

The Development of a Competency-Based Assessment Rubric to Measure Resident Milestones

BEATRICE A. BOATENG, PHD
LANESSA D. BASS, MD
RICHARD T. BLASZAK, MD
HENRY C. FARRAR, MD

Abstract

Background The outcomes-based assessment rubric is a novel systematic instrument for documenting improvement in clinical learning.

Approach This article describes the development of a rubric aimed at introducing specific performance indicators to measure the Accreditation Council for Graduate Medical Education competencies.

Results The potential benefits and implications for medical education include specifying performance indicators and outcomes, ensuring that assessment is coherent and consistent for all residents, measuring resident outcomes based on real-life criteria, providing opportunities for residents to demonstrate proficiency in a specific competency and outcome level, and improving the quality of assessment.

Introduction

The Accreditation Council for Graduate Medical Education (ACGME) has always mandated evaluation of the resident learner for continued accreditation of residency programs. However, with the movement to competency-based education and the Outcome Project, in which programs will be accredited based on patient care and learner outcomes, accurate assessment and evaluations are even more critical. Traditionally, student feedback has been based on knowledge acquisition and the learner's ability to recall key concepts as defined by the faculty. A review of the literature since the advent of the Outcome Project has shown the shortcomings of traditional methods and a renewed focus on more accurate assessment methods.¹

In the area of residency training, faculty is considered experienced enough to assess the knowledge and skills of the student learner while doing hands-on patient care activities. However, various investigators have detailed the challenges faculty face in accurately assessing a resident's clinical competence.²⁻⁵ Challenges include variance in faculty ratings of a resident's performance (ie, individual faculty

rate the same resident differently on the same rotation) and evaluation based on an arbitrarily defined interpretation of model clinical performance.⁴ In spite of these challenges, educators acknowledge that meaningful and accurate assessment of students' knowledge, skill acquisition, and behavioral modifications may improve the quality of learning.

This article is a description of one program's attempt to develop a learner-centered, competency-based assessment rubric for faculty global rating forms that combines 2 assessment models to demonstrate increasing skill acquisition among pediatric house officers. A learner-centered assessment paradigm is intended to promote higher performance, integrate education with experience, and provide constructive feedback to motivate the learner to strive for the desired outcome.

Background

Global rating forms are prevalent assessment tools among residency programs.⁶ Because the ideal evaluation should be real time, relevant and practical, global rating forms will continue to have an important role in assessing residents. Identifiable issues related to resident assessments include inadequate description of evaluation criteria, variations in raters' observations and assessments, unsatisfactory or lack of meaningful feedback, and timeliness of feedback.⁷⁻⁹ Furthermore, assessment tools appear to lack detailed requirements of performance expectations as well as behaviors for each competency or domain.¹ There continues to be an overemphasis on evaluating knowledge acquisition rather than measuring performance progress over time.¹⁰ In shifting toward a performance-based assessment system, benchmarks have emerged as a plausible option for measuring competency-specific behaviors.¹¹ Each specialty

Beatrice A. Boateng, PhD, is an Assistant Professor in the Department of Pediatrics and is with the Office of Education and Evaluation, College of Medicine, University of Arkansas for Medical Sciences, Arkansas Children's Hospital; **LANESSA D. Bass, MD**, **Richard T. Blaszkak, MD**, and **Henry C. Farrar, MD**, are in the Department of Pediatrics, College of Medicine, University of Arkansas for Medical Sciences, Arkansas Children's Hospital.

The authors wish to thank Ms Suzanne Speaker for reviewing this article.

Corresponding author: Beatrice A. Boateng, PhD, Department of Pediatrics, University of Arkansas for Medical Sciences, Arkansas Children's Hospital, 1 Children's Way, No. 512-19A, Little Rock, AR 72202, 501.364.1874, bboateng@uams.edu

Editor's Note: The online version of this article includes additional materials such as data tables, survey or interview forms or assessment tools.

