Self-Regulated Study and the Testing Effect: A Combined Investigation Regarding Material Retention

Teresa Daniele

Follow this and additional works at: http://commons.colgate.edu/car
Part of the Cognition and Perception Commons

Recommended Citation
Available at: http://commons.colgate.edu/car/vol6/iss1/14
Self-Regulated Study and the Testing Effect: A Combined Investigation Regarding Material Retention

By Teresa Daniele, Class of 2010

Abstract

The potential effects of self-regulated study and memory retrieval on learning were investigated in a $2 \times 2$ within-subjects design. Participants studied and later were asked to recall Swahili-English word pairs in a cued-recall paradigm. The independent variables were Self-Regulation (whether participants could drop a word pair from the study list) and Memory Recall (whether participants were asked to generate the second member of a word pair during study). There was a main effect of Self-Regulation (learning increased when participants were unable to discard word pairs), no main effect for Memory Recall, and no significant interaction. The lack of a statistically significant result in the Memory Recall condition perhaps was due to inadvertent recall during the recognition task. Implications and future directions will be discussed.

The ability to navigate the world effectively is crucial to the survival of human beings. A key component of this navigation is the ability to learn about the environment, and subsequently apply this learned knowledge in the future. Further, humans most efficient at this task are better equipped to face challenges and steer their worlds successfully. Thus, it is crucial to know as much as possible regarding how human beings learn most effectively and efficiently. The current research provides insight into how human beings learn, errors they make in their learning, and potential methods to combat these errors.

More specifically, in an academic setting, the current research has both pedagogical implications as well as suggestions for self-regulatory methods. From an educational perspective, insight regarding the most effective means of information presentation could aid in course design, presentation of material to be learned, testing frequency as well as testing design itself. From a learner perspective, a considerable component of an education occurs when one is self-regulating their own study away from the learning environment. Findings that provide insight into how learners regulate their study, mistakes they may make, and factors that aid self-regulation could prove to be extremely important in the educational realm. In order to investigate this vital issue, the current study was designed to further explore phenomena associated with self-regulated study and a finding known as the testing effect.

Testing Effect

The present study ties together two interesting and pragmatic issues regarding the educational arena. First,
the idea known as the testing effect. In an attempt to maximize learner capacity and instruction potential of educational facilities, studies have been conducted exploring various aspects and effects of testing. Testing has been shown to improve retention of material superior to that of additional study time; this phenomenon has become known as the testing effect (McDaniel, Anderson, Derbish, & Morrisette 2007; McDaniel, Roediger III, & McDermott, 2007; Roediger III & Karpicke, 2006). In general, the testing effect is the consequence that direct learning has on a person’s memory; the act of taking a test itself enhances memory through direct processes and not as a result of further steps taken to learn after the testing occurs (McDaniel, Anderson, Derbish, & Morrisette 2007; McDaniel, Roediger III, & McDermott, 2007; Roediger III & Karpicke, 2006). The most robust testing effect occurs when testing requires the participant to produce an answer rather than recognize it (McDaniel, Anderson, Derbish, & Morrisette 2007; McDaniel, Roediger III, & McDermott, 2007; Roediger III & Karpicke, 2006). In other words, when participants are required to recall the material on a test and subsequently produce it (short answer), retention is superior to when participants choose answers they believe to be correct (multiple choice). Further, this enhanced learning based on previous testing has been shown to extend to final test performances in the long term (McDaniel, Anderson, Derbish, & Morrisette 2007).

Feedback has been shown to effect testing paradigms (McDaniel, Anderson, Derbish, & Morrisette 2007; McDaniel, Roediger III, & McDermott, 2007; Metcalfe & Kornell, 2007). Specifically, feedback was shown to positively effect performance and when it was not given, incorrect responses went unchanged (Metcalfe & Kornell, 2007). Some studies suggest this helpful effect of feedback can only be seen in short answer paradigms (McDaniel, Anderson, Derbish, & Morrisette 2007). However, others argue that feedback positively effects many different learning situations, with the most positive results seen in short answer paradigms (McDaniel, Roediger III, & McDermott, 2007). Overall, it is clear from this data that feedback positively supplements the effects of the testing effect.

The testing effect has powerful implications for learning and pedagogy. After reviewing literature, we have found there to be robust evidence for a testing effect. Further, previous studies suggest that there is an interaction of between material retention and self-regulated study (Kornell & Bjork, 2007; Metcalfe & Kornell, 2007; Roediger III & Karpicke, 2006). This brings us to a discussion of another important phenomenon regarding the educational arena: self-regulated study.

Self-Regulated Study

Recent investigations on learning have suggested an effect of self-regulation on material retention (Kornell & Bjork, 2007; Kornell & Bjork, 2008). In many of these studies, participants are given control of when to stop studying
material (Drop condition). Alternatively, other participants are not given this control and required to study all material to be tested (No Drop condition). Kornell and Bjork (2008) chose to test self-regulated study specifically regarding the decision to drop flashcards that contained content to be tested. Materials to be studied were Swahili-English word pairings from Nelson and Dunlosky (1994). If participants chose to drop a flashcard, the word pair was removed from the electronic cycling of material. A robust finding was seen in Experiment 1, where allowing participants to drop flashcards was detrimental to learning. This implies that learners are not properly regulating their study habits, reflecting findings of Kornell and Bjork's (2007) review.

