Evaluating and Enhancing Outcomes Assessment Quality in Higher Education Programs

Kenneth Wolf and Laura Goodwin

Abstract

Accreditation is a mark of distinction indicating that an institution has met high standards set by the profession, and an increasingly important feature of the accreditation process in higher education is "outcomes assessment." This article presents two rubrics for evaluating the quality of an institution's outcomes assessment system. One rubric is for rating the overall quality of an academic program's outcomes assessment system, and the other is for evaluating its student assessment component in particular.

Accreditation is a mark of distinction indicating that an institution or person has met high standards set by the profession. Most reputable universities go through an accreditation process conducted by one of the non-profit regional accrediting bodies such as the Middle States Commission on Higher Education or the Higher Learning Commission of North Central Association. Many professional schools and programs go through their own specialized accreditations as well. An increasingly important feature of the accreditation process for both higher education institutions and professional schools is "outcomes assessment" (Palomba and Banta 1999; Banta et al. 1996), which refers to a process of identifying learning outcomes for students, assessing those outcomes through a variety of measures, and using the information in a feedback loop to guide decisions about program improvement.

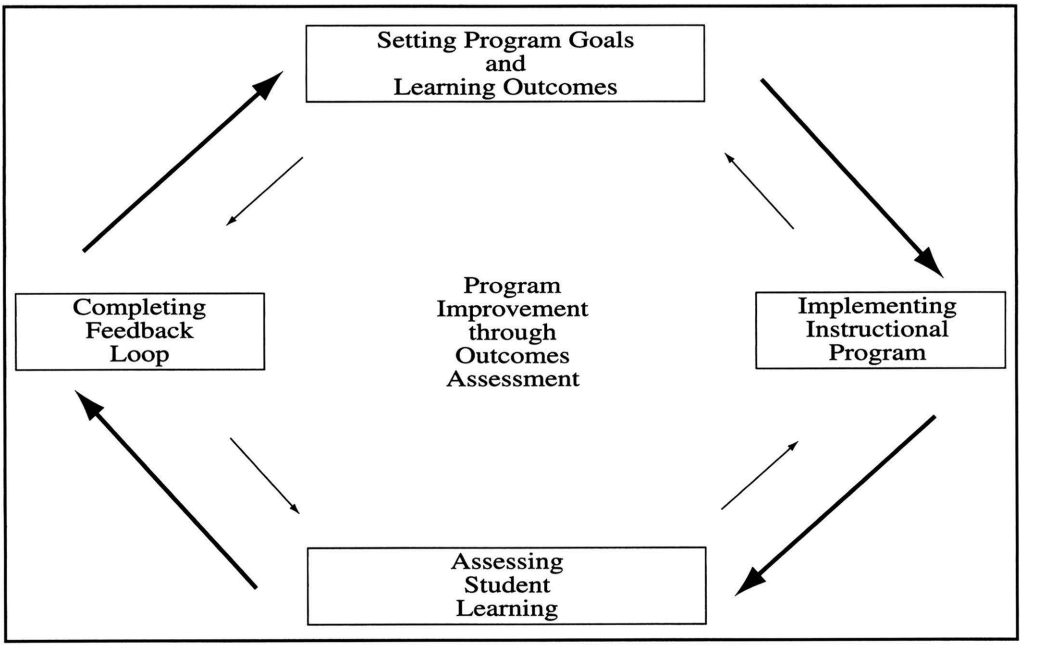
In its resource guide on accreditation, the Middle States Commission on Higher Education (2003, 1) states that "among the principles that guided the revision of the Commission's standards is greater emphasis on institutional assessment and the assessment of student learning." Of the fourteen standards for accreditation, two pertain directly to assessment. Standard 7 highlights institutional assessment (of which student assessment is a central feature), while Standard 14 refers directly to student assessment: "Assessment of student learning demonstrates that the institution's students have knowledge, skills, and competencies consistent with institutional goals and the students at graduation have achieved appropriate higher education goals" (2003, 85).

The emphasis in this article is on student assessment issues in particular as they relate to outcomes assessment at the program level. It is organized around four key features of program improvement through outcomes assessment: (1) identifying learning

outcomes, (2) implementing the instructional program, (3) assessing student learning, and (4) closing the feedback loop. The article presents two rubrics, one for rating the quality of a university program's overall outcomes assessment system and another for evaluating its student assessment component. These rubrics are intended to be a guide for programs as they design and implement their outcomes assessment system.

