

RESEARCH ARTICLE

Open Access



# Test-enhanced learning: analysis of an experience with undergraduate nursing students

Linda Messineo<sup>\*</sup>, Manuel Gentile and Mario Allegra

## Abstract

**Background:** This study is based on the evidence that tests can be used as an educational tool to enhance learning, not just as an evaluation tool. There is a growing body of research that shows that participating in repeated testing improves learning, a phenomenon defined as Test-Enhanced Learning. The aim of the present study was to analyse the effect of the use of a test enhanced learning program integrated into a general psychology course for undergraduate nursing students and its interaction with the students' test anxiety.

**Methods:** 161 undergraduate nursing students attending a General Psychology course followed an educational program based on Test-Enhanced Learning methodology. Students were divided into two groups, an experimental group (TEL group) and a control group (Re-study group). TEL students took a multiple-choice test on the lecture topics. The Re-study group just read study material. Testing and re-study occurred at intervals of about a week after each lesson. TEL students received feedback immediately after each test. About two weeks after the end of the lessons, all the students took a final cumulative test on all the topics. Statistical analysis was used to analyse students' performances. After the administration of the cumulative unit test, all the students took a graded examination.

**Results:** Students in the TEL group performed better than the controls, both in the final cumulative test and in a graded examination. TEL participants experienced better final cumulative test results than students not tested ( $M_{TEL} = 23.11$ ,  $M_{Re-study} = 20.47$ ,  $t(109.86) = -2.57$ ,  $p < 0.05$ ,  $r = 0.24$ ). Test-Enhanced Learning program participation has a positive impact on exam performance ( $\beta_{G\_Step1} = 0.46$ ,  $p < 0.001$ ). Finally, the analysis performed shows a slight moderating effect of test anxiety on Test-Enhanced Learning ( $\beta_{GxTA\_Step3} = 0.15$ ,  $p < 0.05$ ).

**Discussion and Conclusions:** Test-Enhanced Learning can be an effective tool for promoting and enhancing learning. In fact, taking tests after studying produced better long-term retention and then better final test performance than re-reading without testing.

Both students in the TEL group and the Re-study group with a high test anxiety level perform less well than colleagues with lower test anxiety. Nevertheless, students with higher test anxiety may obtain more benefits from participating in a Test-Enhanced Learning process than people with lower test anxiety.

Further studies on larger and more representative samples are necessary in order to investigate the effect of test anxiety on Test-Enhanced Learning.

**Keywords:** Test-Enhanced Learning, Undergraduate nursing students, Psychology education, Test anxiety

<sup>\*</sup> Correspondence: linda.messineo@itd.cnr.it  
Institute for Educational Technologies, National Research Council, CNR-ITD,  
Via Ugo La Malfa 153, Palermo 90146, Italy

## Background

In educational settings, tests are commonly used to assess students' learning. Several studies in cognitive psychology, conducted in laboratories over the past decades, have consistently shown that repeated testing on topics studied produces considerable improvement in later tests compared to re-studying the same contents [1] (for reviews, see [2, 3]). Tests can therefore represent not only an assessment tool of acquired knowledge, but can also change the process of learning and significantly improve long-term retention of contents through the promotion of the process of retrieval and retention of information. These factors facilitate learning and memorizing of the study material tested. This phenomenon is known as Test-Enhanced Learning (TEL) or testing effect. The benefits of taking a test to enhance performance in a later test, similar to the initial test, were studied under a wide range of different conditions (such as free or cued recall), using different materials (such as paired word or picture associated), different learners (such as children, young and adults), and different types of test (such as open or multiple choice tests) [4–11].

Research has also shown that, after taking a test, learners improve their ability to encode information, to reflect on and self-assess their acquired knowledge related to material studied and show an improvement in long-term retention.

Despite the large amount of research carried out in psychology laboratories, the study of the testing effect in educational settings, or with material similar to that used in educational settings, has taken place more recently.

