‘Test–enhanced learning’ by Closed book examination followed by feedback in Biochemistry

Bobby, Z.[1], Nandeesha, H.[1], Thippeswamy, D.N.[1], Archana, N.[2], Prerna, S.[3], Balasubramanian, A.[4]

Abstract

Background: Test–enhanced learning in the form of Closed book examinations an unexplored formative assessment in Biochemistry among medical students.

Purpose: Assessment of the effect of providing feedback after closed book examination for graduate medical students on their understanding on the topic ‘Cancer Biology’.

Methods: A 14 multiple choice question (MCQ) based pre-test was conducted after the regular lecture classes on ‘Cancer Biology’. The MCQs were divided into Sections A and B of 7 questions each based on different learning concepts in the topic. The students were divided into two groups. In phase 1 of the study, Group 1 was subjected to a closed book examination with 7 short answer questions related to the learning concepts in Section A of the pre-test. Group 2 was subjected to a closed book examination with 7 short answer questions related to the learning concepts in Section B of the pre-test. In phase 2, a feedback was provided by the students. Post tests were carried out at the end of each of the phases.

Results: There was a significant gain from the closed book examines per se the students. The feedback further improved the gain. The gain from the second phase of the exercise (effect of feedback) was significantly higher than the gain from the first phase of the exercise (closed book examination).There was complete retention of the gain from the exercise after one week.

Conclusion: Testing effect of closed book examination combined with the feedback is an effective formative assessment for medical students in Biochemistry.

Key words: Formative assessment; Self-directed learning; Active learning; Student-centered education.

Introduction

In India, the first year graduate medical students are mostly taught by didactic lectures, where students get fewer opportunities to clarify their doubts. This will make it difficult for the students to link clinical observations later in their career with basic biochemical concepts. Formative assessments need to be at the center of graduate medical teaching protocols in order to overcome such difficulties.

In general, closed book examinations are used for assessment, which help teachers to evaluate the effectiveness of their curriculum and to assess the extent of learning by students. However, research in the field of cognitive psychology has revealed that tests can also directly enhance learning by enabling better retention of information. This observation is called the ‘Testing effect’ (Larsen *et al*., 2008). Studies have shown that repeated testing of knowledge leads to its increased retention when compared to repeated study (Roediger & Karpicke, 2006).

[1]*Department of Biochemistry, Jawaharlal Institute of Postgraduate Medical Education and Research, Puducherry, India.*
[2]*Department of Biochemistry, College of Medicine, King Khalid University Abha, Saudi Arabia.*
[3]*Department of Biochemistry, Andaman and Nicobar Islands Institute of Medical Sciences, Port Blair, A & N Islands.*
[4]*Department of Biochemistry, Government medical college, Omandurar Govt. estate, Chennai, India.*

Corresponding Author:
Dr. Zachariah Bobby,
Department of Biochemistry, Jawaharlal Institute of Postgraduate Medical Education and Research, Puducherry, India.
Email: zacbobby@yahoo.com

DOI: http://doi.org/10.4038/seajme.v12i2.48

Tests like short-answer tests which require efforts by the students to retrieve information, improve retention than tests that need only recognition, such as multiple-choice tests.

Feedback further enhances the benefits of testing by helping the students to validate their answers. As per the information from the literature, when students are subjected to assessment, it improves their memory than when they are not assessed; yet another explanation of the so called 'Testing effect' (Roediger & Karpicke, 2006; McDaniel *et al*., 2007). When students are subjected to examinations after they read materials or attend lectures, it promotes long term retention of the topics than when they undertake rereading the material. This is known as the direct effect of testing. Indirect effect of examination refers to increased duration of study made available and improved strategies of learning by the students after the announcement of test dates.

In the past we had carried out different revision exercises for the graduate medical students. We found small group discussion followed by presentation as an effective revision exercise for the graduate medical students (Bobby *et al*., 2004). In yet another study we found the formulation of short answer questions as an effective revision exercise (Bobby *et al*., 2007). Identification of mistakes and their correction by small group discussion is also an effective revision exercise for the graduate medical students in our experience (Bobby *et al*., 2014). In our experience, formulation of multiple choice questions is yet another useful revision exercise for the graduate medical students (Bobby *et al*., 2012).

In this study we assessed the direct effect of closed book examination per se and when combined with feedback as a formative assessment at the end of a teaching module in Biochemistry.

Methods

The formative assessment was carried out after a regular teaching module in 'Cancer Biology' for the first year MBBS students of our institute. The study protocol was approved by Institute Human Ethics Committee. The revision protocol was explained to the students. The students were also informed that the marks scored by them during the exercise would not be considered for their internal assessment.

