

Review on the Testing Effect

Liuqing Yang

*School of Public Administration, Sichuan University, Chengdu, Sichuan, 610065, China
bieshiyu@gmail.com*

Keywords: Testing effect, Long-term memory, Learning strategy

Abstract: The retrieval of information from memory through testing produces learning advantages that are superior to studying alone, a phenomenon called the testing effect. As an efficient learning strategy, the testing effect has received more attention in recent years. This paper reviews the researches on the factors influencing the testing effect and points out the future research directions.

1. Introduction

The retrieval of information from memory through testing produces learning advantages that are superior to studying alone, a phenomenon called the testing effect. The study of the testing effect began with Abbott and Gates, but it didn't receive much attention at that moment[1][2]. Only after Roediger and Karpicke published their study did researchers begin to focus on the testing effect[3][4]. Subsequently, the number of research on the testing effect has grown[5].

The veracity of the testing effect has been supported by a lot of experiments. Laboratory studies have shown that tests are effective in improving long-term memory for knowledge compared to restudying[6][7]. Studies with students of different educational levels have shown that tests can similarly improve long-term memory for knowledge in real-life contexts[8][9][10].

Most studies of the testing effect have focused on the effect of testing on memory for knowledge. However, other studies have shown that the testing effect also exists in skill learning. Kromann et al. [11] randomly divided students enrolled in an in-hospital resuscitation program into a treatment group and a control group. Both groups spent four hours in class. In the treatment group, students were given three and a half hours of instruction and training, and half an hour of testing. In the control group, the four hours were used to instruction and training. Both groups were going to be tested two weeks later. The results showed that the performance of treatment group was significantly better than that of the control group. In addition, researchers studied students who took a surgical suture course, and the results also proved the veracity of the testing effect[12].

Dunlosky et al.[13] evaluated the effectiveness of various learning strategies and found that students commonly used learning strategies such as labeling and summarizing learning content, and using keywords to associate language materials, and repeat learning materials, etc. , which are very ineffective. In contrast, retrieval practice was the most effective learning strategy.

In conclusion, the testing effect is proved by existing research.

2. Influencing factors of testing effect

2.1 Feedback

Previous studies have shown that the testing effect can be influenced by feedback, and tests with feedback have better results than tests without feedback. For example, in a study by Lipko-Speed et al.[14], each participant was asked to learn 20 terms (including their names and definitions) , then the participants were showed these terms in different ways (test with feedback, test, restudy, control) . The five terms under the control condition would not appear after the students learned them. The five terms under the restudy condition would appear again for the students to restudy. For the five terms under the test condition, with only their names displayed students were asked to enter the definition of the term. The five terms in the test with feedback condition are presented as same as the test condition, unless the correct answer is presented immediately after the student has entered the definition. One week later, these students were tested. The results showed that the students' memory of the terms under the test with feedback condition was better than that under the test without feedback condition. The results of Rowland's meta-analysis also suggest that feedback can facilitate memory[14]. In addition, some research tables suggest that delayed feedback is more effective than immediate feedback[16][17].

However, the meta-analysis of Adesope shows that the testing effects was not influenced by feedback[18]. After completion of the initial test, the test with feedback does not produce a stronger testing effect than the test without feedback. The authors also suggested that this result may be related to the limitations of their meta-analysis.

In conclusion, although existing studies show that the testing effect is influenced by the feedback, a recent meta-analysis do not support this idea. Since these results may has several deficiencies identified by the authors, future studies would consider doing a meta-analysis to overcome these limitations and re-evaluate the effect of feedback on testing effect. Future research could consider another meta-analysis to reassess whether feedback can influence the testing effect.

2.2 Material

Most of the materials used in studies that support the testing effect are simple textual materials (e. g. , a list of words) . However, in reality, the materials that students learn in the classroom are usually more complex. Is there a testing effect when learning materials are more complex? It is found that the testing effect still exists in complex materials. For example, in materials based on the workings of mechanical systems[19], more integrated texts[20][21][22], spatial information[23][24],and scientific concepts that are more difficult to learn[14], the retrieval practice group all outperformed the restudying group in terms of long-term memory performance in experiments.