In the last few years, studies of this procedure in real or simulated educational settings, especially in schools, have been conducted by [12–18]. Nevertheless, a better understanding of how certain psychological student characteristics (e.g. test anxiety) may affect the implementation of Test-Enhanced Learning in the classroom is needed. For example, in a laboratory setting, it was observed that the benefit of the repeated test in undergraduate students with low working memory capacity and with high test anxiety was very small [19]. Several research studies show that high levels of test anxiety generally produce a decrement in students' learning and performance [20]. Students with high test anxiety feel tense in evaluative situations [21, 22] and commonly obtain poorer test scores and grades [23–25].

In order to improve and enhance the learning process through the use of repeating testing, it is important to better understand how tests can be integrated into real educational contexts and to clarify whether test anxiety may affect retrieval practice benefits.

## Research hypotheses

The aim of the current study was to examine the effect of a TEL program on long-term retention and its

interaction with students' test anxiety in a General Psychology course for undergraduate nursing students.

The main research hypothesis was that an educational program based on the principles of TEL enhances later retention of topics studied during traditional lectures. Participation in TEL activities should lead to an improvement in learning performance both in a final cumulative test (H1) and in the graded performance in a final examination (H2).

The third hypothesis was that test anxiety has an effect on performance in the graded examination (H3).

Finally, the fourth hypothesis was that the TEL effect on the examination results is moderated by anxiety; students with high test anxiety should benefit more than students with lower test anxiety (H4).

## Methods

### Participants

According to the local ethical policy, no formal approval by the Ethics Committee was necessary. We clearly formulated an experimental design which we communicated to the Faculty board of the University of Palermo assuring that ethical standards were met, in compliance with the Declaration of Helsinki, and in accordance with the Ethical Code of the University of Palermo (Rector's Decree no. 400/2012), the Psychology Ethical Code and the Italian Personal Data Protection Code (Leg. Decree no. 196/2003). In particular, the study respected the following aspects: students were asked to provide voluntary written informed consent before participation; students could decide to interrupt their participation at any moment; personal data protection was ensured; there was no conflict of interest.

Subjects were undergraduate nursing students attending a course of General Psychology at the University of Palermo.

The original sample consisted of 201 students. Students were randomly assigned to perform different activities. One hundred and four students were assigned to the experimental group (TEL group), while the control group (Re-study group) was composed of ninety seven students. At the beginning of the TEL program, a number of students in the Re-study group (40 participants) were excluded from the sample because some data was not recorded properly due to a technical problem with the web system implemented for the research study.

The sample of students effectively involved in the study consisted of 161 students (102 females and 59 males, aged from 18 to 31 years). All the statistical analysis was carried out taking into account the unequal sample size of the groups. The following table (Table 1) reports the gender and the age distribution of the two groups.

**Table 1** Demographic characteristics of nursing students by group

	Females (n)	Males (n)	Age
Re-study group <i>n</i> = 57	34	23	<i>M</i> = 20.65, <i>SD</i> = 2.28
TEL group <i>n</i> = 104	68	36	<i>M</i> = 20.72, <i>SD</i> = 2.82
Total participants <i>N</i> = 161	102	59	<i>M</i> = 20.70, <i>SD</i> = 2.63

### Materials and design

The study involved four phases, in each of which different data were collected. The study lasted two months.

A web-platform was implemented for the experimentation. We activated different user profiles for students in the experimental group and students in the control group.

The first measurement (session 0) took place before starting the TEL program and all students were asked to complete the same questionnaire. Specifically, we measured some demographic aspects as well as test anxiety using the Test Anxiety subscale of the Motivated Strategies for Learning Questionnaire. In the second phase (sessions 1–8), TEL students participated in the TEL program taking tests, and the Re-study group read re-study material. Testing and re-study occurred at intervals of about one week after traditional teaching of a specific topic. The TEL component of the course consisted in a series of multiple-choice tests, which students in the experimental group had to take after their weekly lessons in General Psychology. The activities for the experimental group were not mandatory.

In the third phase (session 9), two weeks after the end of the lessons, all the students took a final cumulative test on the content of each lecture. In the final phase (session 10), the grades of all the students in the final examination were measured.