DOI: 10.4300/01.01.0008

board has been charged by the ACGME to propose specialty-specific thresholds, benchmarks, or milestones that would indicate competency at that particular training level and task, instead of specifically relying on the *gestalt* of faculty. Pediatrics has already begun this process with the Milestones Project; specific parameters will be used to measure residency competency throughout the different levels of training.

In 1999, Pangaro¹² developed the RIME (reporter–interpreter–manager–educator) model as a framework for demonstrating professional growth in medical students and residents. The RIME method has been determined to provide meaningful feedback to students.¹³ Additionally, the Dreyfus model for skill acquisition (novice–advanced beginner–competent–proficient–expert) has been recommended by the ACGME as a tool to demonstrate progression in skill acquisition over time.¹⁴ As we move toward a progressive behavioral assessment system, the RIME and Dreyfus models provide us with frameworks for consistency within the evaluation system. The challenge lies in layering the competencies with the evaluation frameworks and demonstrating growth in knowledge, skills, and attitudes over time.

Although various tools have been developed to measure the competencies, Lurie et al¹ concluded that “an explicitly stated set of expectations that would link the ideals of the general competencies to the realities of measurement” remains missing from the ACGME required evaluation process.

Why Use Rubrics?

Rubrics have been used in kindergarten through grade 12 (K-12) and higher education and are gaining recognition in professional education.^{7,15–19} A rubric is a “scoring tool that lays out the expectations for an assignment.”⁸ Rubrics generally have 4 parts: (1) description of the task, (2) the scale to be used, (3) the dimensions of the task, and (4) the description of each dimension on the scale.

Within the health professions, rubrics have been used for assessing literature review skills,⁹ grading papers,²⁰ assessing presentations,^{21,22} improving the quality of online courses,²³ participating in online discussions,²⁴ determining clinical performance in the operating room,¹⁶ and measuring skill development of aseptic techniques.²⁵ Overall, rubrics promote consistency in scoring, encourage self-improvement and self-assessment, motivate learners to achieve the next level, provide timely feedback, and improve instruction. Additionally, rubrics are beneficial to learners,^{18,25} facilitate communication between faculty and learners,⁸ and enable faculty to communicate specific goals, expectations, and performance requirements.

Clinical evaluation remains challenging to even the most seasoned faculty, and rubrics provide a learner-centered assessment approach that focuses on encouraging behavioral change in learners. Although rubrics may not be

useful for multiple-choice examinations, the clinical experience provides an opportunity for performance-based assessments, and rubrics can be used to objectively assess performance. Performance tests are generally used to determine if a learner has mastered specific skills, and the instructor typically makes inferences about the level to which the skill has been mastered. Rubrics provide a potential solution to the subjective grading dilemma faced by clinical faculty.

The Development of the Rotation Rubric

In 2001, we modified an existing end-of-rotation faculty global rating form to evaluate residents based on the 6 ACGME competencies. This instrument included 16 items covering the 6 core competencies, an additional 3 related to resident teaching skills, 1 related to overall competence of the resident on the rotation, and a space for qualitative feedback. The revised instrument used the Dreyfus and RIME models for evaluation, and items were scored using a Likert scale (“performance unacceptable,” “novice,” “advanced beginner,” “competent,” “proficient,” “expert,” and “not observed”). The assessment instrument was used by teaching faculty to evaluate residents at the end of each rotation (TABLE 1). Faculty were educated about the rationale, design, and scoring of the new instrument. This was initially done with a one-time, hour-long grand rounds training format. Semistructured, supplementary training occurred periodically during the academic year (ie, new faculty orientation, faculty meetings, and quarterly resident advisor meetings). Additionally, a detailed explanation of the instruments scale was included on the back of the evaluation form.

Although the instrument was useful in providing a formative assessment of resident performance, a cursory review of the first 12 months of data indicated inconsistencies in faculty evaluations of the same resident on a rotation. Faculty expressed uncertainty about the following aspects of the assessment: What did it mean to be an advanced beginner in professionalism or practice-based learning and improvement? What behaviors or skills constituted a particular level of competency?