Only those students who were willing to participate were enrolled. Out of the 116 students of the first year MBBS, 72 participated after giving written consent. A 14 multiple choice questions (MCQ) based pre-test was conducted after the regular lecture classes but before the revision exercise.

The MCQs were divided into Section A and B of 7 each based on different learning concepts in the topic. The students were divided into two groups after stratified randomization based on their previous internal assessment marks.

In the phase 1 of the study, Group 1 was subjected to a closed book examination with 7 short answer questions related to the learning concepts in Section A of the pre-test. Group 2 was subjected to a closed book examination with the 7 short answer questions related to the learning concepts in Section B of the pre-test. The duration of the closed book examination was one hour. The students were not allowed to refer to their text books or class notes while carrying out the exercise.

A post-test 1 was carried out with the same set of MCQs after the revision exercise. In phase 2 of the study, the students in both groups were provided with printed feedback sheets containing answer keys prepared by the teachers on their respective short answer questions separately. They were provided one hour to study the feedback.

After the feedback, post-test 2 was conducted with the same set of MCQs. The gain scores were calculated by subtracting the pre-test scores/post test scores from the post-test scores. The significance of the overall gain from the exercise at different phases was assessed by comparing the gain of batch X versus the gain of batch Y, where,

- The gain of batch X = (Gain of Group 1 in Section B + Gain of Group 2 in Section A)

- The gain of batch Y = (Gain of Group 1 in Section A + Gain of Group 2 in Section B)

Batch X represents the students who did not undertake the closed book examination and batch Y represents those students who undertook the closed book examination and feedback. The mean and SD of pre and post test scores of batch X and batch Y were also calculated for analyses.

By this experimental design, we attempted to eliminate two types of bias which could creep into the study outcome:

1) Uneven distribution of difficulty index of the learning concepts between two batches of comparison, X and Y.

This was eliminated by randomizing the learning concepts between the two batches. For this, we divided the students into two groups and the learning concepts into two sections, A and B. While Group 1 students undertook closed book examination based on the learning concepts in Section A, they did not undergo the process for the learning concepts in Section B. For the Group 2 students, the order was reversed.

2) Uneven distribution of the intellectual ability of students between the two batches of comparison, X and Y.

This was eliminated by dividing the students into two groups after stratified randomization based on their previous internal assessment marks.

The students were further classified into low, medium and high achievers based on their percentage of marks in internal assessment after the exercise in order to analyze the data. The students who scored less than 50% in their internal assessment were considered as low achievers; those with 50-65% are medium achievers, whereas the high achievers scored more than 65%. To test the retention of the gain, a delayed post test was conducted after one week with the same MCQs.

Statistical Analysis

The data is presented as mean and standard deviation. The comparison of the mean was carried out with Mann Whitney U test or Kruskal-Wallis test. A p value less than 0.05 was considered as the level of significance for all the tests. All the data were subjected D'Agostino and Pearson omnibus normality test before deciding on the appropriate statistical analyses.

Results

The marks obtained by high achievers in post-test 1 were significantly higher than their pre-test marks after the closed book examination (Table 1, Batch Y). This illustrates the direct effect of closed book examination with short answer questions. However such an observation could not be made for the low and medium achievers.

Further, the marks obtained by high achievers in post-test 2 without undertaking the closed book examination was significantly higher than their pre-test marks (Table 1, Batch X). This again illustrates the testing effect arising out of repeated testing with multiple choice questions. The overall gain from the closed book examination per se the students were significantly higher when compared to without closed book examination (Table 2, Gain 1).

Table 1: Mean and SD of Pre and Post-test marks obtained by low, medium and high achievers in Batch X (without closed book examination) and Batch Y (closed book examination followed by feedback).

Total Marks=7	Pre test	Post-test 1	Post-test 2	Delayed post-test
Batch X (Overall) (n=73)	4.15±1.56	4.05±1.58	4.63±1.45	5.26±1.33
Low achievers (n=13)	3.46±1.85	3.69±1.65	4.30±1.25	4.92±1.44
Medium achievers (n=34)	4.15±1.48	3.78±1.59	4.30±1.48	5.12±1.43
High achievers (n=26)	4.50±1.44	4.57±1.44	5.23±1.36*	5.61±1.09
Batch Y (Overall) (n=73)	4.18±1.68	4.54±1.75	6.43±0.97*	6.08±1.13
Low achievers (n=13)	3.23±1.48	3.23±1.48	5.84±1.46*	5.23±1.42
Medium achievers (n=34)	4.12±1.67	4.51±1.83	6.51±0.93*	6.21±1.11
High achievers (n=26)	4.73±1.61	5.23±1.39a	6.61±0.57*	6.34±0.79

** p<0.01 compared to pre-test and post-test 1*
[a] p<0.05 compared to pretest

Table 2: Comparison of the gain from the exercise between Batch X (without closed book examination) and Batch Y (closed book examination followed by feedback).