Some researchers have questioned such results. For example, van Gog and Sweller synthesized some previous experimental studies to argue that the testing effect decreases amd even disappears when facing more complex learning materials[25].

In conclusion, although a large number of studies have shown that the testing effect is applicable to complex materials, there are still skeptical voices. Since the forms of complex materials used in the current study are not abundant enough, future studies can use multiple forms of complex materials to investigate the applicability of the testing effect.

2.3 Test format

The classic testing effect experiment includes two phases of testing. The first stage is the initial

test, which is the test administered to students in the treatment group after they have learned the materials they were asked to learn (students in the control group were then asked to repeat the previously learned materials). The second stage is the final test, which is administered after students completed the learning task. The second stage is designed to measure the final memory effects of the treatment and control groups.

Tests are given in the form of free-recall, cued-recall, multiple-choice, recognition, and short-answer. Research has shown that all of these test formats can produce testing effect. In an experimental study[26], subjects were divided into four groups, a control group (restudying group) and three treatment groups (whose initial test forms were: free recall, cued recall, and recognition). The results showed that all three treatment groups had significantly better long-term memory effects than the control group, which means that all three test formats produced the testing effect.

Testing effect can be influenced by the form of the test. Research shows that free recall is better than recognition[26][27]. The testing effect is also affected by the difference between the form of the initial test and the final test. The results of meta-analysis show that when the initial test and the final test are in the same form, the effect is more obvious than when they are in the different form[18].

2.4 Test timing

Testing effect can be influenced by the test timing. There are two types of test timing, interspersed and at-the-end. In one class, for example, students are asked to take a test each time they finish a part of the class, while the at-the-end test is a test that students take after they have completed the lesson.

Weinstein et al. compared the effects of the interspersed and at-the-end on the testing effect[28]. In their experiment, subjects were divided into three groups and received three treatments: interspersed test, at-the-end test, and none. In the interspersed test group, subjects were given a corresponding test for each part of the materials they learned. In the at-the-end test group, subjects were given a test only after they had learned all the materials, and in the restudy group, subjects only learned without being tested. The test questions used in the interspersed test and the at-the-end test groups were identical, with the former being presented one by one and the latter being presented at once. The results showed that the memory performance of the interspersed test group was better than that of the at-the-end test group during the initial test phase. However, during the delayed testing phase, although both tested groups outperformed the restudy group in terms of memory performance, there was no longer a significant difference between the memory performance of the interspersed test group and the at-the-end test group. Since the testing effect is concerned with the long-term memory effects resulting the test, offering the test at the end of the learning phase may be a more effective timing to test than interspersed testing.

2.5 Individual differences

A large number of studies have proved the testing effect, but these studies have been conducted at the group level, meaning that groups that used tests had better long-term memory effect than groups that used restudying strategies. However, at the individual level, do all individuals who used the test have better long-term memory effects than the restudying? Research denies this idea. Brewer and Unsworth's study showed that not all those who used the test showed improved memory effects[29]. Only two-thirds of students benefit from the retrieval practice, 12% of students who used the test performed no differently from repetition, and 21% of students who used the test performed worse than repetition. Thus, the testing effect does not equally work for all students, and it may be related to individual differences.

Studies of individual differences in cognitive ability have shown that the testing effect is not affected by reading comprehension and processing speed[30]. Furthermore, it has been shown that

situational memory ability can affect the testing effect, with individuals with low situational memory ability showing stronger testing effect[29]. However, in Pan et al.'s study[31], the testing effect were not found to be influenced by situational memory ability, which is also supported by Robey's study[32]. Robey's study showed that the testing effect were not influenced by situational memory ability and general fluid intelligence, but were influenced by the learning strategies used in the learning process. Specifically, students who used less effective or no memory strategies gained more from testing than students who used effective memory strategies. The testing effect are similarly unaffected by working memory capacity[33].

