

REVIEWS

Progress Examinations in Pharmacy Education

Cecilia M. Plaza, PharmD, PhD

Director of Academic Affairs and Assessment, American Association of Colleges of Pharmacy

Submitted January 2007; accepted March 20, 2007; August 15, 2007.

Interest in the use of the progress examination has grown in the current culture of accountability in higher education. The Accreditation Council for Pharmacy Education's (ACPE's) Standards 2007 calls for comprehensive, knowledge- and performance-based examinations as part of a school or college of pharmacy's evaluation and assessment of student learning. Progress examinations have been used primarily in medical education. The purpose of this paper is to provide a brief overview of the literature on progress examinations and considerations for their potential use within an effective assessment plan.

Keywords: progress examinations, assessment

INTRODUCTION

Anderson and colleagues suggested that assessment should have the dual purpose of accountability and the improvement of student learning as well as the 3 overarching characteristics, "... (1) it consists of a systematic and continuous process; (2) it emphasizes student learning, with the cornerstone being what students can do; and (3) it focuses on the improvement of educational programs."¹ Assessment measures should be used for continuous curricular improvement and employ a variety of valid and reliable systematic measures that are both summative and formative in nature.²⁻³ Outcomes assessment involves collecting information on a desired outcome and comparing it to previously established mission statements, goals, and objectives.⁴ Outcomes assessment is a critical component of accreditation in that it has strengthened accreditation and accreditation has likewise sustained the assessment movement, thus creating a synergistic relationship.⁵ There has been an increased focus on the accountability component of assessment and accreditation in higher education as evidenced by the publications from the US Department of Education and the National Association of State Universities and Land-Grant Colleges (NASULGC).⁶⁻⁸ Within this increased culture of accountability, there has been a push for greater standardization, and within this movement towards standardization is the issue of the use of progress examinations. ACPE Standards 2007, Guideline 15.1 states that a school or college of pharmacy's evaluation of student

learning should, "... incorporate periodic, psychometrically sound, comprehensive, knowledge-based and performance-based formative and summative assessments, including nationally standardized assessments... that allow comparisons and benchmarks with all accredited and peer institutions."² While there are several key foundational resources available within pharmacy education that address the issue of assessment, there is little within these resources on the use and implications of progress examinations.^{3,4,9-11} The purpose of this paper is to provide an overview of the literature on progress examinations within the health sciences and considerations for their potential use within an effective assessment plan.

A progress examination can be defined as a method of assessing both the acquisition and retention of knowledge at one or more points in the curriculum relative to curricular goals and objectives.¹² Progress examinations have been suggested for a variety of uses: peer-comparisons among schools (also referred to as benchmarking in the educational literature), comparisons among students to identify those who could benefit from remediation, formative assessment as part of an overall assessment plan, high-stakes assessment to determine progression in the curriculum, low-stakes assessment, and as an adjunct to program evaluation.

OVERVIEW OF THE LITERATURE

In undertaking any assessment endeavor, such as progress examinations, there are numerous elements that must be considered, chief of which are reliability and validity.¹³⁻¹⁴ A glossary of terms related to reliability and validity is provided in Appendix 1. While Standards 2007 have brought the issue of progress-type examinations back into the forefront, the Basic Pharmaceutical Sciences Exam (BPSE) developed by Pharmat, Inc, was

Corresponding author: Cecilia M. Plaza, PharmD, PhD.
American Association of Colleges of Pharmacy 1426 Prince Street, Alexandria, VA 22314 Tel: 703-739-2330, ext 1018.
Fax: 703-836-8982. E-mail: cplaza@aacp.org

used by some schools and colleges of pharmacy in the 1980s as a progress examination prior to entry into the final year of the curriculum.¹⁵ The BPSE which covered pharmacology, toxicology, pharmaceuticals, and medicinal chemistry, was designed to measure pharmacy student knowledge relative to both the content and goals in the preclinical curriculum.¹⁵ Fassett and Campbell examined the BPSE as a potential predictor of student performance during clinical training but found there was no correlation between scores on the BPSE and performance in the “clinical” coursework or with performance on experiential rotations.¹⁵ However, the authors concluded that the BPSE appeared to serve as a comprehensive evaluation of basic science knowledge in that there was a strong correlation with BPSE scores and basic pharmaceutical sciences course grades (Spearman’s rho = 0.75, $p = 0.001$) and therefore could be used for national comparison among institutions in aggregate. The BPSE fell out of use during the 1980s and is no longer available for use by colleges and schools of pharmacy.

