Cheating: Its Implications for American Board of Family Medicine Examinees

Kenneth D. Royal, PhD, and James C. Puffer, MD

Cheating is undesirable and unethical, but, unfortunately, sometimes it does occur. Recent events at 3 American Board of Medical Specialties specialty boards[1,2] have illustrated that the medical certification industry is not immune from this phenomenon. Although there are numerous moral and professional implications involved with cheating, we wish to address the implications of cheating from a psychometric perspective. Our intent is to highlight some of the less obvious ways in which all American Board of Family Medicine (ABFM) diplomates possibly could be impacted should those diplomates and candidates resort to cheating on examinations.

So, what is cheating? Cizek[3] defines it as "any action that violates the rules for administering a test, any behavior that gives an examinee an unfair advantage over other examinees, or any action on the part of an examinee or test administrator that decreases the accuracy of the intended inferences arising from the examinee's test score or performance." The ABFM goes to great length to ensure a fair test for all examinees. When examinees register for ABFM exams, they make a promise to adhere to both the ethical and legal standards associated with the administration of the examination. This pact between the ABFM and the candidates minimizes the risk of a compromised examination score(s). Unfortunately, when members of either party fail to adhere to the agreed-upon standards, problems can arise.

Submitted 29 February 2012; revised 29 February 2012; accepted 1 March 2012.

From the American Board of Family Medicine, Lexington, KY.

Funding: none.

Conflict of interest: none declared.

Corresponding author: Kenneth Royal, PhD, 1648 McGrathiana Parkway, Suite 550, Lexington, KY 40511 (E-mail: kroyal@theabfm.org).

Cheating as a Threat to the Validity of Examination Scores

Validity is perhaps the most important aspect of any test.[4] The concept of validity refers to the extent to which interpretations and inferences gleaned from a score are accurate. When cheating occurs, estimates of an examinee's performance are no longer accurate. Perhaps the most obvious example of cheating as a threat to validity occurs when an individual has an undue advantage and receives a score that is higher than his or her true estimate. The inflated score essentially would be a misrepresentation of that individual's performance, thus yielding an inaccurate estimate of performance.

More subtle ways in which cheating can affect validity exist as well. The most overt threat to examination validity would be associated with the leakage of examination items. Most testing organizations, the ABFM included, possess item banks with a large pool of items readily available for inclusion on an examination. Items often vary with regard to how many times they may be used; some items are only used once, whereas others may be used perpetually provided they remain valid from a content perspective and continue to function in a psychometrically sound manner. Some overlap of items across administrations almost always exists, although the amount of overlap varies considerably across testing organizations. In any instance, examination items that are leaked from the item bank could give those with access a significant advantage. Regardless of how the test is constructed, if a single item has been compromised it could result in some examinees receiving a score that misrepresents their actual estimates of performance. Of course, the more items that are leaked, the greater the threat to the validity of the examination.

Because most high-stakes examinations are scored with some form of item response theory (IRT) methodology, the difficulty of the items plays an important role in discerning a measure of

the examinee's performance. As such, cheaters have the ability to impact the accuracy of item calibrations by making items seem easier than they actually are. Although isolated incidents of cheating would have negligible effects on these calibrations, wide-scale cheating, on the other hand, would severely affect these calibrations. In fact, the more rampant the cheating, the greater the negative consequences for all other examinees because they would in turn need to get more items correct to pass the examination. Thus, one could surmise that anyone who cheats on a high-stakes examination is not only selfishly influencing his or her own score, but is doing so at the expense of others.

The notion of item difficulty calibrations becoming altered can lead to other adverse effects. For instance, exams are typically equated, or brought onto the same scale, by using a number of common items across the exams. These common items are referred to as *item anchors*. If the items used in the anchoring process have been tainted by widespread cheating, the newly constructed test will tend to be considered easier from a measurement perspective. Under most IRT traditions, easier tests require more correct answers to pass. In the aforementioned scenario, all examinees would be affected and would need to answer more items correctly to pass the examination. With some IRT scoring methods, items are scored in such a way that credit is given (or not) based on one's response to each individual item. In instances where a particular item has been affected by inaccurate calibrations, examinees who correctly answer the question will receive less credit than they actually deserve, and examinees who incorrectly answer the question will be punished more severely as the scoring method attempts to fine-tune a performance estimate. Regardless of the scoring method used, widespread cheating in such a scenario would have the potential to impact all examinees negatively.

Deterring Cheating: A Call for Assistance

The ABFM works diligently to ensure that a fair and psychometrically sound examination is administered and that all resulting scores are valid. In addition to some of the more straightforward safeguards against cheating provided by our testing vendor and standardized examination process, our psychometric staff have a number of sophisticated methods and techniques to detect cheating. For security purposes we will not reveal the specifics of the various tools and techniques we use, but we give all examinees assurance that we work hard to ensure the accuracy of our examination results. Unfortunately, however, limitations to our means of detecting cheating exist. It is for this reason that we ask our candidates and diplomates to help ensure everyone is given a fair test so that all score results are as accurate as possible. We ask that anyone with knowledge of misconduct related to the administration of the ABFM examination report this information immediately to the ABFM Test Security Group. For more information about suspected cheating and how you may contact the ABFM, please refer to the Suspected Cheating page on our website.[5]

Conclusion

Threats to the validity of the ABFM's examination results are minimized when cheating does not occur. Any instance of cheating could generate significant consequences, not only for the examinee(s) who benefitted from the unfair advantage, but also for the honest and ethically responsible examinees that did not. The old adage that "one bad apple destroys the entire bunch" in many ways applies equally to the accuracy of information yielded from test scores. Although the overwhelming majority of family physicians conduct themselves in ethically responsible ways, we as a certification organization remain vigilant with regard to cheating, and we respectfully ask that anyone with knowledge of others who have cheated (or are planning to cheat) on ABFM examinations report this information to us as soon as possible.

References

1. Hobson K. Medical board says MDs cheated. *Wall Street Journal* 9 June 2010, A6.
2. Zamost S, Griffin D, Ansari A. Doctors cheated on exams. Updated 13 January 2012. Available from: http://www.cnn.com/2012/01/13/health/prescription-for-cheating/index.html. Accessed 28 February 2012.
3. Cizek G. An overview of issues concerning cheating on large-scale tests. Paper presented at: The Annual Meeting of the National Council on Measurement in Education, April 2001, Seattle, WA.
4. Puffer JC. Correspondence. J Am Board Fam Med 2003;16(3):467.
5. American Board of Family Medicine. Suspected cheating. Available from: https://www.theabfm.org/cert/suspected-cheating.aspx. Access 28 February 2012.