In 2008, pediatrics teaching faculty completed a survey on their perceptions of, attitudes toward, and the applications of the ACGME core competencies.²⁶ Faculty acknowledged the usefulness of the competencies in improving patient care outcomes; however, they were less confident about assessing systems-based practice and practice-based learning and improvement, indicating that a greater understanding of these competencies could lead to improvement in their assessment of residents within these competencies.

Against this backdrop and in an effort to improve faculty recognition and assessment of systems-based practice and practice-based learning and improvement, the evaluation form was revised in 2009 to include benchmarks

matched to the Dreyfus and RIME scales. Additionally, we sought to address questions raised by the previous evaluation tool with regard to identifying recognizable behaviors that matched a specific competency level of the core competencies. The RIME scale was modified to include the master level (R-I-Manager-Master-E). The instrument was reduced from 20 to 13 items. Professionalism was the most challenging competency to incorporate within the Dreyfus/RIME models. It is fluid in definition and may contain as many as 3 to 7 identifiable factors.^{27,28} Professionalism was viewed as evident in all the competencies and a separate scale was used for its assessment (TABLE 2). Seven professionalism factors were identified through the ACGME core competencies and the American Board of Pediatrics Program Directors guide to teaching and assessing professionalism.²⁹ A 6-point frequency Likert scale was used for professionalism: “never,” “rarely,” “sometimes,” “often,” “all the time,” and “not observed.” In modifying the assessment tool and developing the rotation rubric, benchmarks were developed by addressing the following questions:

1. Which elements (knowledge, skill, and attitude) make up a particular competency, as it applies to pediatrics?
2. What are the recognizable behavioral anchors? Are these behaviors observable?
3. What are the benchmarks?
4. How will the behavioral anchors be measured?

With these questions in mind, we developed a rotation rubric with the goal of providing an improved formative assessment tool that would provide both the assessor (faculty) and the assessed (resident) with clearly recognizable, observable skills and behaviors.

Methods

The rubric was developed within an instructional framework that focuses on achieving desired outcomes.³⁰ This was achieved in 3 steps:

Step 1: Identifying the Desired Outcomes

The primary focus in developing the rubric was outcomes. What were the essential behaviors, knowledge, and skills that we would like our residents to have at the end of their educational experience? We defined outcomes as the expression of the learner’s capability as demonstrated through particular skills and behaviors. These were developed within the Dreyfus and RIME assessment models with the desired outcome at the end the educational experience defined as a proficient to expert physician.

Step 2: Identifying Evidence of Observable Skills and Behaviors

The most challenging aspect of developing the rubric was defining the observable behaviors associated with learning and mastery. We sought to include the 3 main behavioral

domains associated with learning (cognitive, affective, and psychomotor) with the ACGME core competencies. The cognitive domain, which deals with the acquisition and retention of knowledge, can be evaluated through existing test instruments along with observation of its use by residents in providing quality patient care. The affective domain deals with attitudes toward patient care and is evident through the patient care, interpersonal and communication skills, and professionalism competencies. Lastly, the psychomotor domain, which deals with skill acquisition, can be measured across all ACGME core competencies. In addition to behaviors associated with the ACGME core competencies, we identified knowledge, attitudes, and clinical skills specific to pediatrics.

Step 3: Reviewing Behavioral Descriptions

Finally, a multidisciplinary team reviewed the descriptive narratives of the rubric. The team consisted of the pediatric program director (H.C.F.), the associate program director for assessment (R.T.B.), the associate program director for curriculum development (L.D.B.), and the pediatric education coordinator who holds a doctorate in education (B.A.B.). The rubric went through several iterations to refine the descriptions, the scale, and the descriptive narratives of the instrument to make it more clinically relevant and pediatrics-specific.

Implications

Our initial 2001 competency-based Likert scale assessment tool revealed inherent limitations of Likert scale evaluations. Our instrument could not provide specific indicators that faculty could observe to objectively assess residents at the end of the rotations.

Although rubrics have been used in K-12 and higher education, their application in residency education aligns with a shift toward developing outcomes-based assessments for residency education. Outcomes-based assessments are a systematic process through which a program can articulate what it intends to accomplish with regard to services and learning. The assessment process can be measured over time, and the outcomes used to plan program improvements.