Kirschenbaum and colleagues examined programmatic curricular outcomes assessment at schools and colleges of pharmacy in the United States and Puerto Rico and found that 4 institutions used an end-of-semester comprehensive written examination that did not affect course grades, 13 used a “high-stakes” end-of-year examination, and 17 used a “low-stakes” examination to assess curricular outcomes ($N = 68$).¹⁶ Locally developed examinations have been used at individual schools and colleges pharmacy such as the Milemarker Assessment at the University of Houston, which is a case-based, multiple-choice progress examination administered at the conclusion of each didactic year in the professional curriculum.¹⁷

Use of Progress Examinations

Progress examinations have been primarily used in medicine to assess knowledge. Newble and Jaeger examined the effect of assessments and examinations on learning in medical students at the University of Adelaide in Australia following a change in their curriculum to include both an experiential component in the curriculum and low-stakes clinically based assessments in the final year that did not influence progression in addition to the existing year-end high stakes multiple-choice examinations.¹⁸ The year-end high-stakes multiple-choice examination served as the basis for pass or fail decisions in the final year of the curriculum. They cautioned that, “Should the examination system be seen by the students to require predominantly recall of factual information then they will tend to adopt a surface-level or rote-learning approach.”¹⁸ Based on questionnaires sent to recent gra-

duates, they found that the study habits of students did indeed change to conform to the format of the year-end multiple-choice progress examination (ie, attempting to study for the examination through memorization) which had items written at the recall level of knowledge rather than spending time in clinical experiences offered to students. This was found to be an unintended consequence of using the high-stakes examination, which in essence competed with a low-consequence but high-learning situation of clinically-based assessments. A mismatch between the educational objectives and the operationalization of the assessment program can result in a hidden curriculum based on the assessment.¹⁸⁻¹⁹

Blake et al examined the psychometric properties of the progress examination introduced into McMaster University’s medical curriculum in Canada as well as the effect on student learning.²⁰ The progress examination was intended primarily for formative student self-assessment rather than as a peer-comparison measure with other schools or as a high-stakes assessment. The progress examination consisted of the same 180 multiple-choice items administered 3 times per year to all classes. To assess the psychometric properties of the progress examination, the researchers examined the reliability across multiple administrations as well as construct validity. The reliability estimates increased the more times a student had taken the examination suggesting a potential test-retest threat to internal validity. The stated overriding construct of the progress examination was that it was, “. . .capable of demonstrating consistent progress over time, reflecting increased knowledge of students.”²⁰ However, this was based on the assumption that the total score really represented knowledge versus increased test-taking savvy or other sources of construct-irrelevant variance. A 6-item questionnaire was administered to students to ascertain their perceptions on how, if at all, the progress examination affected their approach to learning. Students gave a low rating to the progress examination as a factor in changing their approach to learning.

Van der Vleuten et al described the use of a progress examination designed specifically for problem-based learning curriculum at Maastricht medical school in the Netherlands.¹⁹ They suggested that progress examinations should share some common features: (1) the material covered is sufficiently comprehensive so that students cannot study specifically for the examination; and (2) student assessment is based on successive overall performance on progress examinations rather than performance on a single examination. The examination consisted of 250 true/false items written by faculty members and included an additional selection choice of “I don’t know” intended to minimize guessing. A new examination was

constructed with items compiled from faculty members, for each of the 4 administrations during an academic year. The examination was administered to all students regardless of year in the curriculum. To compare across examinations, scores were expressed as percentages since the selection of the "I don't know" option did not count in the calculation of the total score which was calculated based on the number correct minus the number incorrect. Since the expression of scores as a percentage is not test equating in the measurement sense, it was not possible to determine whether the various iterations of the examination were truly equivalent in terms of difficulty and content. In a related study, researchers at Maastricht medical school attempted to establish the potential role and value of knowledge-based progress examinations in medical education and examined the convergent validity of the progress examination relative to clinically based assessments.²¹ Acknowledging the dissonance between the level of learning addressed in a progress examination (knowledge/recall) and the educational goals or instructional methods of the curriculum, the progress test correlated with a clinical reasoning test.

