delayed assessment were found for either Cued vs. Uncued Recall conditions, or Open vs. Closed Book conditions, nor were there any significant interaction effects (see Table 3).

**DISCUSSION**
The present study examined the learning benefits of cued vs. uncued recall under both open and closed book recall conditions on quiz performance at immediate and delayed (one-week) testing phases. The findings indicated superior performance on a 20-item knowledge quiz for open book recall, but only for immediate testing. That is, students benefited from the opportunity to use text materials when they were tested shortly following the retrieval task, as compared to other students who performed the retrieval task while relying entirely on memory. Open book recall may provide additional retrieval cues that help reinforce retention of recently read material, as well as providing an additional opportunity for encoding of information to be learned from reading material.

The present study had the advantage of testing knowledge of concepts drawn from reading college level text materials rather than laboratory tasks such as paired-associate learning that are often used to focus on research on testing effects. The study also benefited from using alternate forms of a knowledge quiz to control for testing effects and to model naturally occurring study conditions in which different types of study and retrieval tasks are combined in questions to the actual exam questions. Introducing a delay between initial practice and delayed assessment also models the type of delay students frequently encounter between practice quizzes and exams.

Research on the testing effect has yielded mixed evidence on the short-term benefits of testing versus retesting. Although some studies, including the present one, failed to find evidence in favor of testing effects or retrieval practice (closed book recall) relative to retesting (open book recall), a recent meta-analysis suggests evidence of reliable short-term testing effects (Rowland, 2014). The two recall conditions (Cued vs. Uncued recall) in the present study produced comparable results, which are consistent with findings from other researchers that the learning benefits of retrieval do not appear to depend on the format of the retrieval task (Blunt & Karpicke, 2014; Smith & Karpicke, 2014). However, the effectiveness of the retrieval tasks in the present study may have been mitigated because textbook materials themselves have built-in cues in the form of headings and study questions. It is conceivable that cue recall might better facilitate performance for text materials that is lacking such internal scaffolding.

The results of the present study differed from those of Blunt and Karpicke (2014). Although these other researchers did not test for differences in immediate retention, they showed a long-term (one week) benefit for closed book memory retention over an open book retesting condition with an equivalent effortful task following the reading assignment. Our results point to an immediate benefit of assisted (open book) recall relative to closed book recall (Model 1) in the immediate assessment phase, which was not significant, f(2, 132) = 2.48, p = .09. However, open book recall individually contributed incrementally to prediction of test performance, f(1, 136) = 2.07, p < .05. Open book condition was also marginally significant, f(2, 136) = .97, p = .05, when the interaction term was included in the model (see Model 3 in Table 2). The interaction effect failed to significantly contribute to prediction of immediate test performance, f(1, 132) = .51, p = .48. No significant effects at the delayed assessment were found for either Cued vs. Uncued Recall conditions, or Open vs. Closed Book conditions, nor were there any significant interaction effects (see Table 3).

**TABLE 1. Means and Standard Deviations for Quiz Performance (% correct)**

<table>
<thead>
<tr>
<th>Predictor</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immediate Assessment</td>
<td>74.24</td>
<td>14.94</td>
</tr>
<tr>
<td>Delayed Assessment</td>
<td>66.82</td>
<td>16.93</td>
</tr>
</tbody>
</table>

**TABLE 2. Hierarchical Regression Analysis of Study Conditions on Immediate Quiz Performance (N=137)**

<table>
<thead>
<tr>
<th>Predictor</th>
<th>B</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>.16**</td>
<td>.01 **</td>
<td>-16.30</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Word Count</td>
<td>.06</td>
<td>.01</td>
<td>7.17*</td>
<td>.003</td>
</tr>
<tr>
<td>Cued vs Uncued</td>
<td>.01</td>
<td>.01</td>
<td>3.20</td>
<td>.001</td>
</tr>
<tr>
<td>Open vs Closed Book</td>
<td>.01</td>
<td>.01</td>
<td>3.20</td>
<td>.001</td>
</tr>
</tbody>
</table>

**TABLE 3. Hierarchical Regression Analysis of Study Conditions on Delayed Quiz Performance (N=137)**

<table>
<thead>
<tr>
<th>Predictor</th>
<th>B</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>.16**</td>
<td>.01 **</td>
<td>-16.30</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Word Count</td>
<td>.06</td>
<td>.01</td>
<td>7.17*</td>
<td>.003</td>
</tr>
<tr>
<td>Cued vs Uncued</td>
<td>.01</td>
<td>.01</td>
<td>3.20</td>
<td>.001</td>
</tr>
<tr>
<td>Open vs Closed Book</td>
<td>.01</td>
<td>.01</td>
<td>3.20</td>
<td>.001</td>
</tr>
</tbody>
</table>
book memory retrieval, but no differences at one-week follow up. Several important differences between the two studies are noteworthy. Our study involved a longer reading assignment (1,714 words versus two paragraphs of each less than 300 words in the Blunt and Karpicke study). Secondly, the present study employed a single recall task after the reading phase, whereas the Blunt and Karpicke participants read a paragraph, performed the learning activity (recall or rereading), and then re-read the paragraph and again repeated the learning activity. The same procedure was then repeated for the second paragraph. Testing effects are likely to be stronger in conditions in which memory retrieval is repeated and each practice is followed by re-exposure (feedback) to the reading material. However, the interruptions in a studying routine necessitated by repeatedly reading and performing a retrieval task for each paragraph may not generalize to typical studying conditions.