Kornell and Bjork (2007) suggest that participants do not make choices that maximize learning. They argue that participants decide to stop studying material in either of two situations. The first is when they judge they have mastered the material, and the second is when they judge that they will not be able to master the material. Further, participants select easiest to-be-learned material when allowed to restudy previously presented material; they do not focus on the more difficult material they do not yet know. Lastly, participants seem to improperly judge their retention ability; participants drop items they may know in the short term, not taking into account their loss of memory over time. Many participants did not believe testing was a chance to learn (only 18% surveyed believed it to be; Kornell & Bjork, 2007). If participants can harness proper self-regulatory mechanisms when learning, the improvement in retention could be substantial. Alternatively, if certain factors are found to mediate the effects of self-regulation, perhaps improved retention could be achieved.

Experiment 4 of Kornell and Bjork (2008) failed to reproduce a difference between No Drop and Drop conditions that was present in Experiment 1. A key difference between these experiments concerns whether participants were required to produce the answer during study. When participants did not produce answers during learning, greater differences between self-regulation conditions (Drop and No Drop) are seen. This begs the question of why there was considerably less of a difference between Drop and No Drop conditions in Experiment 4 (where material was produced during learning) than in Experiment 1 (where material was reread)? Two possible explanations include: (i) metacognitively, participants are judging their learning more accurately because they are abruptly made aware of their lacking knowledge if they are not able to produce the material, or (ii) a memory-based explanation suggests learning is more effective due to material production, which reflects the idea of the testing effect (McDaniel, Anderson, Derbish, & Morrisey 2007; McDaniel, Roediger III, & McDermott, 2007; Roediger III & Karpicke, 2006). More simply, did production of answers during learning cause participants in Kornell and Bjork’s (2008) study to actually learn better (implying the testing effect) or are they simply deciding how well they know the material more accurately?
Looking more closely at this finding, it is necessary to note another aspect of Kornell and Bjork’s (2008) study. In Experiment 4, Kornell and Bjork placed participants into three conditions requiring either the user to decide when to drop word-pairs (User), the computer to drop pairs after one correct answer (Autodrop 1), or the computer to drop pairs after two correct answers (Autodrop 2). Each of these conditions (User, Autodrop 1, Autodrop 2) had a No Drop condition associated with it, in which the participant was required to continue studying all of the material the entire study period. Oddly, there were major differences between the No Drop conditions despite being the same procedurally (the only difference was which Drop condition (User, Autodrop 1, Autodrop 2) the No Drop condition was paired with). The experimenters did not address this finding in their paper, but it is worthy of note and further investigation. To explore the main question regarding the possible mediation of generation on self-regulated study, it is important to replicate the differences between Experiments 1 and 4 of Kornell and Bjork (2008) in a single paradigm. These differences suggest an interaction between retrieval type and self-regulated study but are complicated by the varied No Drop results described above.

Experiment 1

This experiment was designed to: (i) replicate the testing effect findings and self-regulation data in a single testing paradigm, (ii) gain an early sense of possible interactions occurring between these two variables, and (iii) gain insight as to why differences were seen among the No Drop conditions in Kornell and Bjork’s (2008) Experiment 4.

Methods

Participants. 20 undergraduates participated in one 50-min session as part of a requirement for an introductory psychology course. All participants reported being native English speakers with no experience with Swahili.

Stimuli and Procedure. The experiment performed was a 2 x 2 within-subjects design where participants studied and later recalled Swahili-English word pairs using a cued-recall paradigm. One hundred Swahili-English pairs formed the stimulus set, with pairs selected randomly by participant and condition (Nelson & Dunlosky, 1994).

The independent variables were Self-Regulation (whether participants could drop a word pair from the study list) and Memory Recall (whether participants were asked to generate the second member of a word pair during study). A computer with a standard LCD monitor was used for stimulus presentation and data collection. Stimuli were presented in text mode in a simulated “flash card” situation, where one word was presented on the screen at a time.

The experiment was blocked, with each block representing a condition (Drop-Generate, Drop-No Generate, No Drop-Generate, and No Drop-No Generate). The order of blocks was randomized for each participant. Before the main experiment began, participants had a practice block in which data were presented.
not recorded. The inter-block interval (time between blocks) was 10 seconds.

Blocks consisted of 10 to-be-remembered word pairs. Each block had two phases (initial presentation and study). The initial presentation was the same for all blocks, and the study phase varied by condition. In the initial presentation the pairs were presented (one word at a time), with the Swahili word presented for 3 seconds, the English word presented for 3 seconds and a 1 second inter-trial interval (time between each pair). After this initial presentation, the study phase sequence varied by condition (see below). In all conditions, there were 7.5 minutes allocated to study time with a countdown timer presented on the computer screen indicating time left for the block. Participants were informed that all 7.5 minutes would need to be utilized before the computer would allow them to continue to the test phase. After all blocks were completed, a cued-recall test was performed for all 40 word pairs, presented in random order.