Finally, in order to compensate for the expected TEL benefit, appropriate procedures of compensation were designed for the students in the Re-study group (e.g. the integration of curricular activities at the end of the TEL program).

The research phases listed above are described in the following paragraphs. Table 2 shows the scheduling of the experimental design with an indication of the different conditions used in the experiment in the different sessions.

**Table 2** Experimental sessions

Group	Session 0	Session 1-8	Session 9	Session 10
Re-study	Pre-test	Re-study	Final cumulative test	Final exam
TEL	Pre-test	TEL	Final cumulative test	Final exam

The experiment spanned all the weeks of the course and continued for a further two weeks after the course for the final cumulative test.

### Phase one: Demographic aspects and test anxiety

In the first phase, all students accessing the platform for the first time completed a personal data questionnaire as well as the Motivated Strategies for Learning Questionnaire to measure test anxiety. The personal questions were concerned with basic demographic information such as gender and age. Students were also asked if they had followed other university courses and if they had already attended a general psychology course. Students who had already studied psychology were excluded from the sample. Test anxiety was measured according to the following definition: “the set of phenomenological, physiological and behavioural responses that accompany concern about possible negative consequences or failure on an exam or similar evaluative situation” [26].

To measure students’ test anxiety, we used the Italian version of the Motivated Strategies for Learning Questionnaire (MSLQ) [27, 28]. MSLQ is a self-report questionnaire developed to evaluate college student motivation and learning strategies. It consists of 81 items on a 7-point Likert-scale, and it is divided into two components. The first section, consisting of 31 items, assesses the motivational orientation of students, while the second section evaluates the students’ learning strategies. Each of the sections is composed of different scales, which can be used separately or together according to the aim of the research. One subscale of the motivational scale is the test anxiety scale and this was administered to all students, in both the experimental and control groups, during the initial phase of the experimentation. The Test Anxiety subscale rates students on two subcomponents: a cognitive component, related to negative thoughts that disturb their performance, and an emotional component that activates affective and psychological aspects. These components are the best predictors of a decrease in performance in an exam [27, 28].

### Second phase: test about the lecture

The General Psychology course in the nursing degree lasted two months and consisted of weekly traditional type lectures. The course in General Psychology aims to introduce students to the scientific study of mental and behavioural human processes. All the students participated in the psychology course covering the basic subjects in general psychology. For the experimentation, only the following eight topics were selected: attention, sensation, perception, motivation, emotions, learning, memory, cognition and intelligence. For each of these eight psychology topics, two groups of 10 multiple choice questions (with 4 possible answers) were defined.

One set of questions was used for the definition of the test used during the TEL program (experimental group) and another set of questions was used for the final cumulative test (experimental and control group). After each weekly lesson, dealing with a single issue, an email was sent to students with instructions on how to use the TEL system. Emails were different for the experimental group and the control group. The experimental group took a weekly session test on topics studied during the weekly lecture, while the control group re-read the content of the lecture. Test was active for a limited time of three days from activation, and students were required to take the test outside lesson times. Students could take a test only once, and there was a time limit for its completion. Immediately after the test, the system provided feedback on the correctness or incorrectness of all the answers. Feedback included all items with all answer options, and the answers given by the students were highlighted. If the answer was correct, the question-answer pair was highlighted in green, and a message “correct answer” appeared. If the answer was incorrect, the question-answer was highlighted in red, and the correct answer in green and the message “incorrect answer” appeared. Feedback is a crucial element in increasing the power of testing in terms of retention boosting [29, 30] and, moreover, research has shown that providing delayed feedback for all the questions at the end of the whole test, is better than giving immediate feedback after each individual answer [31–33].

The control group accessed a section in which they read question-answer pairs. Every time a re-study student accessed material, after reading the entire set of question-answer pairs, he had to click on “I have read the topic” and then concluded the study session. The learning platform was designed to allow tracking of all the students’ actions.

**Third phase: final cumulative test**

Two weeks after the end of the course (session 9), all students in both the experimental group and the control group completed a final cumulative test of eighty questions, on the eight psychology topics, similar to items used during the TEL program.

