How Reflecting on Previous Recall Affects the Efficacy of Later Restudy Periods

Elias Theodosis and Jeffrey D. Karpicke

Introduction

Early studies by people such as Tulving (1967) have investigated the effects of repeated testing, repeated studying and alternating study and test periods. Tulving's experiments show that the act of recalling material not only evaluates learning, but also has an impact on that learning. More recent studies have shown the importance of testing for long-term retention (Hogan & Kintsch, 1971; McDaniel & Masson, 1985; Roediger & Karpicke, 2006b; etc.). It has also been shown that alternating study and test trials is perhaps more effective than either repeated studying or repeated testing for long-term retention (Karpicke, 2007).

In many studies that use this standard procedure (study, test, study, test, etc.) participants rarely reach above 70% - 80% recollection of the material. This is often regardless of the number of trials they are given. Although this is quite high, it also raises the question of why participants do not approach closer to the goal of 100% recollection. We believed this might be a result of their inability to accurately judge and recall their performance during previous test periods. This would decrease the benefits received from those restudy periods and lead to less progress in their overall learning.

We also know that students commonly learn by repeatedly studying and testing themselves over their materials. So, if we improved this method further then we could potentially enhance the effectiveness of a person's time spent studying.

We examined the impact of giving feedback during restudy periods in a standard repeated study and test procedure. By doing this we aimed to improve the efficacy of these restudy periods. This should result in higher initial acquisition of the material and greater overall retention (which we tested after a week delay).

Method

120 Purdue University undergraduate students were asked to commit a short scientific passage to memory in one of four conditions.

In the control condition they alternated back and forth between studying the text and playing a computer game. The other three conditions were asked to alternate between studying and testing over the text (free recall).

• The first of these conditions simply follows this study-test pattern for three trials. This is referred to as the standard condition.
• The participants in the second condition were given feedback on their previous test performance, via computerized highlighting of missed ideas, while restudying the text the second and third times. This feedback was given through a computer-based scoring algorithm called Quickscore (Grimaldi & Karpicke, 2012) that was integrated into the study. This is defined as the computer-highlighted condition.
• The final condition requires participants to supply their own feedback by highlighting the ideas that they missed during each previous test period. Therefore, this is called the self-highlighted condition.

The participants in all four conditions were asked to take a final test over the material after a one week delay.

Session 1: Initial Learning Curve

During the initial learning period, the participants in the standard condition appear to be learning the material at a faster rate than the other conditions. This leads to the standard condition having a higher overall proportion recalled (.63) than either the self-highlighted condition (.53) or the computer-highlighted condition (.51) by trail 3.

The read only condition was not included in the session 1 data because they were not asked to recall the material until the final test.

Session 2: Final Test

The final test shows that participants in the standard condition seem to retain a much higher proportion of the material (.46) than the other groups. Another surprising aspect is that the self-highlighted condition (.30) does not seem to be noticeably different than the computer-highlighted condition (.27).

You can also see that the read only condition is by far the worst (.12), which was expected.

Conclusions

Participants that simply read the text recalled more than those that received feedback through highlighting. This seems to suggest that highlighting previously missed concepts might be disrupting encoding.

The fact that this seems to disrupt encoding does not mean that the feedback itself is harmful. It could mean that it is distracting their attention away from retaining the ideas that are highlighted or not highlighted as part of the feedback.

If participants are simply focusing on, and retaining, the non-highlighted ideas then evaluating results on a trial-by-trial basis should show this effect. (this evaluation is in progress)

Further studies could attempt to find and employ a means of feedback that is less intrusive for encoding.

The read only condition meets expectations by being the lowest at overall retention. This is a clear example of the benefits of testing over simply reading the material.

Address correspondence to Jeff Karpicke (karpicke@purdue.edu)