The study phase varied by condition. In the Generate conditions, participants were required to type in the English word that they believed corresponded to the Swahili cue word, followed by a presentation of the correct English word. In the No Generate conditions, participants simply looked at the correct translation after viewing the Swahili cue word (similar to the study phase). In the Drop conditions, participants were able to control when to stop studying a pair by dropping it from rotation after the Generate or No Generate aspect of the study phase. In the No Drop conditions, all pairs were presented in rotation and participants did not have the choice of dropping them.

Final testing presented the Swahili cue words in random order and participants were asked to type in the corresponding English word. The Swahili cue was shown and participants were urged to hurry if they took longer then 12 seconds to produce the English response.

Results

The average number of correct responses from the test phase are presented in Table 1 and Figure 1. The number correct in the test phase were subjected to a two-way, repeated measures Analysis of Variance (ANOVA), with Self Regulation (Drop or No Drop) and Memory Recall (Generate or No Generate) as the independent variables. There was a significant main effect for Self Regulation (F=5.830, p=.026), with participants demonstrating better learning when they were not allowed to control their study. This result replicates previous findings and is consistent with the idea that there are discrepancies between metacognitive control and ideal learning (e.g., Kornell & Bjork, 2008; Kornell & Bjork, 2007). There was no main effect of Memory Recall (F= <1), nor an interaction between Self Regulation and Memory Recall (F= <1).

Discussion

The main effect seen for Self Regulation shows that the ability to control ones studying is detrimental to material retention. Mean correct responses for final testing were
significantly lower when participants were in the Drop condition. This replicates those of previous studies further suggesting that self regulation is vigorously harmful to material retention (Kornell & Bjork, 2008). Further, this finding suggests participants are consistently making decisions to stop studying material that results in lower material retention.

There was no significant main effect found for Memory Recall. In the Generate condition participants were meant to actively recall the corresponding English word pair. Participants in the No Generate condition were meant to simply reread the English word, without internally recalling what they believed to be the corresponding answer. It is possible that participants in the No Generate condition (in which they were supposed to reread the English word) were in fact recalling English word pairs when presented with the Swahili cue words prior to seeing the English word presentation.

In order to ameliorate this potential confound, future work will utilize a two-alternative forced-choice paradigm. In this paradigm, participants would be presented with the Swahili cue word and then presented with two possible English word answers (one correct, one incorrect). This two-alternative forced-choice paradigm was shown in a previous study to be less effective in enhancing memory than a recall paradigm (McDaniel & Masson, 1985). Therefore, it is hopeful that this paradigm will enhance the manipulation difference between the Memory Recall conditions (Generate and No Generate).

In addition to this adjustment, it may be most effective to present Swahili cue words for a shorter period of time (1.5s as in Kornell and Bjork (2008) Experiment 1) or at the same time as the two English choices are presented. This would further eliminate any possibility that participants would internally recall when they were meant to simply recognize. In the Generate condition, participants would be presented with the Swahili cue word and then provided a blank space to type in their presumed answer.

There was evidence (though not statistically significant) of an interaction between Self Regulation and Memory Recall. Specifically, a greater improvement in mean correct responses was seen when participants were required to produce answers in the Drop conditions than in the No Drop conditions. In other words, when participants were able to control when they dropped material, generating answers made more of a difference in their learning then when they were unable to drop material. This has important implications for the main question of a possible interaction between self regulation and production of answers causing the differences seen between Kornell and Bjork’s (2008) Experiment 1 and Experiment 4.

Lastly, the current study showed evidence of ceiling effects. Increasing the amount of words per block from ten to twelve (adjusting the time per block accordingly) might assure participants are sufficiently challenged. This adjustment, along with the other proposed changes, may allow a more definitive exploration of the two alternative explanations detailed above.
Concluding Remarks

The issues of self-regulated study and the testing effect are pragmatic and vital to gaining valuable insight into the realm of learning both within and beyond an educational setting. As stated above, future studies will further address the questions surrounding memorial and metacognitive issues involved in self-regulated study and the testing effect. The findings from this work will be of importance to both basic scientists and those interested in educational applications.

Acknowledgements

Dr. Douglas Johnson of Colgate University was, and continues to be, an integral part of the development and implementation of this study. Dr. Johnson played a central role in the extensive literature review, the proposal for research, the decision-making process regarding focus, protocol, and aims, development of the current study, creation of test protocol, data analysis, subsequent study adjustment and development of this paper. I am extraordinarily thankful for his involvement in this study and very much look forward to future endeavors.

References


Figure 1: Average Correct Responses per Block in Final Testing

![Bar Chart]

**Table 1: Average Correct Responses in Final Testing**

<table>
<thead>
<tr>
<th></th>
<th>Drop</th>
<th>No Drop</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generate</td>
<td>7.2</td>
<td>7.9</td>
</tr>
<tr>
<td>No Generate</td>
<td>6.7</td>
<td>7.8</td>
</tr>
</tbody>
</table>