Key Features of an Outcomes Assessment System

Figure 1 presents a conceptual diagram illustrating the interrelationship of the four components of program improvement through outcomes assessment. The process begins at the top of Figure 1 with identifying learner outcomes, then moves to implementing the instructional program, continues with assessing student learning, and concludes with completing the feedback loop. While the large arrows portray a clockwise direction that goes from beginning to end in a cycle of continuous improvement, the smaller arrows going counterclockwise indicate that the process may backtrack on itself as it unfolds, illustrating the iterative and recursive nature of effective assessment systems.



1. Setting Program Goals and Student Learning Outcomes

Program goals describe in broad strokes the vision of a program and set the framework for courses and other learning experiences. As well, the program goals should align with both the overall university mission and the more specific course-level learning outcomes and objectives for students. An example of a program's goals from the University of Colorado at Denver and Health Sciences Center in the Business School's master's degree in Finance follows:

> The program's educational goals are to provide students with a solid knowledge of financial theory and economic principles to allow students to

have the analytical skills needed for a career as a financial manager or finance specialist and to provide students with specialized electives to gain more in-depth knowledge in particular finance areas to advance in their particular area of choice (Cooperman 2005, 2).

While program goals guide the development of the program, student learning outcomes are the hub of the outcomes assessment process. Clear and explicit learning outcomes strengthen the teaching and learning process by making the learning target visible for all to see. As Stiggins (2001) points out, it's difficult for students to hit the target if they don't know what it is! As well, if teachers keep student learning goals front and center as they plan and teach, they are more likely to help students develop the essential knowledge and skills considered necessary for performance in the discipline or field.

At the program level, the number of outcomes that seems to work best is somewhere between a large handful and a small armful. Typically, a program needs to have three or more outcomes so that key components of the program can be meaningfully delineated, but not so many outcomes that people can't remember them without looking them up on the program's Web site every time they need to refer to them. Many programs reference, if not adopt whole cloth, the learning outcomes cited by their professional organizations.

The program goals for the University's master's degree program in Finance are realized through five broad learning outcomes. Students will be able to:

1. Evaluate the investment and financial decisions of a firm utilizing financial theory and decision-making tools.
2. Apply quantitative data analysis methods to help analyze financial decisions.
3. Apply economic analysis to a firm's decision-making and understand the impact of the economic environment in evaluating financial decisions.
4. Demonstrate an understanding of the external and internal economic environment and demonstrate knowledge of different financial instruments, how the U.S. economy works, how the Federal Reserve System operates and carries out monetary policy, and theories of yield curves and interest-rate determination.
5. Perform a financial analysis of a firm and perform a pro forma forecasting analysis demonstrating knowledge of financial accounting and financial analysis (Cooperman 2005, 2).

Each of these five outcomes is operationalized in greater detail to further guide both teaching and assessment. For example, the first learning outcome addressing whether students will be able to evaluate the investment and financial decisions of a firm is further defined by the following three attributes.

Students will be able to:

a. Evaluate the investment and financing decisions of a firm, using the tools and theories underlying cost of capital, capital budgeting, and capital structure.
b. Evaluate the investment decisions using tools and theories based on concepts of risk and return and portfolio diversification.

c. Explain the basic issues in strategic finance decisions including capital budgeting, cost controls, risk management and performance evaluation (Cooperman 2005, 3).

Learning outcomes expressed at the *course level* may be identical to those at the *program level*, or they may represent only a portion of the program outcomes since a single course or learning experience is unlikely to address all of the outcomes for a program. However, while course level outcomes may sometimes go beyond those for the program, they should be consistent with and supportive of the overall program learning outcomes for students. Alignment of program and course goals and outcomes is vital for an effective outcomes assessment system.

2. Designing and Implementing the Instructional Program

Once program faculty are clear about what they want students to know and be able to do, they can begin to design learning experiences (e.g., courses, internships, studios) that best support students in achieving the requisite knowledge and skills. Comparing outcomes and learning opportunities can help programs map out whether students have the opportunity to learn each of the program outcomes.

While the teaching and learning process is the critical component in the program improvement cycle, the purpose of this article is to discuss the assessment of student learning, and not teaching and learning more broadly. Hence, this section is brief, and does not address issues such as student characteristics and motivation, curriculum coherence and development, or instructional design and delivery. However, given that the assessment of student learning is an integral and embedded strand in the overall instructional program, the general topic of assessment merits a brief comment in light of its interwoven relationship with instruction before being more formally discussed in the following section.