The present study may have practical implications for assisting students in developing more effective study habits, as well as several important limitations. Students typically prepare for course examinations by rereading required text material and reviewing class notes. However, rereading or review may lack the degree of effortful processing needed to enhance memory retention. As evidence of the testing effect demonstrates, practicing recall of recently read information can strengthen retention. The question we posed is whether unassisted (closed book) recall is a more effective study strategy for strengthening retention of newly acquired information than aided (open book) recall under equally effortful conditions. Our results suggest that students may benefit in short-term retention from having a second look at newly learned material when performing a retrieval exercise. Open book tasks may provide additional opportunities for encoding new information and for retrieval cues to jog memory of previously encoded information. Unlike passive re-reading, open book recall requires more effortful processing in the form of reciting as much information as one can recall within a truncated time period or by answering a set of study questions.

Perhaps most importantly, our findings failed to demonstrate a lasting benefit of an open book recall task. These results are consistent with other research showing that open book testing as a method of retrieval practice produces better initial retention than closed book testing, but no differences when retention is measured after a delay (Agarwal, Karpicke, Kang, Roediger, & McDermott, 2008). Our study differed by providing open book access during a recall task rather than a testing task. Not surprisingly, we also found poorer retention over time when comparing immediate and delayed assessments.

These results present something of a cautionary tale with respect to student preparation for exams. Students who perform recall tasks with access to text materials may experience a temporary boost in retention on practice exams, but our results suggest this effect may be short-lived. Moreover, overconfident metacognitive judgments of knowledge may lead students to prepare less thoroughly or practice less effortful retrieval in preparing for course exams. They may erroneously believe their performance on practice exams will carry over to course exams.

The present study had several important limitations. First, our focus was on comparing two types of recall tasks, cued and uncued, which were practiced while students either had access to the learning materials or did not have access to these materials. Given the robustness of retrieval practice effects, we did not employ a more-extensive control. Secondly we examined effects on test performance based on a single retrieval task. As evidence from studies on the testing effect demonstrates, repeated retrieval can enhance longer-term retention (Roediger & Butler, 2011). Further research is needed to determine whether repeated retrieval tasks performed closer to the time of delayed assessment, or the use of spaced retrieval tasks, can lead to more durable learning effects relative to additional study or rereading of text materials.

REFERENCES


# Appendix
Sample Feature-to-Concept, Concept-to-Feature, and Apply Questions from Alternate Versions of Multiple Choice Quizzes for Three Concepts

## Incipient Cooperation Stage of Moral Development

### Feature-to-concept:
- According to Piaget, which of the following stages of moral development lasts from age 7 to age 10 and is marked by children's games becoming more social?
  - a. incipient cooperation stage*
  - b. autonomous cooperation stage
  - c. heteronomous morality
  - d. concrete morality

### Concept-to-feature:
- Piaget describes the incipient cooperation stage of moral development as lasting from age 7 to 10 and as the stage during which
  - a. children's games more rigidly adhere to rules
  - b. children's games become more social in nature*
  - c. children's games become less social in nature
  - d. children realize that game rules can be modified if the players agree to the changes

## Autonomous Cooperation Stage of Moral Development

### Concept-to-feature:
- The autonomous cooperation stage of Piaget's moral development model is characterized by which of the following:
  - a. the view that rules are invariant and unchangeable
  - b. the belief that when rules are broken, punishment will immediately follow
  - c. an increased ability to understand the formal rules of games
  - d. the understanding that rules are created by people and subject to change*

### Feature-to-concept:
- Children who understand that rules and laws are created by people and are subject to change according to the will of people would likely be in which of the following stages of Piaget's moral development model?
  - a. incipient cooperation stage
  - b. autonomous cooperation stage*
  - c. heteronomous morality stage
  - d. concrete morality stage

## Immanent Justice

### Apply:
- Five-year-old Juan cut in front of several classmates waiting in line for the water fountain at his school. Even though his teacher and classmates did not see him cutting the line, Juan was certain that he would be punished for breaking a classroom rule. Which of the following principles does Juan's belief demonstrate?
  - a. immanent justice*
  - b. concrete morality
  - c. autonomous cooperation
  - d. incipient cooperation

### Apply:
- Which of the following behaviors best exemplifies the principle of immanent justice?
  - a. Jane who cheated on a test, but was not worried about getting caught since no one saw her do it
  - b. Phillipe who agrees with his friends to change the formal rules of a card game to make the game more fun
  - c. Jacobo who stole his classmate's candy without getting caught, but still worries that he will be punished for doing so*
  - d. Leticia who loves playing games after school so that she can interact with friends.

Note: Asterisk indicates correct answer.