Remmen and colleagues suggested that a written progress examination could be used as an alternative to a more time- and personnel-intensive performance-based Objective Structured Clinical Examination (OSCE).²² This study compared the performance of 106 medical students on a 132-item true/false/"don't know" progress examination with their performance on an OSCE consisting of 12 stations that were each 13 minutes in duration and staffed by medical school faculty members. The OSCE was graded using both a checklist and an overall global rating. Both assessments covered basic physical diagnostic and therapeutic skills. The correlation between the performance on the written progress examination (broken down by school) and the performance on the OSCE (broken down by grading method) ranged from 0.35 to 0.48 which the researchers then corrected for attenuation with a resulting range of 0.64 to 0.87. No *p* values were reported for any of the correlations between the 2 types of assessments. The authors suggested that the reported reliabilities from the written progress examination of 0.32 for Antwerp students and 0.58 for Ghent students were too low for summative assessment but could be improved through adding more items. The Spearman Brown Prophecy formula reveals that the effect of doubling the number of items to 264 items only predicts an increase in the reliability from 0.32 to 0.48 for 1 of the schools. Thus, the relative gains in reliability by increasing the number of items tends to be modest. The results of this study also need to be considered in light of their small sample size.

The use of progress examinations in combination with OSCEs for formative student self-assessment has been described as an adjunct to program evaluation at the Ohio University College of Osteopathic Medicine.¹² The 6-hour progress examination consisted of 330 multiple-choice items with 5 response choices administered twice annually across the entire 7-year curriculum. While all students in all years take the same examination each time it is administered, a new set of examination items are used for each administration. The OSCE, administered only to fourth-year medical students, was not used for formal grading or advancement determination purposes. At the time of publication, the progress examination has been administered twice and the OSCE once. They concluded that using the 2 different forms of assessment together provided more valuable feedback to students as well as to the institution, with the written progress examination providing information on accumulated knowledge and the OSCE providing information on accumulated clinical and interpersonal skills.

Researchers at Utrecht medical school in the Netherlands proposed a progress test using short answer questions as an alternative to true/false examinations.²³ The short answer examination consisted of 40 cases with both clinical and basic science components administered 3 times a year during the final 2 years. The examination was based on the concept of mastery learning where the goal is to have all students achieve some predetermined level, in this case a score of 80% on at least 3 separate administrations of the examination. The examination was developed using a blueprint to establish content validity and then assessed for face validity using a committee of experts. Since the examination was based on mastery learning, students could sit for the examination as many times as necessary over the 6 possible administrations until mastery was achieved. The primary disadvantage to this form of progress testing is the time required to grade the responses and develop psychometrically sound items.

Use of Progress Examinations for Benchmarking

Medical schools in the Netherlands have also been involved in using progress examinations in benchmarking among institutions as well as collaboration among institutions in the development of the examinations.²⁴⁻²⁵ The Maastricht Progress Test was piloted as a method of international benchmarking across medical schools in Europe.²⁴ The Maastricht Progress test consisted of 250 true/false/don't know items administered 4 times annually.^{19,24-25} For international benchmarking, the method of administration varied among schools so that for some students the examination was voluntary without

any credit awarded, for other students the examination was voluntary but credit was awarded, and for yet other students the examination was mandatory.²⁴ This led to construct-irrelevant variance, limiting the interpretation of the comparisons. Led by the Maastricht medical school, a partnership was established among 3 medical schools, with a fourth medical school purchasing the examination for the development and administration of a progress examination to achieve some economy of scale.²⁵ The authors noted that, unlike the United States, the Netherlands does not have a national medical licensure examination so this collaboration could be considered a partial move towards a national examination, but one where schools themselves retained ownership and control over the process.²⁵ A centralized review committee that included students was established to evaluate test items with the goal of each institution contributing 300 items. It took several test cycles before there was parity among the institutions in the contribution of usable items. Plans for the future included moving away from a true/false format in favor of multiple-choice in addition to less emphasis on factual knowledge as well as the need for expertise in methodologies to equate tests across administrations since differences in difficulty levels was a concern. While not mentioned in this paper, item response theory could be used to place different forms of an examination on the same scale to allow for comparisons among students of the similar overall ability such as those in the same professional year (horizontal equating) and students of different abilities such as those in different professional years (vertical equating).²⁶

DISCUSSION

There are several potential advantages to using progress type examinations. One advantage is the opportunity to have students review material more often (eg, reviewing materials prior to progressing to the next professional year with progress examination given annually).²⁷ A related potential advantage is that progress examinations can provide implicit emphasis on the cumulative nature of pharmacy education.²⁷ Progress examinations can also help colleges and schools identify students requiring remediation when used for formative assessment, as well as for assessing the curriculum overall.²⁷