Effective assessment should take place before, during, and after instruction. Pre-assessments of student knowledge and perceptions can help a teacher customize instructional content and strategies to a particular group of students. Ongoing assessments of student learning and attitudes can help teachers make adjustments in their instruction along the way. Post-assessments provide essential information about whether students have mastered the material. All three of these assessments are part and parcel of the instructional process and when carefully woven in are sometimes indistinguishable from instruction or the learning experience itself.

3. Assessing Student Learning

Assessments of student learning can provide evidence of whether students have mastered a program's outcomes. However, the quality of this evidence depends in large part on the quality of the assessments on which it is based. A number of issues are important to consider in ensuring high quality assessment, including assessment formats and processes, summative and formative assessment, direct and indirect measures, and validity and reliability as they pertain to outcomes assessment.

Table 1 presents a rubric for rating the quality of a program's assessment system in light of these and other issues.

Table 1. Rubric for Rating Assessment Quality at the Program Level

	Inadequate	Adequate	Superior
Alignment with Outcomes	Alignment between assessments and learning outcomes is weak or unclear.	Assessments are aligned with program goals and learning outcomes.	Assessments are fully aligned with program goals and outcomes, and embedded at both the course and program level.
Soundness of Measures	Assessments may be unsound for a variety of reasons including: • problematic design issues such as with instructions or types of questions; • questionable validity due to mismatch with outcomes; • may not be appropriate to the purpose or for the type of learning being measured.	Assessments are sound in purpose, design, and application.	The purposes for assessment are clear, and the assessment formats advance those purposes. Assessments are designed with attention to the specifics of each format. For example, multiple choice response options are parallel, essay questions are at an appropriate reading level for the student, and performance assessments have clear instructions.
Embedded I in Learning Experiences	Assessments are summative only and occur only at the conclusion of learning experiences.	Assessments occur before, during, and after learning experiences, and are both formative and summative.	Assessments occur before, during, and after learning experiences, and are both formative and summative in purpose, with ongoing and productive feedback to students. Assessments are sometimes indistinguishable from learning experiences themselves, and often advance the learning being measured.
Direct Measures	Direct measures of student learning are not collected.	Both direct and indirect measures of student learning are collected.	Both direct and indirect measures of student learning, along with a variety of other sources of information, are collected.
Multiple Measures	Most outcomes are measured with a single assessment.	Most outcomes are assessed with more than one measure.	All key outcomes are assessed through multiple measures at different intervals in a student's educational experience.

Table 1. ***continued***

	Inadequate	Adequate	Superior
Various Formats	Most outcomes are assessed through a single format, typically an objective test.	Most outcomes are assessed through multiple formats such as exams, portfolios, studios.	All key outcomes are assessed through a number of different formats and especially with performance-based assessments.
Assessment Criteria (e.g., Rubrics)	Assessments have unclear or unstated criteria for scoring the performance.	Assessments have clear and explicit criteria, including rubrics where appropriate.	All key assessments in the program have clear and explicit criteria, often in the form of a rubric, that are shared with all judges and students. Where possible, rubrics and sample performances are shared with students in advance of their performances.
Data Analysis	Analysis of assessment data may be unsound, or inappropriate inferences are being drawn from the data.	Analysis of assessment data is sound, uses both quantitative and qualitative methods, and makes appropriate inferences about the nature of student learning.	Analysis of assessment data is sound, draws on both quantitative and qualitative methods, and through innovative approaches insightfully offers views into patterns in student learning that might not have been otherwise visible.
Assessment Processes	Assessment processes are not clear or are problematic, such as when criteria or approaches are applied unequally across individuals.	Assessment processes and criteria are applied in a consistent and fair fashion across performances.	Assessment processes and criteria are applied in a consistent and fair fashion across performances, with careful attention to issues that can diminish reliability or fairness such as the training of raters, assessment context, or availability of resources.
Self Assessment	Students are not invited to self-assess, nor are peers encouraged to offer feedback.	Students are given multiple opportunities to self-assess and receive feedback from peers.	Students are given structured opportunities (as with the use of a rubric) to self-assess, and peers and other professional are regularly enlisted to give constructive feedback.
Technology Use	Assessments do not take advantage of available technology.	Assessments draw on appropriate technology.	Assessment process utilizes appropriate and innovative technologies, such as data analysis software, clickers, or electronic portfolios.

Table 1. ***continued***

	Inadequate	Adequate	Superior
Measurement Error	Significant measurement error exists in the assessment process, due to problems such as unclear instructions, cultural bias, or non-adaptive testing methods.	Considerations are given to minimizing measurement error in assessment design, implementation, and analysis.	Measurement error is minimized through a systematic review of common problems related to areas such as assessment design, scoring methods, student characteristics, and testing conditions.