	Gain I (Post-test 1– Pre-test)	Gain 2 (Post-test 2– Post-test 1)	Gain 3 (Post-test 2– Pre-test)
Batch X (n=73)	-0.08±1.19	0.56±1.20	0.28±1.04
Batch Y (n=73)	0.35±1.19*	1.87±1.56**	2.23±1.50**

**P < 0.05 and **p<0.01 compared to the gain obtained by Batch X*

The effect of feedback on the questions asked in the closed book examination was significant as evident from the increment in the marks obtained by the students in Post-test 2 when compared to Post-test 1 (Table 1). Further the gain obtained from the phase 2 of the exercise as a result of feedback given to the students was significantly higher when compared to without feedback (Table 2, Gain 2). The gain from the second phase of the exercise (effect of feedback) was significantly higher than the gain from the first phase of the exercise (closed book examination) for low and medium achievers (Table 3). There was almost complete retention of the gain from the exercise after one week for all the three different categories of students as evident from the delayed post-test marks (Table 1).

Table 3: Comparison of the gains from the different phases of the exercise obtained by low, medium and high achievers in closed book examination followed by feedback (Batch Y)

	Gain I (Post-test 1– Pre-test)	Gain 2 (Post-test 2– Post-test 1)	Gain 3 (Post-test 2– Pre-test)
Low (n=13)	0.00±1.08	2.61±1.75*	2.61±1.89*
Medium (n=34)	0.39±1.24	2.0±1.62*	2.39±1.41*
High (n=26)	0.50±1.20	1.38±1.26	1.88±1.39

** p <0.01 compared to gain 1*

We collected subjective feedback from the students on the exercise. Among the students, 99% wanted to have such sessions in future. As per their opinion, the 'gain' from the exercise was 'very much' for 80%, 'little' for 18% and 'nothing' for 2%. Among the students, 94% did not feel that the exercise was a burden for them. In the Post-test 1, 61% of the students could not find any difference in the difficulty index between MCQs related to section A and section B of the closed book examination.

Thirty three percent of the students felt that the MCQs related to the section which was given to them for the closed book examination were easier. Only 6% of the students felt the other set of MCQs easier which were related to the section which was not given to them for the closed book examination. In the Post-test 2, 22% of the students did not find any difference in the difficulty index between the MCQs related to the section A and section B of the closed book examination. Seventy six percent of the students felt that the MCQs related to the section which was given to them for the closed book examination and feedback were easier. Only 2% of the students felt the other set of MCQs easier which were related to the section which was not given to them for the closed book examination and feedback.

Discussion

In education, it is a wide spread concept that taking a test assesses the contents of memory without changing them. However, studies have proven that learning process takes place when tests are conducted giving rise to the thinking that the process of retrieval improves memory (Bjork & Storm, 2011). This theory has not been evaluated in the field of medical education. In our study, we tested the idea in the field of Medical Biochemistry with the first year graduate medical students.

Our results support the 'testing effect' of closed book examination. No study had been carried out in the past to assess the concept of test-enhanced learning in Medical Biochemistry. Closed book examination per se

and with feedback is a novel revision exercise for the graduate medical students.

To explain the testing effect, several theories have been put forth over the years (McDaniel *et al*., 2007). As per one theory, the effort put in to retrieve information from memory enhances the memory and leads to better long-term retention of the information. However according to another explanation, the benefit of testing is derived from the additional exposure to the material, which is known as 'total time hypothesis'.

The total time hypothesis was subsequently found to be incorrect. Several researches have proven that taking an assessment promotes improved retention than re-studying the material for an equivalent duration (Roediger & Karpicke, 2006; A. Glover, 1989). Retrieval is the critical process that leads to improved retention in the 'Testing effect'. Retrieval during a test is an active process. The ultimate application task of knowledge gained also involves active retrieval of information. During testing there is partial practicing of the skill by active retrieval. By contrast, studying the information does not involve practicing the critical retrieval skills. In effect in testing we move towards practicing what we preach. Hence repeated testing is superior to repeated rereading of information.

In our study, we conducted only one test to test the 'test-enhanced learning'. However repeated testing has been shown to promote better retention than taking a single test (Karpicke & Roediger, 2008; Wheeler & Roediger, 1992). The benefits of repeated testing are even greater when tests are distributed over time (Karpicke & Roediger, 2007).