While progress tests have many potential advantages in an assessment plan, they are not without disadvantages. One major concern with the use of progress examinations is the issue of consequential validity: the consideration of the potential effect the assessment in question can have on learning.^{18-19,21} Knowledge-based examinations written at the factual or recall level of learning can potentially encourage memorization over higher levels of learning as

students are not required to integrate, apply, synthesize, or evaluate information. There is also the potential to create a disconnect between having a knowledge-based examination while promoting lifelong learning. Whether a sufficiently comprehensive examination prevents studying to the examination is debated.

There are numerous sources of construct-irrelevant variance such as the difficulty in controlling conditions of administration and differences in curricular sequencing across institutions.²⁸ Controlling conditions of administration is of special concern when the progress examination is being considered for use in benchmarking with other institutions when there are different consequences for students at various institutions, which will affect how a student approaches the examination (eg, credit awarded for participation in the examination versus an examination required for progression). Different curricular sequencing such as year round versus traditional semester programs can also affect the ability to use the examination for benchmarking across institutions, reflecting opportunity to learn issues and potential recency effects.

Item format also needs to be considered with any examination administration since there are advantages and disadvantages with each item type.²⁹⁻³⁴ True-false items, for example, tend to focus exclusively on recall of factual knowledge, can potentially reinforce retention of false information that is then difficult to unlearn, and are highly prone to guessing.³¹ There is debate over the value of attempting to control for guessing as potential gains in reliability and validity are modest and are based on the assumption that guessing is “blind” which is often not the case.³⁵ Numerous additional questions should also be considered prior to implementing a new assessment plan as presented in Table 1.³³⁻³⁴ Each type of validity evidence should be considered prior to the implementation of a progress examination as part of an assessment plan.

As the Academy considers the use of progress examinations for assessment purposes, the emphasis these examinations receive within an assessment plan is of particular concern. Making a progress examination high-stakes could be seen as a move towards a United States Medical Licensing Examination (USMLE) Step 1 type of examination. The USMLE Step 1 examination, sponsored by the Federation of State Medical Boards and the National Board of Medical Examiners, is used by medical schools to assess the first 2 years of the medical curriculum which concentrate on the basic sciences and a passing score is required for progression.³⁶ The USMLE Step 1 examination has been criticized as dictating the content of the first 2 years of medical school curricula so that students are prepared for the material on the examination, thus limiting the amount of innovation possible in

Table 1. Different Types of Validity Evidence for Educational Assessments*

Type of Evidence	Relevant Questions
Content representativeness <i>and</i> relevance	<ul style="list-style-type: none"> ● How well is important content in the domain represented? ● How well does the assessment task represent the defined curriculum? ● How well does the assessment task reflect current thinking about what should be learned? (ie, is it worth it?)
Types of thinking skills and processes required	<ul style="list-style-type: none"> ● How well does the assessment represent the thinking skills desired in the curriculum? ● Are the thinking skills addressed in the assessment procedure the intended thinking skills? (eg, recall versus critical thinking)
Internal structural evidence	<ul style="list-style-type: none"> ● Do assessment tasks each contribute meaningful information to assess the desired outcome? ● Do assessment tasks that are intended to provide different information indeed provide unique information?
External structural evidence	<ul style="list-style-type: none"> ● Are results consistent with similar assessments of students? ● How well does performance predict current or future performance? (predictive validity)
Reliability evidence (as it contributes to validity)	<ul style="list-style-type: none"> ● Test-retest (if appropriate) ● Interrater agreement (if appropriate) ● Equivalence, alternate form reliability (if appropriate)
External validity (generalization evidence)	<ul style="list-style-type: none"> ● Do results differ when given to students with different backgrounds but the same ability? ● Do results change with special instruction or coaching and if so should the interpretation of the results be changed?
Consequential validity (value of the intended and/or unintended consequences)	<ul style="list-style-type: none"> ● What are the consequences (both good and bad) to students if the assessment is interpreted and used in a particular way? ● What are potential unintended consequences of the planned interpretation and use of the assessment?
Practicality evidence (cost, efficiency, and instructional features)	<ul style="list-style-type: none"> ● Do the assessment results explain individual differences in a meaningful way? ● Would an alternative assessment be more efficient? ● Do teachers agree with the theoretical concepts in the assessment procedure as reflecting what students are learning in the classroom? ● Does the assessment provide timely results to guide instruction? ● What costs are associated with the assessment (eg, time, money, unintended consequences)?