**Fourth phase: grade in the final examination**

Eight weeks after the experimentation (session 10), following the administration of the cumulative unit test, all the students took the final exam and their grades were analysed. The final exam was an oral exam, covering all the content already studied during the course.

**Results**

A Welch’s *t*-test was performed to confirm Hypothesis 1. According to Welch [34], this test is more reliable than a Student’s *t*-test when the two samples have unequal

variances and unequal sample sizes. Moreover, the effect size of the TEL was calculated.

Following Cohen & Cohen’s approach [35], moderated regression analysis was used for testing interaction effects. Hierarchical multiple regression analyses were performed to detect the main effects of testing effect and interaction effects with test anxiety as the moderator variable.

In order to test interaction effects, a multiplicative term was created for the standardised independent variables [36, 37].

The standardised independent variables were introduced into the equation in three successive steps [38].

In the first step, the testing condition was introduced to check the influence of TEL participation on the final cumulative test score as well as exam performance. Next (step 2), the moderator variable (test anxiety), and finally (step 3) the two-way interaction (testing condition x moderator) were added. The significant interaction effects would support Hypothesis 4.

However, we also took into account the main effects, given that, as Jaccard et al. [38] point out, the main effects of the independent variables generally constitute significant information. The significant effect of the moderator variable supports Hypothesis 3. The changes in *R*<sup>2</sup> in the moderator model would support Hypothesis 4, suggesting that test anxiety moderates the link between testing conditions and the final cumulative test.

In order to interpret the standardised variables a priori, unstandardized regression coefficients (B) [37] are presented in Tables 3.

The final test consisted of 10 multiple-choice questions on each of the 8 topics tested and re-studied during the TEL activity. The questions in the final test and the exam were not the same as those used in testing during the TEL program. On average, TEL participants

**Table 3** Results of the moderated regression analysis

	<i>R</i>	<i>R</i> <sup>2</sup>	Δ <i>R</i> <sup>2</sup>	B	B SE	β
Step1	0.21***	0.21				
Intercept				12.68***	0.69	
Group				5.37***	0.85	0.46
Step2	0.28***	0.27	0.07***			
Intercept				12.57***	0.66	
Group				5.55***	0.82	0.48
Test anxiety				-1.44***	0.39	-0.26
Step3	0.30***	0.28	0.02*			
Intercept				12.47***	0.65	
Group				5.62***	0.81	0.48
Test anxiety				-2.63***	0.71	-0.47
Group × Test anxiety				1.71*	0.85	0.15

Note. \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001

experienced better ( $M=23.11$ ,  $SE=0.59$ ) final cumulative test scores than non-tested students ( $M=20.47$ ,  $SE = 0.84$ ). This difference is significant,  $t(109.86) = -2.57$ ,  $p < 0.05$ ; it represents a medium sized-effect,  $r = 0.24$  (H1).

Twelve students that completed the study did not take the examination; they were excluded from the sample in order to verify the H2, H3 and H4 hypothesis. Therefore, data from 149 students were used.

The MSLQ Test Anxiety subscale consisted of 5 items ( $\alpha = 0.76$ ). The results of the regression analysis are reported in Table 3. The aim of the first model was to verify the TEL effect on the examination results.

Group ( $\beta_{G\_Step1} = 0.46$ ,  $p < 0.001$ ) was a significant predictor of exam performance. The overall model fit was  $R^2 = 0.21$  (H2).

In the second step of Cohen & Cohen's approach, Group ( $\beta_{G\_Step2} = 0.48$ ,  $p < 0.001$ ) and Test Anxiety ( $\beta_{TA\_Step2} = -0.26$ ,  $p < 0.001$ ) were significant predictors of exam performance. The overall model fit was  $R^2 = 0.27$  and also  $\Delta R^2 = 0.07$ ,  $p < 0.001$  highlighting a better fit of the second model (H3).