Assessment formats. Assessments of student learning come in a variety of formats. Stiggins (2001, 86), for example, describes four general assessment approaches: selected response, essay, performance assessments, and personal communication. The selected response format includes short answer and multiple-choice type tests. Essays are a time-honored form of the constructed response format and refer to assessments of extended written products such as term papers. The performance assessment format refers to activities in which students are judged based on a product that they have developed (such as a sculpture) or a performance that they have enacted (such as a dance). The personal communication format refers to assessments based on activities such as oral exams or interviews with students, for example.

It is important to note that some formats are more effective than others depending upon the purpose for the assessment and the type of learning being assessed. To that point, Stiggins (2001) also describes five types of learning—knowledge, reasoning, performance skills, products, and dispositions—and discusses which assessment format is best at assessing which type of learning. Knowledge, for example, is most efficiently tapped, in most cases, with a straightforward selected response measure such as a multiple-choice test since this kind of measure can sample a topic broadly and efficiently. On the other hand, performance skills, such as when an architecture student builds a model of a proposed new city library, are best measured through a performance assessment.

Imagine if an architecture student's skill at building a model of the library was assessed through a multiple-choice test instead of a performance assessment. Not only would this kind of assessment not allow the student to demonstrate his or her skill in applying the principles of building design and construction, the learning opportunity inherent in most performance assessments would be lost as well. On the other hand, using a performance assessment in which students create a dramatization of the Lewis and Clark expedition would be overkill if the learning outcome were only for students to be able to name the states traversed by the expeditionary party. As one pundit put it,

"With performance assessments, the juice should be worth the squeeze" (McTighe and O'Connor 2005, 15).

Performance assessments. Given both its value and challenge, the performance assessment format deserves further attention. As discussed earlier, performance assessments are those in which students demonstrate their learning through complex performances that often are similar to what they would be asked to do as professionals. Examples of performance assessments include such tasks as compiling a portfolio of paintings, performing a dance recital, or teaching elementary students. Performance assessments are especially attractive for assessing learning because students are typically engaged in activities that simulate what they would be asked to do in an authentic context in the real world. Thus, in the doing of the assessment, students are learning the skills that they need to show mastery of the outcomes.

A note of warning, however. While they have great value as learning activities, high quality performance assessments take time to design. For example, writing directions for the assessments with enough detail so that students know what is expected of them and developing a rubric that captures all of the key dimensions of the performance are two tasks that can be very time-consuming. In addition, scoring these kinds of assessments can be time-consuming as well and difficult to carry out because the performances themselves can be quite complex. A schoolteacher's portfolio, for example, might contain samples of student work, videotapes of teaching, lesson plans, and reflective commentaries by the teacher. Assessing each of these can be challenging since the portfolio contains so many different types of information. What can help make sense of all of this information in an efficient way is a rubric or scoring guide in which key features of the expected performance (i.e., the performance criteria) are described at various levels of quality.

Rubrics. If learning outcomes are the hub of the outcomes assessment process, then rubrics are the linchpin. Rubrics are scoring guides used by trained raters to assist them in assessing the quality of a product or performance (Arter and McTighe 2001).

We always have criteria in mind when we evaluate something, whether it is a movie or a springboard diving performance. It's just that these criteria aren't always explicit, sometimes even to ourselves. When we judge a springboard diver's performance as good or bad, for example, we are basing that judgment on something. We have some criteria in mind. Maybe it's the number of body rotations or the splash the diver makes on entry. Maybe it's something that really has nothing to do with the performance itself such as the diver's smile or nationality.

As we become more informed about springboard diving, however, we may begin to draw on the five criteria used by the springboard diving association: Starting Position, Take Off, Approach, Flight, and Entry (Federation Internationale de Natation n.d.). These criteria are then elaborated in a rubric that describes what is meant by each of them. Each of these criteria, as well as the overall performance, is then described on a

scale that covers six levels from "complete failure" to "very good."

In an outcomes-based (or criterion-referenced) system, performance levels for key assessments are established in advance by deciding at the course or program or institutional or professional level how good is good enough. For a student to be proficient in a program of study, for example, what level of knowledge and skills (and sometimes dispositions) are expected? Setting these performance levels can be a laborious but necessary process that draws on the professional standards or expectations for the field or discipline, and that requires extensive reflection, discussion, implementation, and revision by the program faculty.

