

# Post-Exam Reviews: A Consideration of Costs and Unintended Consequences

Kenneth D. Royal<sup>1</sup> · Ashley G. Henderson<sup>2</sup> · Mari-Wells Hedgpeth<sup>1</sup>

Published online: 9 June 2015

© International Association of Medical Science Educators 2015

Post-exam review is the process of reviewing an examination's items with students. The process typically involves an instructor presenting each examination item (typically a multiple-choice item) and announcing the correct answer, either as a general oral forum or by posting the exam temporarily. Discussion then ensues as to why the correct answer was indeed correct, or why it was a better answer than the alternative options. Medical educators typically conduct post-exam reviews with good intentions, as students routinely indicate they find such reviews very helpful.

In recent years, however, many institutions have reported review sessions have become quite contentious for both faculty and students, as competitive students that perceive a flaw in an item are quick to question the accuracy of their score results, and faculty that labored in preparing high-quality items are often offended by these criticisms and react defensively. Because of the potentially combustible situation, many medical educators are reconsidering the practice of post-exam reviews in their classrooms.

However, there are numerous additional reasons why medical educators may wish to reconsider this practice. The purpose of this article is to further elaborate on why post-exam reviews may not be as valuable as many medical educators (and students) might think, and illustrate many of the unintended negative consequences that might result from this practice. While there are stories abound about the pros of this practice, our intention is to equip medical educators with many of the cons surrounding this practice so that faculty

may thoughtfully decide if it is still worth pursuing in their respective classrooms.

## Validity Concerns

Items that will be used again on future examinations should only under rare circumstances be released to examinees. The exception to this rule is if the items are to be used on formative assessments, such as quizzes, where the stakes are very low for students. The problem stems from an issue of security, as “leaked” items increase the odds for cheating and poses significant threats to score and inferential validity [1]. Threats such as having pre-knowledge of test content, collusion with others, cheating aids, memorizing content, etc., all pose considerable threats to future examination scores, thus potentially invalidating those results. This could quickly have a ripple effect on a medical school's assessment system, and ultimately its reputation, as underprepared students could unjustly pass and proceed to become licensed physicians. Medical educators should ensure item content is secure and protected from unwanted exposure.

## Testing Purpose

The purpose of any examination is to measure one's knowledge, skills, abilities, or some other latent trait. However, many medical educators and students perceive all examinations also to be instructional tools. Unfortunately, this perception is misguided. Formative, low-stakes assessments that bear no real consequences for students and can be adapted in real-time (e.g., quizzes and practice examinations) may be useful instructionally [2, 3]. However, summative assessments with greater stakes (e.g., mid-term, final or placement

✉ Kenneth D. Royal  
kdroyal2@ncsu.edu

<sup>1</sup> North Carolina State University, Raleigh, NC, USA

<sup>2</sup> University of North Carolina, Chapel Hill, NC, USA

examinations) are intended to measure learning after instruction is completed [4]; thus, summative assessments generally should not be considered an instructional tool. In addition, as described below, giving answers to help with this presumed instructional tool is also not instructive.

### **Mastery Learning**

It is not a good practice to review items with students unless the assessment is formative and low-stakes in nature. The confusion regarding this topic likely emanates from the mastery learning literature. In the late 1960s, Benjamin Bloom coined the term “mastery learning” to describe an educational approach in which students are helped to master various learning concepts before proceeding to more advanced concepts [5]. Activities that allow students opportunities to learn from their mistakes, such as reviewing quizzes, offering repeat attempts at exams, etc. are typical examples of methods used in mastery learning. Further, research on mastery learning has shown the approach to be quite effective in the context of formative assessments [6]. However, applying the fundamentals of mastery learning to summative assessments (assessments that are intended to measure learning *after* instruction is complete) is generally both inappropriate and misguided.

### **Questionable Learning Experience**

It is questionable that post-exam review actually constitutes a learning experience for students. Although there is no doubt that providing the correct answer on a specific multiple choice item will increase the probability of getting that same item correct in the future, there is evidence that it will not help a person get a related question correct with any more accuracy than when the correct answer was not provided [7]. Further, even if students did receive some instructional benefit by learning a correct answer to a specific item, how long would such benefits reasonably last? Research suggests feedback is generally more helpful for correcting errors than maintaining correct responses [8]. Thus, it could be argued that post-exam review is both an inefficient and ineffective use of valuable class time, in addition to often invoking discourse.

### **Institutional Culture**

Class attendance is a significant problem at many medical schools [9]. The problem is likely to be exacerbated when students expect little new material to be presented in class. Using valuable time, reviewing examination items may discourage students from attending class and could potentially convey a message of questionable faculty/institutional values.

If an instructor’s primary purpose is to produce students with the highest scores possible, then post-exam review would likely help achieve this mission. However, most instructors are less concerned about students receiving high cosmetic marks and more concerned about students acquiring and retaining a sufficient fund of medical knowledge. Thus, post-exam review sessions may inadvertently impact the culture of a medical school in negative ways.