Finally, the overall model fit of the third model was  $R^2 = 0.28$ ;  $\Delta R^2 = 0.02$ ,  $p < 0.05$  which highlights a slight improvement. Group ( $\beta_{G\_Step3} = 0.48$ ,  $p < 0.001$ ) and Test Anxiety ( $\beta_{TA\_Step3} = -0.47$ ,  $p < 0.001$ ) and the interaction factor Group  $\times$  Test Anxiety ( $\beta_{G \times TA\_Step3} = 0.15$ ,  $p < 0.05$ ) were all significant predictors of exam performance (H4).

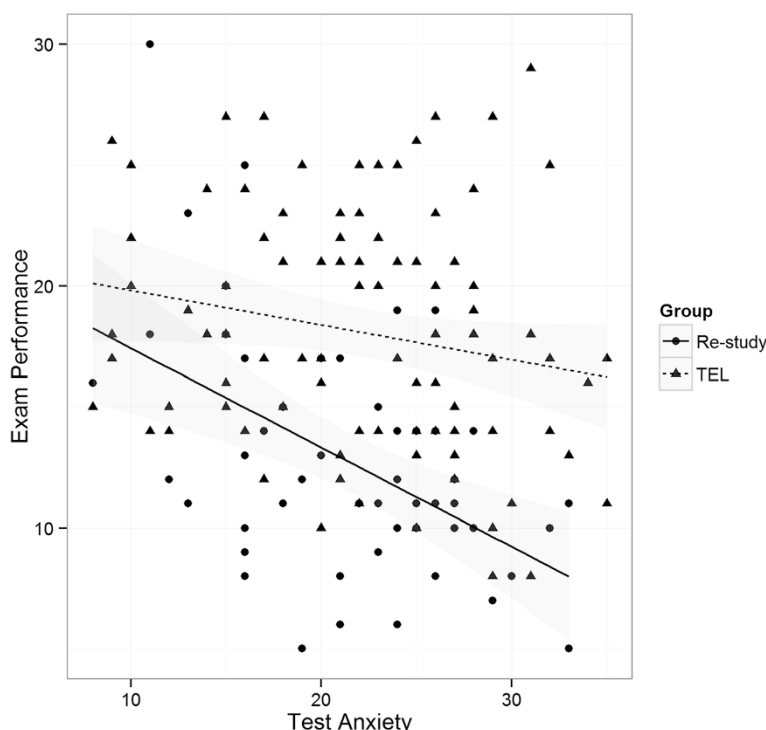
In Fig. 1, the relationship between the exam performance variable and the test anxiety variable is shown. The dots represent the position of observations from the Re-study group, while the triangles are the observations for the TEL group. For each group a regression line and its CI is reported. Both lines show a negative trend in exam performance for increasing test anxiety. Nevertheless, the slope for the TEL group relation is less steep than the slope for the Re-study group in absolute terms, suggesting a minor effect of test anxiety on TEL students.

Table 4 shows the effect of the TEL program on the students at three different levels of test anxiety (TA). Low TA group includes students with values  $<$  (mean -  $SD$ ), medium TA group includes students in the interval (mean  $\pm$   $SD$ ), and high TA group includes students with values  $>$  (mean +  $SD$ ). In the Table 4, the number of students in each condition ( $n$ ) is also indicated. Figure 2 shows the boxplots for each group.

**Discussion and Conclusions**

The results of this study demonstrate that using tests enhances learning and has a significant effect on improving long-term retention of the different topics studied.

These results are consistent with other studies carried out in this area, which have shown that testing



**Fig. 1** Relationship between the exam performance variable and the test anxiety



**Table 4** Statistics about distribution of students (Re-study and TEL) at relatively low, medium and high levels of the test anxiety

	Exam performance		
	Low TA	Medium TA	High TA
Re-study group	<i>n</i> = 14 <i>M</i> = 16.43, <i>SD</i> = 6.49	<i>n</i> = 29 <i>M</i> = 11.86, <i>SD</i> = 3.74	<i>n</i> = 10 <i>M</i> = 9.8, <i>SD</i> = 2.57
TEL group	<i>n</i> = 20 <i>M</i> = 19.05, <i>SD</i> = 4.29	<i>n</i> = 46 <i>M</i> = 18.80, <i>SD</i> = 4.94	<i>n</i> = 30 <i>M</i> = 16.23, <i>SD</i> = 5.16

has considerable benefits for memory retention. In cognitive psychology this effect is referred to as the “testing effect”. The majority of initial studies of this effect were carried out in psychology laboratories with the use of non-significant material of different kinds (e.g. verbal and visual material). Despite the large number of studies in this area, it is only recently that studies have been made into the possibility of using this procedure in education. These studies have shown that it is possible to use this methodology to improve memory in real educational settings. The testing effect is a robust effect and even with fairly

small samples, as in the case of our study, it is possible to observe the benefits of using tests.