The rubric described in this article is a novel formative assessment instrument that provides faculty with identifiable observable skills in residents. An outcomes-based assessment rubric provides tools for objectively assessing resident learning and encouraging lifelong learning skills while focusing on achieving clearly defined outcomes. Practice can also be improved through outcomes-based assessment.³¹

Within medical education, rubrics offer additional benefits including:

1. Specifying performance indicators and outcomes
2. Ensuring that assessment is coherent and consistent for all residents

3. Measuring resident outcomes based on real-life criteria
4. Providing opportunities for residents to demonstrate proficiency in a specific competency and outcome level
5. Improving the quality of assessment

Rubrics provide an understanding of the relationship between objectives and outcomes. The alignment of outcomes and assessments is the foundation for drawing valid inferences about learning.³² This implies that we need to rethink assessment in graduate medical education; resident progress should be based on observable performance indicators. As per ACGME requirements, residents are provided with rotation-specific curricula, competency-based goals and objectives, and the assessment tools required for each rotation in order to create awareness of rotation requirements and expectations. It is hoped that at the end of their training, the culmination of rotation assessments and feedback received throughout the training will result in more competent physicians who are better equipped for independent practice.

Summary

The instrument presented here was designed to be used primarily by teaching faculty to assess residents at the end of each rotation. The 7-point scale used combined both the Dreyfus and RIME models of assessing skill acquisition and included unacceptable performance and not observed. For this rubric, the dimensions were the ACGME competencies and teaching skills. The “expected level” component for each training level is included for faculty training only and then removed for actual assessments of residents. This instrument is currently being piloted in the general pediatric clinic, before widespread institutional use.

Further considerations include the adjustment of rubric terminology to match competency-based rotation goals and objectives for each rotation. The rubric will also enable us to compare resident assessments across rotations as well as against other existing assessments such as the in-training examination, Clinical Skills Exam (CHEX), peer, self-evaluation, or 360 evaluation. We hope that customizing the rubric for each rotation will guide the faculty toward more accurate, less subjective consistent assessments.