*Adapted from Nitko³³ and Messick³⁴

curricular structure. The cost associated with a multi-stage licensure process similar to that used in medicine would also be a major consideration of the ability of students to shoulder an increased financial burden. Use of multiple assessments such as knowledge/content/recall examinations in addition to clinical skills evaluations such as OSCEs may collectively provide more information on student learning, educational methods, and institutional comparisons. The use of multiple assessments must be balanced with regard to the benefit of the infor-

mation provided versus the increased cost of administering different assessment. The use of multiple assessments must be balanced: the benefit of the information provided must be weighed against the increased cost of administering different assessments.

ACKNOWLEDGEMENTS

The author would like to thank Dr. Ken Miller at AACP for his invaluable feedback and review of this paper.

The ideas expressed in this manuscript are those of the primary author and do not represent those of the American Association of Colleges of Pharmacy.

REFERENCES

1. Anderson HM, Anaya GA, Bird E, Moore DL. A review of educational assessment. *Am J Pharm Educ.* 2005;69:Article 12.
2. Accreditation Council for Pharmacy Education. Accreditation Standards and Guidelines for the Professional Program in Pharmacy Leading to the Doctor of Pharmacy Degree. The Accreditation Council for Pharmacy Education Inc. Available at: http://www.ACPE_Revised_PharmD_Standards_Adopted_Jan152006.pdf. Accessed on March 10, 2006.
3. Abate MA, Stamatakis MK, Hagggett RR. Excellence in curriculum development and assessment. *Am J Pharm Educ.* 2003;67:Article 89.
4. Hollenbeck, RG. Chair report for the academic affairs committee. *Am J Pharm Educ.* 1999;63:7S-13S.
5. Wright BD. Accreditation and the scholarship of assessment. In: Banta TW, ed. *Building a Scholarship of Assessment*. San Francisco: Jossey-Bass; 2002.
6. U.S. Department of Education. *A Test of Leadership: Charting the Future of U.S. Higher Education*. Washington, DC, 2006.
7. McPherson P, Shulenburg D. Improving student learning in higher education through better accountability and assessment. National Association of State Universities and Land-Grant Colleges. Available at: http://www.nasulgc.org/Accountability_DiscussionPaper_NASULGC.pdf. Accessed on December 4, 2006.
8. McPherson P, Shulenburg D. Elements of accountability for public universities and colleges. National Association of State Universities and Land-Grant Colleges. Available at: http://www.nasulgc.org/Accountability_DiscussionPaper_Revised_NASULGC.pdf. Accessed on December 4, 2006.
9. Boyce EG, Maldonado WT, Murphy NL, et al. Building a process for program quality enhancement in pharmacy education: report of the 2003-04 academic affairs committee. *Am J Pharm Educ.* 2004;68:Article S7.
10. Boyce EG. A Guide for Doctor of Pharmacy Program Assessment. American Association of Colleges of Pharmacy. Available at: [http://www.aacp.org/Docs/MainNavigation/Resources/5416\)pharmacyprogramassessment_forweb.pdf](http://www.aacp.org/Docs/MainNavigation/Resources/5416)pharmacyprogramassessment_forweb.pdf). Accessed on December 4, 2006.
11. *Handbook on Outcomes Assessment*. Alexandria, Virginia: American Association of Colleges of Pharmacy; 1995.
12. Portanova R, Adelman M, Jollick JD, Schuler E, Ross-Lee B. Student assessment in the Ohio University College of Osteopathic Medicine CORE system: progress testing and objective structured clinical examinations. *J Am Osteopathic Assoc.* 2000;100:707-12.
13. Gray PJ. The roots of assessment: tensions, solutions, and research directions. In: Banta TW, ed. *Building a Scholarship of Assessment*. San Francisco: Jossey-Bass; 2002.
14. Allen MJ, Yen WM. *Introduction to Measurement Theory*. Prospect Heights, Ill: Waveland Press; 1979.
15. Fassett WE, Campbell WH. Basic pharmaceutical science examination as a predictor of student performance during clinical training. *Am J Pharm Educ.* 1984;48:239-42.
16. Kirschenbaum HL, Brown ME, Kalis MM. Programmatic curricular outcomes assessment at colleges and schools of pharmacy in the United States and Puerto Rico. *Am J Pharm Educ.* 2006;70:Article 08.
