The Impact of Metacognitive Beliefs on the Testing Effect

Avery Wall and Neil Mulligan, PhD, Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill

Background

The testing effect refers to the improved memory seen when a person uses retrieval practice (e.g., taking a test) compared to when they simply restudy some information (e.g., rereading it) (Roediger & Karpicke, 2006). It is a robust phenomenon, having been found across a variety of material types and test formats (Rowland, 2014). However, most students view testing as a way to assess learning rather than improve it, and few students use self-testing in their study routines (Karpicke & Roediger, 2008; Karpicke et al., 2009). Given these negative views on testing, it is important to understand whether these beliefs can impact the testing effect. Beliefs may impact the testing effect through two mechanisms: effort and testing anxiety. According to the effort hypothesis, the increased effort required to retrieve information produces the testing effect, so it follows that those with negative beliefs may not put in the effort required to produce a testing effect (Pyc & Rawson, 2009). Higher test anxiety has been linked to lower exam performance, so anxiety from being evaluated may impede the benefits of testing (Putwain & Remedios, 2014). The goal of this study is to determine whether beliefs about retrieval practice influence the testing effect—specifically, do individuals who hold negative beliefs about testing still benefit from testing?

Methods

Study Strategies Survey

All participants rated the efficacy of various study strategies, including self-testing (retrieval) and rereading (restudy).

Positive Bias (n = 26)
Participants read a short text that falsely presented retrieval practice as more effective than restudy.

Negative Bias (n = 24)
Participants read a short text that falsely presented restudy as more effective than retrieval practice.

Control (n = 24)
Participants did not read anything about retrieval practice and restudy.

Word Pairs
Participants saw 40 English-Swahili word pairs in the learning phase. Participants then reviewed the pairs in the study phase. Restudy pairs were presented again. Retrieval pairs had the Swahili word presented and participants tried to recall its English counterpart.

Texts
Participants read two short scientific texts. After reading a text, the participant moved on to the study phase. For the retrieval text, participants were told to recall everything they could remember about the text.

Final Tests
After a two-day delay, participants were tested on all the materials. For the word pairs, participants were shown the Swahili words one at a time and were asked to recall their English counterparts. For the texts, participants wrote down everything they could remember about the text.

Study Strategies Survey
Participants took the same study strategies survey in order to measure how their beliefs had changed after the experiment.

Results and Discussion

We looked at how participants’ beliefs in the efficacy of restudy and retrieval practice changed over the course of the experiment, and we observed a three-way interaction between bias, time, and study strategy. This means that the effect of time on strategy belief was influenced by bias condition. The negative bias condition reported much lower beliefs in the efficacy of retrieval practice on the post-manipulation survey, whereas those in the positive bias condition reported lower beliefs in restudy. This means that participants’ beliefs were effectively manipulated according to their condition, which supports the validity of the experimental design.

Participants’ recall rates on the final test were analyzed, and the results for the word pairs are shown on the left. No testing effects were observed in any condition, as retrieval practice recall was not significantly higher than restudy recall, and this was supported by the lack of a main effect of study condition. Bias condition also did not affect recall rates, as there was no main effect of bias condition, nor was there an interaction between bias and study condition. These findings were also true for the text results. The lack of a testing effect in the control condition was surprising given that this condition served as a replication of other studies which had found testing effects. Possible explanations and implications are discussed to the right.

One possible cause for the lack of testing effects is low recall rates in the study phase, as low retrieval practice success is associated with smaller or null testing effects (Rowland, 2014). Participants in our study had much lower recall rates compared to participants in the studies upon which ours was based. This could be due to the online nature leading to reduced participant effort, perhaps due to participants not facing any consequences (e.g., reduced payment) for poor performance or a lack of interaction with the researcher. This has important implications for education because it indicates that increased testing in online classes may not provide the same benefits that it would for in-person classes. For example, students who need to complete online quizzes for a completion grade may not put in as much effort as those completing a quiz in-person or for a grade, so they may not get a testing effect. Future research should aim to get a larger sample with more strictly sourced participants; for example, utilizing current students as participants would help increase the validity of the results.

Acknowledgments

This project was supported by a David Bray Pook Memorial Research Award and a Lindquist Undergraduate Research Grant from the Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill.