### **Questionable Quality Feedback**

It is undeniable that student feedback is important. However, not all feedback is created equal [10]. Simply sharing a correct answer to a single, nuanced item from a potential universe of thousands is unlikely to be very helpful for students [11]. Because items appearing on summative assessments will likely be re-used in the future, it is unwise to disclose the content verbatim to students after the assessment. Thus, a more helpful means for providing students with feedback that potentially reduces misunderstandings and error is to review substantive concepts, as opposed to specific item content. After examinations are scored, faculty can easily identify apparent gaps in content knowledge, and can subsequently devote valuable class time to providing remediation on those specific areas. This process is likely to be far more effective and efficient than reviewing items verbatim.

### **Perceived vs. Empirically Evidenced Flaws**

Many students that perceive a flaw in an item are quick to question the accuracy of their score results. In actuality, no examination is perfect. Even professional quality examinations will contain items that examinees will perceive as flawed. This is an inescapable element of testing. However, when assessments are properly conducted, only the items that have been thoroughly evaluated with regard to both content and psychometric characteristics will become a “scoreable” item. When medical educators have both content and psychometric evidence to support the quality and functioning of their items, there is no need to subject the item to students’ scrutiny and criticism. By eliminating post-exam review, faculty can eliminate potentially uncomfortable exchanges with students about examination items and instead spend valuable class time focusing on more meaningful instruction.

### **Licensure and Certification**

Students do not receive item-level feedback from various medical boards, nor will they receive such feedback throughout their careers. Testing companies go to great pains to ensure

their items are protected and secured [1]. Medical school is the perfect place to begin training students for what lies ahead as they seek medical licensure and certification.

## Expense

Finally, item production is difficult and expensive. Many educators are unaware that each exam item appearing on a high-stakes exam must be written, reviewed, edited, and psychometrically evaluated. In the professional testing industry, the monetary costs associated with one test item are approximately \$1000 [12]. Further, the process is very time-consuming and typically takes approximately 2 years from the time the item is originally produced before it becomes a “scoreable” item on an exam, assuming it ever appears at all. Medical school examination items are also difficult to write and expensive to produce. Freely exposing items that were so expensive and laborious to produce is both counterintuitive and counterproductive.

## Conclusion

The practice of offering post-exam review is very popular in medical education. Both faculty and students often perceive the practice as one that is particularly helpful for improving learning outcomes, although there is little research evidence to support this contention. Many faculty have grown weary of post-exam reviews due to potential conflicts with students. We contend there are additional reasons why medical educators may wish to further reconsider this practice. It is our opinion that the various costs associated with typical post-exam review sessions (e.g., monetary expenses, faculty time, class time, instructional opportunity losses, etc.) far outweigh any presumed benefits. Further, unintended consequences (e.g., invalid score results and inferences, increased opportunities for cheating, institutional cultural problems, etc.) potentially pose a wide array of additional problems. We encourage

medical educators to be cognizant of the costs and unintended consequences associated with traditional post-exam reviews and thoughtfully decide if it is still worth pursuing in their respective classrooms. If faculty wish to continue with post-exam review sessions, we strongly encourage faculty to modify their approach by focusing on substantive content areas in which students demonstrate weakness, as opposed to reviewing individual examination items.

## References

1. Royal K, Puffer J. Cheating: its implications for American Board of Family Medicine examinees. *J Am Board Fam Med.* 2012;25(3): 400–1.
2. Black P, William D. Assessment in classroom learning. *Assess Educ.* 1998;5(1):7–74.
3. Shepard L. The role of classroom assessment in teaching and learning. In: Richardson V, editor. *Handbook of research on teaching.* 4th ed. Washington, DC: AERA; 2001. p. 1066–101.
4. Fisher D, Frey N. *Checking for understanding: formative assessment techniques for your classroom.* Alexandria: Association for Supervision and Curriculum Development; 2007.
5. Bloom B. *Learning for mastery.* Los Angeles: University of California Press; 1968.
6. Guskey T. Closing achievement gaps: revisiting Benjamin S. Bloom’s “learning for mastery”. *J Adv Acad.* 2007;19:8–31.
7. Little J, Bjork E, Bjork R, Angello G. Multiple-choice tests exonerated, at least of some charges: fostering test-induced learning and avoiding test-induced forgetting. *Psychol Sci.* 2012;23(11):1337–44.
8. Hays M, Kornell N, Bjork R. Costs and benefits of feedback during learning. *Psychon Bull Rev.* 2010;17:797–801.
9. Millis RM, Dyson S, Cannon D. Association of classroom participation and examination performance in a first-year medical school course. *Adv Physiol Educ.* 2009;33(3):139–43.
10. Hattie J, Timperley H. The power of feedback. *Rev Educ Res.* 2007;77:81–112.
11. Marsh E, Lozito J, Umanath S, Bjork E, Bjork R. Using verification feedback to correct errors made on a multiple-choice test. *Memory.* 2012;20(6):645–53.
12. Fitzgerald C. Risk management: calculating the bottom line of developing a certification or licensure exam. 2005. From <http://www.caveon.com/articles/fitzgerald3.htm>. Accessed 17 Apr 2014.