Therefore, TEL methodologies could certainly be an effective tool for promoting and enhancing learning. In fact, taking tests after studying produced better long-term retention followed by better final test performance than re-reading without testing.

This study shows that both students in the TEL group as well as those in the Re-study group with a high test anxiety level perform less well than colleagues with lower test anxiety. Nevertheless, students with higher test anxiety may obtain more benefits from participating in a TEL learning process. A slight moderating effect of test anxiety on TEL was observed. However, further investigation is necessary in order to improve the understanding of test anxiety - TEL program interaction.

One limitation of the current study is that a large number of students in the Re-study group were excluded from the sample (40 participants) because some data was not recorded properly due to a technical hitch with the web system implemented for the research study.

The lack of measurement of test anxiety close to the final cumulative tests and the final examination represents another limitation of this study, but at the same time a way to further investigate the relationship between TEL and test anxiety.

In conclusion, new studies for better understanding how a TEL program can be integrated into real contexts are needed. In an interesting recent study, Hinze and Rapp investigated the effects of performance pressure on retrieval practice [39]. They found that only low-stake tests led to performance benefits over a re-study control group. A possible next step of particular relevance in this research area could be to replicate our study in low-stake and high-stake situations, in which, for example, performance pressure during test is induced, and investigate how test anxiety may affect retrieval practice benefits in the two different situations.

**Competing interests**

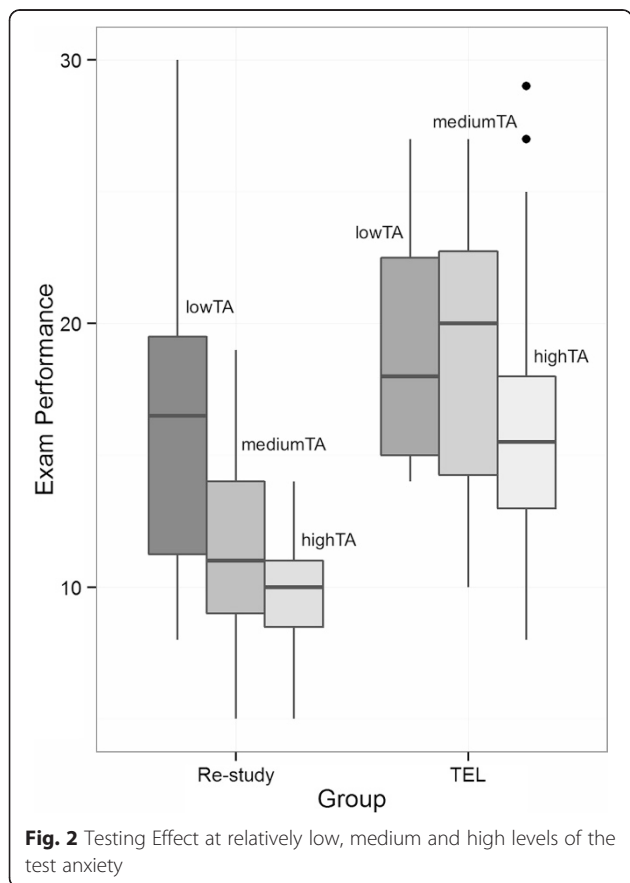
The authors declare that they have no competing interest.

**Authors’ contributions**

LM conceived of the study, designed the TEL program and the experiment, collected quantitative data, and had primary responsibility for drafting the manuscript. MG provided statistical modelling of the data and wrote the results section of the manuscript. MA contributed to the design of the study, analysis of results and drafting of the manuscript. All authors read and approved the final manuscript.