Studies of individual differences in personality traits have shown that the testing effect are not influenced by cognitive need and grit [33].

Studies of individual differences in individual states have shown that the testing effect is not affected by stress[34].

In conclusion, existing studies suggest that the testing effect is not influenced by many individual differences. However, there are still relatively few studies on the relationship between testing effect and individual differences, and future studies can further explore the effects of other individual differences on the testing effect.

3. Conclusion

Through a large number of studies in the last decade or so, researchers have gained a relatively in-depth understanding of the testing effect. Research on the applicability of the testing effect has shown that the testing effect is very generalizable. Researchers have found the testing effect in different contexts (laboratory, classroom), with different test materials, test formats, and test timing. Nevertheless, there are still some questions that researchers need to answer, such as the relationship between individual differences in other areas and the testing effect.

References

- [1] Abbott, E. E. (1909). *On the analysis of the factor of recall in the learning process. The Psychological Review: Monograph Supplements*, 11(1), 159.
- [2] Gates, A. I. (1917). *Recitation as a factor in memorizing. Archives of psychology*, 6, 40.
- [3] Roediger, H. L., & Karpicke, J. D. (2006a). *Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention. Psychological Science*, 17(3), 249–255.
- [4] Roediger, H. L., & Karpicke, J. D. (2006b). *The Power of Testing Memory: Basic Research and Implications for Educational Practice. Perspectives on Psychological Science*, 1(3), 181–210.
- [5] Karpicke, J. D. (2017). *Retrieval-Based Learning: A Decade of Progress. In Learning and Memory: A Comprehensive Reference* (pp. 487–514).
- [6] Roediger, H. L., & Butler, A. C. (2011). *The critical role of retrieval practice in long-term retention. Trends in Cognitive Sciences*, 15(1), 20–27.
- [7] Roediger, H. L., & Butler, A. C. (2011). *The critical role of retrieval practice in long-term retention. Trends in Cognitive Sciences*, 15(1), 20–27.
- [8] Batsell, W. R., Perry, J. L., Hanley, E., & Hostetter, A. B. (2017). *Ecological Validity of the Testing Effect: The Use of Daily Quizzes in Introductory Psychology. Teaching of Psychology*, 44(1), 18–23.
- [9] Dobson, J. L., & Linderholm, T. (2015). *Self-testing promotes superior retention of anatomy and physiology information. Advances in Health Sciences Education*, 20(1), 149–161.
- [10] Lyle, K. B., & Crawford, N. A. (2011). *Retrieving Essential Material at the End of Lectures Improves Performance on Statistics Exams. Teaching of Psychology*, 38(2), 94–97.
- [11] Kromann, C. B., Jensen, M. L., & Ringsted, C. (2009). *The effect of testing on skills learning. Medical Education*, 43(1), 21–27.
- [12] Sennhenn-Kirchner, S., Goerlich, Y., Kirchner, B., Notbohm, M., Schiekirka, S., Simmenroth, A., & Raupach, T. (2018). *The effect of repeated testing vs repeated practice on skills learning in undergraduate dental education. European Journal of Dental Education*, 22(1), e42–e47.