17. Sangsiry SS, Nadkarni A, Lemke T. Perceptions of PharmD students towards a cumulative examination: the Milemarker process. *Am J Pharm Educ.* 2004;68:Article 93.
18. Newble DI, Jaeger K. The effect of assessments and examinations on the learning of medical students. *Med Educ.* 1983;17:165-71.
19. van der Vleuten CPM, ver Wijnen GM, Wijnen WHFW. Fifteen years of experience with progress testing in a problem-based learning curriculum. *Med Teach.* 1996;18:103-9.
20. Blake JM, Norman GR, Keane DR, Mueller B, Cunningham J, Didyk N. Introducing progress testing in McMaster University's problem-based medical curriculum: psychometric properties and effect on learning. *Acad Med.* 1996;71:1002-7.
21. Boshuizen HPA, van der Vleuten CPM, Schmidt HG, Machiels-Bongaerts M. Measuring knowledge and clinical reasoning skills in a problem-based curriculum. *Med Educ.* 1997;31:115-21.
22. Remmen R, Scherpbier A, Denekens JO, et al. Correlation of a written test of skills and a performance based test: a study in two traditional medical schools. *Med Teach.* 2001;23:29-32.
23. Rademakers J, Ten Cate TJ, Bar PR. Progress testing with short answer questions. *Med Teach.* 2005;27:578-82.
24. Albano MG, Cavallo F, Hoogenboom R, et al. An international comparison of knowledge levels of medical students: the Maastricht Progress Test. *Med Educ.* 1996;30:239-45.
25. van der Vleuten CPM, Schuwirth LWT, Muijtjens AMM, Thoben AJNM, Cohen-Schotanus J, van Boven CPA. Cross institutional collaboration in assessment: a case on progress testing. *Med Teach.* 2004;26:719-25.
26. Bond TG, Fox CM. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Mahwah, NJ: Lawrence Erlbaum Associates, 2001.
27. Ryan GJ, Nykamp D. Use of cumulative examinations at U.S. schools of pharmacy. *Am J Pharm Educ.* 2000;64:409-12.
28. Haladyna TM, Downing SM. Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice.* 2004;23:17-27.
29. Holt GA, Holt KE. Teacher-made exams: part 1. *J Pharm Teaching.* 1990;1:33-40.
30. Holt GA, Holt KE. Teacher-made exams: part 2. *J Pharm Teaching.* 1990;1:69-81.
31. Holt GA, Holt KE. Teacher-made exams: part 3. *J Pharm Teaching.* 1990;1:55-73.
32. Holt GA, Holt KE. Teacher-made exams: part 4. *J Pharm Teaching.* 1991;2:59-83.
33. Nitko AJ. *The Educational Assessment of Students*, 4th ed. Englewood Cliffs, NJ: Prentice-Hall; 2003.
34. Messick S. Standards of validity and the validity of standards in performance assessment. *Educational Meas: Issues Pract.* 1995;Winter: 5-8.
35. Nunnally JC, Bernstein IH. *Psychometric Theory*, 3rd ed. New York, NY: McGraw-Hill, Inc, 1994.
36. United States Medical Licensing Examination. Federation of State Medical Boards, Dallas, Tx, and the National Board of Medical Examiners, Philadelphia, Pa. Available at: <http://www.usmle.org>. Accessed January 23, 2007.

Appendix 1. Glossary*

Term	Definition
construct validity	In the broadest sense is an examination of the degree to which the test measures or captures the essence of what is being measured (the construct).
construct-irrelevant variance	The degree to which test scores are affected by processes unrelated to the construct in question (eg, uniformity and types of test preparation).
content validity	The degree to which the test or instrument items are representative of the content it is supposed to measure. Face validity is a type of content validity that uses “expert” or “professional” appraisal of whether the items are representative of the intended content.
predictive validity	The degree to which an instrument predicts some behavior or trait.
convergent validity	The degree to which 2 or more instruments that are supposed to measure or predict the same behavior agree.
reliability	The consistency of scores or the proportion of the total variance due to true differences among individuals. The less the error in measurement, the higher the reliability. There are numerous ways to estimate reliability based on the nature of the data (eg, inter- or intra-rater reliability, Cronbach’s alpha, Spearman-Brown prophesy).

*References 23, 28, 33