Another important factor that influences the efficiency of testing is the format in which the test is given. Recognition tests involve selecting the correct response from a number of presented alternatives like multiple–choice (MCQ), true/false, etc. On the contrary, the production tests require the students to construct a response which include short-answer questions (SAQ), fill-in-the-blank, essay, etc. Studies have shown that production tests have better 'Testing effect' than recognition tests. This can be explained by the fact that there is more involvement of effortful retrieval of information from memory in production tests than recognition tests (Foos & Fisher, 1988). This explains why we used SAQs as the tool for test-enhanced learning and MCQs for assessing the effectiveness of the exercise.

One important factor that greatly improves the effectiveness of testing is feedback. In the present study, feedback enhanced the retention for low, medium and high achievers. In fact the effect of feedback was more than the effect of closed book examination for low and medium achievers. Providing feedback enhances the benefits of testing by correcting errors and confirming correct responses (Butler & Roediger, 2008). The most important aspect of feedback is the content of the feedback message. We provided feedback immediately after the test. However research shows that at least in some cases, delayed feedback provides superior retention (Butler & Roediger, 2008).

The subjective feedback from the students was encouraging and reflects the effect of closed book examination as an instrument in the learning process.

Conclusion

Closed book examination followed by feedback is a novel revision exercise in Medical Biochemistry and may be tried in other medical disciplines. Tests should no longer be considered as neutral tools of measurement but rather as active instruments to aid in the acquisition and retention of knowledge.

Conflict of Interest

The authors declare no conflict of interest.

References

Bjork, E.L. & Storm, B.C. (2011) Retrieval experience as a modifier of future encoding: Another test effect, *Journal of Experimental Psychology: Learning, Memory and Cognition*, 37, 5, pp. 1113–1124.

Bobby, Z., Koner, B.C., Sen, S.K., Renuka, P., Nandakumar, D.N., Nandeesha, H., & Sathiamoorthy, R. (2004) Small group discussion followed by presentation as a revision exercise at the end of a teaching module in biochemistry, *National Medical Journal of India*, 17, 1, pp. 36–38.

Bobby, Z., Koner, B.C., Sridhar, M.G., Nandeesha, H., Renuka, P., Setia, S., & Asmathulla, S. (2007) Formulation of questions followed by small group discussion as a revision exercise at the end of a teaching module in biochemistry, *Biochemistry and Molecular Biology Education*, 35, 1, pp. 45–48.

Bobby, Z., Nandeesha, H., Sridhar, M.G., Soundravally, R., Setiya, S., Babu, M.S. & Niranjan, G. (2014) Identification of mistakes and their correction by a small group discussion as a revision exercise at the end of a teaching module in biochemistry, *The National Medical Journal of India*, 27, 1, pp. 22–23.

Bobby, Z., Radhika, M.R., Nandeesha, H., Balasubramanian, A., Prerna, S., Archana, N. & Thippeswamy, D.N. (2012) Formulation of multiple choice questions as a revision exercise at the end of a teaching module in biochemistry, *Biochemistry and Molecular Biology Education*, 40, 3, pp. 169–173.

Butler, A.C. & Roediger, H.L. (2008) Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing, *Memory & Cognition*, 36, 3, pp. 604–616.

Foos, P.W. & Fisher, R.P. (1988) Using tests as learning opportunities, *Journal of Educational Psychology*, 80, 2, pp. 179–183.

Glover, J.A. (1989) The "Testing" Phenomenon: Not Gone but Nearly Forgotten, *Journal of Educational Psychology*, 81, 3, pp. 392–399.

Karpicke, J.D. & Roediger, H.L. (2007) Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention, Journal of Experimental Psychology: Learning, *Memory and Cognition*, 33, 4, pp. 704–719.

Karpicke, J.D. & Roediger, H.L. (2008) The critical importance of retrieval for learning, *Science*, 319, pp. 966–968.

Larsen, D.P., Butler, A.C. & Roediger, H.L. (2008) Test-enhanced learning in medical education, *Medical Education*, 42, 10, pp. 959–966.

McDaniel, M.A., Roediger, H.L. & McDermott, K.B. (2007) Generalizing test-enhanced learning from the laboratory to the classroom, *Psychonomic Bulletin & Review*, 14, 2, pp. 200–206.

Roediger, H.L. & Karpicke, J.D. (2006) Test-enhanced learning: taking memory tests improves long-term retention, *Psychological Science*, 17, 3, pp. 249–255.

Roediger, H.L. & Karpicke, J.D. (2006) The Power of Testing Memory: Basic Research and Implications for Educational Practice, *Perspectives on Psychological Science*, 1, 3, pp. 181–210.

Wheeler, M.A. & Roediger, H.L. (1992) Disparate effects of repeated testing: Reconciling Ballard's (1913) and Bartlett's (1932) results, *Psychological Science*, 3, 4, pp. 240–245.