

BOARD NEWS

The Consequential Validity of ABFM Examinations

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

(*J Am Board Fam Med* 2014;27:430–431.)

Measurement scholar Samuel Messick,¹ defines validity as “an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores. . . ” (p. 13). Messick’s definition of validity differed from those of previous validity theorists in that he acknowledged that test scores often affect social policy and thus argued social consequences should be examined. Messick referred to this form of validity as “consequential validity.” Shepard^{2,3} further clarified social consequences to include both the positive/negative and intended/unintended consequences that may result from score-based inferences. The purpose of this article is to discuss consequential validity as it pertains to American Board of Family Medicine (ABFM) examinations.

To date, the ABFM has published numerous articles^{4–10} that evidence the adequacy and appropriateness of inferences based on examination scores. Many of these articles are validity studies that involve rigorous data analyses with state-of-the-art psychometric methods, whereas others advocate responsible score reporting and interpretation. Given that Messick’s¹ framework for validity also includes the social consequences that may result from score inferences, it is important to address this aspect of validity as well. Unlike other indicators of validity, consequential validity has less to do with data analysis and more to do with making inferences. Thus, the extent to which ABFM examination scores are appropriately interpreted and used depends largely on others. Our intention is to clarify some key inferences that should and should not be made about ABFM examination score results.

ABFM examinations measure a physician’s fund of medical knowledge within the context of the clinical practice of the specialty of family medicine. The examinations do not measure other important aspects of family medicine, such as one’s clinical or procedural skills, the ability to communicate with patients, professional attitudes and behaviors, the ability to practice within a system of care, and the ability to learn from the practice of family medicine to continuously improve patient care. Unfortunately, many consumers of ABFM examination score results often make inappropriate inferences about what exactly the scores mean. For example, consumers rightly infer that a passing score conferring certification is a surrogate for quality.^{11,12} Consumers also rightly infer that a passing score and subsequent certification should facilitate privileges within a hospital setting or credentials within a medical group. Unfortunately, consumers sometimes wrongly infer that a nonpassing score indicates that a physician is not worthy of being certified, and thus by extension, does not or is not capable of providing high-quality care. In addition, some consumers incorrectly infer that a higher examination score is more indicative of a better physician (compared with a physician who has a lower score); it is well understood, however, that multiple factors determine whether a physician is “good.”

It is critical that consumers understand that simply because a physician fails the Maintenance of Certification for Family Physicians (MC-FP) examination does not mean she or he is incapable of providing high-quality care or is incapable of becoming more knowledgeable about the important body of knowledge that defines the specialty of family medicine. Knowledge is fluid; thus everyone has the propensity to become more knowledgeable. In fact, over the years the ABFM staff has heard from hundreds of physicians who initially failed the MC-FP examination and who subsequently developed an improved study plan and passed on their

Conflict of interest: The authors are from the ABFM.

next attempt. Despite the initial stumble, most of these physicians continue to provide quality care to their patients today. Moreover, certification is voluntary. A number of excellent physicians practice family medicine without board certification. Thus, the lack of certification does not imply poor quality; it simply implies the physician has not evidenced his or her knowledge and commitment to continuous improvement by way of a formal certification process.

While fully aware that an examination in and of itself is unable to provide sufficient information about the quality of a physician, the ABFM, along with all American Board of Medical Specialties member boards, adopted a more comprehensive approach to assessing physician performance in 2000. This new paradigm, called Maintenance of Certification, assesses 6 general competencies: professionalism, medical knowledge, communication and interpersonal skills, patient care, systems-based practice, and practice-based learning and improvement. These are assessed by the ABFM within a 4-part construct that (1) assesses professionalism, licensure, and personal conduct; (2) measures the ability of the physician to self-assess and develop a program of life-long learning; (3) assesses by examination cognitive expertise; and (4) assesses the physician's performance in practice and the ability to develop mechanisms to continuously improve quality based on the assessment. We would argue that this expanded approach to physician assessment provides additional information from which appropriate inferences can be made about the quality of care that a physician delivers and has far greater consequential validity within the construct as defined by Messick.¹

Conclusion

Empirical data analyses with rigorous research methodologies are critical for providing evidence that an examination is functioning well and measuring the intended construct. The ABFM has produced a considerable body of research that evidences the accuracy and trustworthiness of the score results produced by its examinations. Similarly, the ABFM has continually emphasized that the purpose of the examination is to measure a physician's fund of medical knowledge in clinical

family medicine and has emphasized appropriate and responsible score interpretations. Unfortunately, some consumers continue to attach additional meaning to these score results that can affect a physician in unintended ways. To preserve the integrity of the score inferences and their impact on physicians, it is important that all consumers of ABFM examination score results make appropriate and responsible inferences about what exactly the scores mean.

References

1. Messick S. Validity. In: Linn RL, ed. *Educational measurement*. 3rd ed. New York: Macmillan Publishing Co; 1989:13–103.
2. Shepard LA. Evaluating test validity. *Rev Res Educ* 1993;19:405–50.
3. Shepard LA. The centrality of test use and consequences for test validity. *Educ Meas Issues Pract*. 1997;16:5–24.
4. O'Neill TR, Royal KD, Puffer JP. Performance on the American Board of Family Medicine certification examination: are superior test taking skills alone sufficient to pass? *J Am Board Fam Med* 2011;24:175–80.
5. Royal KD, Puffer JC. Understanding the “sum of subtest to overall score discrepancy” on the Maintenance of Certification for Family Physicians examination. *J Am Board Fam Med* 2012;25:260–1.
6. Royal KD, Puffer JC. The reliability of American Board of Family Medicine examinations: implications for test-takers. *J Am Board Fam Med* 2012;25: 131–3.
7. Royal KD, Puffer JC. Dimensionality of the Maintenance of Certification for Family Physicians examination: evidence of construct validity. *J Am Board Fam Med* 2013;26:342–4.
8. Royal KD, Puffer JC. Cheating: its implications for ABFM examinees. *J Am Board Fam Med* 2012;10: 274–5.
9. Royal KD, Puffer JC. Criterion-referenced examinations: implications for the reporting and interpretation of examination results. *J Am Board Fam Med* 2013;26:225–6.
10. Royal KD, Puffer JC. A closer look at recertification candidate pass rates. *J Am Board Fam Med* 2013;26: 478–9.
11. Puffer JC. The American Board of Family Medicine certification examination: a proxy for quality. *Fam Med* 2011;43:436–7.
12. O'Neill TR, Puffer JC. Maintenance of certification and its association with the clinical knowledge of family physicians. *Acad Med* 2013;88:780–7.