- [13] Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). *Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology*. *Psychological Science in the Public Interest*, 14(1), 4–58.
- [14] Lipko-Speed, A., Dunlosky, J., & Rawson, K. A. (2014). *Does testing with feedback help grade-school children learn key concepts in science?* *Journal of Applied Research in Memory and Cognition*, 3(3), 171–176.
- [15] Rowland, C. A. (2014). *The effect of testing versus restudy on retention: A meta-analytic review of the testing effect*. *Psychological Bulletin*, 140(6), 1432–1463.
- [16] Butler, A. C., Karpicke, J. D., & Roediger, H. L. (2007). *The effect of type and timing of feedback on learning from multiple-choice tests*. *Journal of Experimental Psychology: Applied*, 13(4), 273–281.
- [17] Metcalfe, J., Kornell, N., & Finn, B. (2009). *Delayed versus immediate feedback in children's and adults' vocabulary learning*. *Memory & Cognition*, 37(8), 1077–1087.
- [18] Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). *Rethinking the Use of Tests: A Meta-Analysis of Practice Testing*. *Review of Educational Research*, 87(3), 659–701.
- [19] McDaniel, M. A., Howard, D. C., & Einstein, G. O. (2009). *The Read-Recite-Review Study Strategy: Effective and Portable*. *Psychological Science*, 20(4), 516–522.
- [20] Chan, J. C. K. (2009). *When does retrieval induce forgetting and when does it induce facilitation? Implications for retrieval inhibition, testing effect, and text processing*. *Journal of Memory and Language*, 61(2), 153–170.
- [21] Chan, J. C. K. (2010). *Long-term effects of testing on the recall of nontested materials*. *Memory*, 18(1), 49–57.
- [22] Chan, J. C. K., McDermott, K. B., & Roediger, H. L. (2006). *Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material*. *Journal of Experimental Psychology: General*, 135(4), 553–571.
- [23] Carpenter, S. K., & Kelly, J. W. (2012). *Tests enhance retention and transfer of spatial learning*. *Psychonomic Bulletin & Review*, 19(3), 443–448.
- [24] Rohrer, D., Taylor, K., & Sholar, B. (2010). *Tests enhance the transfer of learning*. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(1), 233–239.
- [25] van Gog, T., & Sweller, J. (2015). *Not New, but Nearly Forgotten: The Testing Effect Decreases or even Disappears as the Complexity of Learning Materials Increases*. *Educational Psychology Review*, 27(2), 247–264.
- [26] Carpenter, S. K., & Delosh, E. L. (2006). *Impoverished cue support enhances subsequent retention: Support for the elaborative retrieval explanation of the testing effect*. *Memory & Cognition*, 34(2), 268–276.
- [27] Rawson, K. A., & Zarny, A. (2019). *Why is free recall practice more effective than recognition practice for enhancing memory? Evaluating the relational processing hypothesis*. *Journal of Memory and Language*, 105, 141–152.
- [28] Weinstein, Y., Nunes, L. D., & Karpicke, J. D. (2016). *On the placement of practice questions during study*. *Journal of Experimental Psychology: Applied*, 22(1), 72–84.
- [29] Brewer, G. A., & Unsworth, N. (2012). *Individual differences in the effects of retrieval from long-term memory*. *Journal of Memory and Language*, 66(3), 407–415.
- [30] Karpicke, J. D., Blunt, J. R., & Smith, M. A. (2016). *Retrieval-Based Learning: Positive Effects of Retrieval Practice in Elementary School Children*. *Frontiers in Psychology*, 7.
- [31] Pan, S. C., Pashler, H., Potter, Z. E., & Rickard, T. C. (2015). *Testing enhances learning across a range of episodic memory abilities*. *Journal of Memory and Language*, 83, 53–61.
- [32] Robey, A. (2019). *The benefits of testing: Individual differences based on student factors*. *Journal of Memory and Language*, 108, 104029.
- [33] Bertilsson, F., Wiklund-Hörnqvist, C., Stenlund, T., & Jonsson, B. (2017). *The Testing Effect and Its Relation to Working Memory Capacity and Personality Characteristics*. *Journal of Cognitive Education and Psychology*, 16(3), 241–259.
- [34] Szöllösi, Á., Keresztes, A., Novák, B., Szászi, B., Kéri, S., & Racsmány, M. (2017). *The Testing Effect is Preserved in Stressful Final Testing Environment: Stress and testing effect*. *Applied Cognitive Psychology*, 31(6), 615–622.