**Acknowledgements**

The authors gratefully acknowledge the collaboration of students participating in this study. They also thank Giovanni Todaro for assistance with data collection and management.



**Fig. 2** Testing Effect at relatively low, medium and high levels of the test anxiety

Received: 2 July 2014 Accepted: 14 October 2015

Published online: 24 October 2015

## References

- Carrier M, Pashler H. The influence of retrieval on retention. *Mem Cognit*. 1992;20:633–42. doi:10.3758/BF03202713.
- Roediger HL, Butler AC. The critical role of retrieval practice in long-term retention. *Trends Cogn Sci*. 2011;15(1):20–7. doi:10.1016/j.tics.2010.09.003.
- Rawson KA, Dunlosky J. Optimizing schedules of retrieval practice for durable and efficient learning: how much is enough? *J Exp Psychol*. 2011;140:283–302. doi:10.1037/a0023956.
- Izawa C. Function of test trials in paired-associate learning. *J Exp Psychol*. 1967;75:194–209.
- Hogan RM, Kintsch W. Differential effects of study and test trials on long-term recognition and recall. *J Verbal Learn Verbal Behav*. 1971;10:562–7. doi:10.1016/S0022-5371(71)80029-4.
- Wheeler MA, Roediger HL. Disparate effects of repeated testing: reconciling Ballard's (1913) and Bartlett's (1932) results. *Psychol Sci*. 1992;3:240–5. doi:10.1111/j.1467-9280.1992.tb00036.x.
- Butler AC, Marsh EJ, Goode MK, Roediger III HL. When additional multiple-choice lures aid versus hinder later memory. *Appl Cognit Psychol*. 2006;20:941–56. doi:10.1002/acp.1239.
- Karpicke JD, Roediger III HL. Repeated retrieval during learning is the key to long-term retention. *J Mem Lang*. 2007;57:151–62. doi:10.1016/j.jml.2006.09.004.
- Agarwal PK, Karpicke JD, Kang SHK, Roediger HL, McDermott KB. Examining the testing effect with open- and closed-book tests. *Appl Cognit Psychol*. 2008;22:861–76. doi:10.1002/acp.1391.
- Pyc MA, Rawson KA. Testing the retrieval effort hypothesis: does greater difficulty correctly recalling information lead to higher levels of memory? *J Mem Lang*. 2009;60:437–47. doi:10.1016/j.jml.2009.01.004.
- Butler AC. Repeated testing produces superior transfer of learning relative to repeated studying. *J Exp Psychol Learn Mem Cogn*. 2010;36:1118–33. doi:10.1037/a0019902.
- Leeming FC. The exam-a-day procedure improves performance in psychology classes. *Teach Psychol*. 2002;29:210–2. doi:10.1207/S15328023TOP2903\_06.
- Roediger HL, Karpicke JD. Test-enhanced learning: taking memory tests improves long-term retention. *Psychol Sci*. 2006;17:249–55. doi:10.1111/j.1467-9280.2006.01693.x.
- Roediger HL, Karpicke JD. The power of testing memory: basic research and implications for educational practice. *Perspect Psychol Sci*. 2006;1:181–210. doi:10.1111/j.1745-6916.2006.00012.x.
- Butler AC, Roediger HL. Testing improves long-term retention in a simulated classroom setting. *Eur J Cogn Psychol*. 2007;19:514–27. doi:10.1080/09541440701326097.
- Lyle KB, Crawford NA. Retrieving essential material at the end of lectures improves performance on statistics exams. *Teach Psychol*. 2011;38:94–7. doi:10.1177/0098628311401587.
- McDaniel MA, Roediger HL, McDermott KB. Generalizing test-enhanced learning from the laboratory to the classroom. *Psychon Bull Rev*. 2007;14:200–6. doi:10.3758/BF03194052.
- McDaniel MA, Agarwal PK, Huelser BJ, McDermott KB, Roediger HL. Test-enhanced learning in a middle school science classroom: the effects of quiz frequency and placement. *J Educ Psychol*. 2011;103:399–414. doi:10.1037/a0021782.