References

- 1 Lurie SJ, Mooney CJ, Lyness JM. Measurement of the general competencies of the accreditation council for graduate medical education: a systematic review. *Acad Med*. 2009;84(3):301–309.
- 2 Herbers JE Jr, Noel GL, Cooper GS, Harvey J, Pangaro LN, Weaver MJ. How accurate are faculty evaluations of clinical competence? *J Gen Intern Med*. 1989;4(3):202–208.
- 3 Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA*. 2002;287(2):226–235.
- 4 Pulito AR, Donnelly MB, Plymale M, Mentzer RM. What do faculty observe of medical students' clinical performance? *Teach Learn Med*. 2006;18(2):99–104.
- 5 Siegel BS, Greenberg LW. Effective evaluation of residency education: how do we know it when we see it? *Pediatrics*. 2000;105(4, pt 2):964–965.
- 6 Silber CG, Nasca TJ, Paskin DL, Eiger G, Robeson M, Veloski JJ. Do global rating forms enable program directors to assess the ACGME competencies? *Acad Med*. 2004;79(6):549–556.
- 7 Moni RW, Moni KB. Student perceptions and use of an assessment rubric for a group concept map in physiology. *Adv Physiol Educ*. 2008;32(1):47–54.
- 8 Stevens DD, Levi AJ. Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback and Promote Student Learning. 1st ed. Sterling, VA: Stylus Publishing LLC; 2005.
- 9 Blommel ML, Abate MA. A rubric to assess critical literature evaluation skills. *Am J Pharm Educ*. 2007;71(4):1–8.
- 10 Garfunkel LC, Sidelinger DE, Rezet B, Blaschke GS, Risko W. Achieving consensus on competency in community pediatrics. *Pediatrics*. 2005;115(suppl 4):1167–1171.
- 11 Carraccio C, Englander R, Wolfsthal S, Martin C, Ferentz K. Educating the pediatrician of the 21st century: defining and implementing a competency-based system. *Pediatrics*. 2004;113(2):252–258.
- 12 Pangaro L. A new vocabulary and other innovations for improving descriptive in-training evaluations. *Acad Med*. 1999;74(11):1203–1207.
- 13 Espy E, Nuthalapaty F, Cox S, et al. To the point: medical education review of the RIME method for the evaluation of medical student clinical performance. *Am J Obstet Gynecol*. 2007;197(2):123–133.
- 14 Dreyfus SE, Dreyfus HL. A Five-Stage Model of the Mental Activities Involved in Directed Skill Acquisition. University of California: Berkeley, Operations Research Center, 1980.
- 15 Moni RW, Beswick E, Moni KB. Using student feedback to construct an assessment rubric for a concept map in physiology. *Adv Physiol Educ*. 2005;29(4):197–203.
- 16 Nicholson P, Gillis S, Dunning AM. The use of scoring rubrics to determine clinical performance in the operating suite. *Nurse Educ Today*. 2009;29(1):73–82.
- 17 Isaacson JJ, Stacy AS. Rubrics for clinical evaluation: objectifying the subjective experience. *Nurse Educ Pract*. 2009;9(2):134–140.
- 18 Allen D, Tanner K. Rubrics: tools for making learning goals and evaluation criteria explicit for both teachers and learners. *CBE Life Sci Educ*. 2006;5(3):197–203.
- 19 Licari FW, Knight GW, Guenzel PJ. Designing evaluation forms to facilitate student learning. *J Dent Educ*. 2008;72(1):48–58.
- 20 Daggett LM. A rubric for grading or editing student papers. *Nurse Educ*. 2008;33(2):55–56.
- 21 Musial JL, Rubinfeld IS, Parker AO, et al. Developing a scoring rubric for resident research presentations: a pilot study. *J Surg Res*. 2007;142(2):304–307.
- 22 O'Brien CE, Franks AM, Stowe CD. Multiple rubric-based assessments of student case presentations. *Am J Pharm Educ*. 2008;72(3):1–7.
- 23 Blood-Siegfried JE, Short NM, Rapp CG, et al. A rubric for improving the quality of online courses. *Int J Nurs Educ Scholarsh*. 2008;5:Article 34. Accessed April 6, 2009 at: <http://www.bepress.com/ijines/vol5/iss1/art34>.
- 24 Lunney M, Sammarco A. Scoring rubric for grading students' participation in online discussions. *Comput Inform Nurs*. 2009;27(1):26–33.
- 25 Brown MC, Conway J, Sorensen TD. Development and implementation of a scoring rubric for aseptic technique. *Am J Pharm Educ*. 2006;70(6):1–6.
- 26 Bass LD, Brown CM, Patil S, Lloyd EC, Blaszk RT, Boateng BA. Pediatric faculty perceptions of the ACGME competencies. In Proceedings from the Accreditation Council for Graduate Medical Education; March 5–8, 2009. Chicago, IL. Abstract no. 27.
- 27 Arnold EL, Blank LL, Race KE, Cipparrone N. Can professionalism be measured? The development of a scale for use in the medical environment. *Acad Med*. 1998;73(10):1119–1121.
- 28 Blackall GF, Melnick SA, Shoop GH, et al. Professionalism in medical education: the development and validation of a survey instrument to assess attitudes toward professionalism. *Med Teach*. 2007;29(2–3):e58–62. Available at: <http://dx.doi.org/10.1080.01421590601044984>
- 29 The American Board of Pediatrics. Program directors guide to teaching and assessing professionalism. Available at: <https://www.abp.org/abpwebsite/publicat/professionalism.pdf>. Accessed April 6, 2009.
- 30 Wiggins GP, McTighe J. *Understanding by Design*. 2nd ed. Alexandria, VA: ASCD, 2005.
- 31 Reynolds AL, Chris S. Improving practice through outcomes based planning and assessment: a counseling center case study. *J Coll Student Dev*. 2008;49(4):374–387.
- 32 Killen R. Validity in outcomes-based assessment. *Perspect Educ*. 2003;21(1):1–14.