- Tse CS, Pu X. The effectiveness of test-enhanced learning depends on trait test anxiety and working-memory capacity. *J Exp Psychol Appl*. 2012;18(3):253–64. doi:10.1037/a0029190.
- Sub A, Prabha C. Academic performance in relation to perfectionism, test procrastination and test anxiety of high school children. *Psychol Stud*. 2003;48:77–81.
- Zeidner M. Test anxiety: the state of art. New York, NY: Plenum; 1998.
- Spielberger C, Vagg P. Test anxiety: a transactional process model. In: Test anxiety theory assessment and treatment. Washington, DC: Taylor & Francis; 1995. p. 3–14.
- Efklides A, Papadaki M, Papantoniou G, Kiosseoglou G. Effects of cognitive ability and affect on school mathematics performance and feelings of difficulty. *Am J Psychol*. 1997;110:225–58.
- Efklides A, Papadaki M, Papantoniou G, Kiosseoglou G. Individual differences in school mathematics performance and feelings of difficulty: the effects of cognitive ability, affect, age, and gender. *Eur J Psychol Educ*. 1999;XIV:57–69. doi:10.1007/BF03173111.
- Chapell MS, Blanding ZB, Silverstein ME, Takahashi M, Newman B, Gubi A, et al. Test anxiety and academic performance in undergraduate and graduate students. *J Educ Psychol*. 2005;97:268–74. doi:10.1037/0022-0663.97.2.268.
- Siebert JE, O'Neil HF, Tobias S. Anxiety, learning, and instruction. Hillsdale, NJ: New York; Toronto: Lawrence Erlbaum Associates; 1977.
- Pintrich PR, Smith DAF, García T, McKeachie WJ. A manual for the use of the motivated strategies questionnaire (MSLQ). Ann Arbor, MI: University of Michigan, National Center for Research to Improve Postsecondary Teaching and Learning; 1991.
- Pintrich PR, Smith DAF, García T, McKeachie WJ. Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). *Educ Psychol Meas*. 1993;53:801–13. doi:10.1177/0013164493053003024.
- McDaniel MA, Fisher RP. Tests and test feedback as learning sources. *Contemp Educ Psychol*. 1991;16:192–201. doi:10.1016/0361-476X(91)90037-L.
- Butler AC, Roediger HL. Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Mem Cognit*. 2008;36:604–16. doi:10.3758/MC.36.3.604.
- Butler AC, Karpicke JD, Roediger HL. The effect of type and timing of feedback on learning from multiple-choice tests. *J Exp Psychol Appl*. 2007;13:273–81. doi:10.1037/1076-898X.13.4.273.
- Metcalfe J, Kornell N, Finn B. Delayed versus immediate feedback in children's and adults' vocabulary learning. *Mem Cognit*. 2009;37:1077–87. doi:10.3758/MC.37.8.1077.
- Smith TA, Kimball DR. Learning from feedback: spacing and the delay-retention effect. *J Exp Psychol Learn Mem Cogn*. 2010;36:80–95. doi:10.1037/a0017407.
- Welch BL. On the comparison of several mean values: an alternative approach. *Biometrika*. 1951;38:330–6.
- Cohen J, Cohen P. Applied multiple regression/correlation analysis for the behavioural sciences. 2nd ed. Hillsdale, NJ: Erlbaum; 1983.
- Kleinbaum DG, Kupper LL, Muller KE. Applied regression analysis and other multivariable methods. 2nd ed. Boston: PWS-Kent; 1988.
- Aiken LS, West SG. Multiple regression: testing and interpreting interactions. Thousand Oaks, London: Sage; 1991.
- Jaccard J, Turrisi R, Wan CK. Interaction effects in multiple regression. CA: Sage: Newbury Park; 1990.
- Hinze SR, Rapp DN. Retrieval (sometimes) enhances learning: performance pressure reduces the benefits of retrieval practice. *Appl Cognit Psychol*. 2014;28:597–606. doi:10.1002/acp.3032.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