A good rubric works on multiple levels. In springboard diving, for example, a rubric provides important guidance to judges, coaches, and athletes. A rubric makes it clear to judges how to rate a dive, though they still need to draw on their extensive professional knowledge in applying these criteria. Coaches study the criteria so that they can provide effective instruction to their athletes. And the athletes themselves examine the criteria to help them in planning and perfecting their dives.

While rubrics make the assessment of student products and performances more valid and reliable, their real value lies in advancing the teaching and learning process. A good rubric makes the learning target more tangible for students, and it also functions as a guidepost for teachers as they design and carry out their instruction.

But having a rubric doesn't necessarily mean that the evaluation task will then be simple or clear-cut. The best rubrics allow evaluators and teachers to draw on their professional knowledge and experience and then use that knowledge and experience in ways that the rating process doesn't fall victim to the whims of personality or the limitations of human information processing.

Assessment processes. The process in which student products and performances are assessed is an important consideration as well. Should an individual or group assess the performance? If by a group, should group members independently rate the performance and then average their scores, or should they discuss the performance in a group gathering and then arrive at a consensus? Should the performance be scored holistically (i.e., assigning an overall score for a performance) or analytically (i.e., assigning separate scores for each of the key dimensions of the performance)? Each of these approaches is valid, but it is important to recognize that there are trade-offs. For example, holistic assessment generally results in greater reliability among raters in scoring over analytic approaches but does not offer the student much in the way of specific feedback. To best ensure fairness and avoid bias, assessment criteria and processes should be applied consistently across all key performances in a class or program so that no individual is treated any different from any other individual in terms of the scoring process.

Formative and summative assessment. Summative assessments, those conducted at the end of a learning experience to give a grade or score as a measure of how well students have mastered the program or course content, are the main focus of outcomes assessment. However, formative assessments, those conducted during the learning experience to monitor student progress and to guide instruction, advance the entire teaching and learning process.

Formative assessments of students can be informal, such as asking students questions during a lesson to gauge their understanding, or more formal, such as giving students a list of terms from the topic under study and asking them to organize the terms into a conceptual map. Formative assessment can also include self and peer assessment. Accurate and insightful self-assessment is a hallmark of a professional and a skill that should be developed in students. Students might be asked, for example, to use a rubric to rate their own projects as a way to deepen their understanding of the performance criteria and sharpen their ability to self-critique.

Peer assessment, which should be used only for feedback and not grading, can also give students another perspective on the quality of their work. Rather than evaluate the work per se, peers might be directed to ask fellow student questions as a way of getting them to think more deeply about their work. For example, peers might inquire: What are you trying to convey through your collection of photographs? Why are some in black and white, and others in color? Why don't you show anyone's feet? Rubrics can help peers make more informed assessments of their colleagues' work because it can give them a frame of reference and the language for describing performances at various levels.

Direct and indirect measures. Assessments can be in the form of direct measures of student learning as in exams, projects, or presentations, or indirect measures such as knowledge surveys, exit interviews of graduates, and number of job placements. Direct measures demonstrate that actual learning has occurred, while indirect measures only imply that learning has taken place.

When students build an architectural model, for example, they are providing direct evidence of their knowledge and sills. But when students self-report their level of knowledge on a topic, for example, they are presenting their perceptions of what they know and can do, but they have not provided actual evidence of their knowledge and skills. Indirect measures of student learning, however, can provide useful insights into the "why's" and "how's" of learning, while direct measures for the most part convey "if" learning has occurred. In an outcomes-based system, the emphasis is on examining direct evidence of whether or not students have mastered the program outcomes.

Validity and reliability. Validity refers to the soundness of the inferences that are made on the basis of the assessments, while reliability pertains to the consistency of the assessments. Validity is an issue if the assessments aren't aligned with the outcomes that they are intended to measure. For example, if the outcome is that "student will be

able to drive a car" and the assessment is a written quiz on the rules of the road, there is a validity problem if we infer that students know how to drive if they pass the written assessment. Outcomes and assessments need to be in alignment.

Reliability is a problem if a teacher's ratings of student performances or products are not consistent across students, such as when two students get different scores for identical or highly similar performances. Inter-rater reliability, or the consistency of measurement across raters, is also an important consideration if more than one rater is used to evaluate the same performance. For high inter-rater reliability, different raters should arrive at the same or similar scores for the same performance. A good way to increase reliability in both of these situations is to have a rubric, train raters, and use a consistent set of procedures for evaluating the performance.

4. Completing the Feedback Loop at the Program Level

All of this assessment information is not reaching its full potential if it is not used to guide program improvement. Closing the feedback loop—using the assessment of student learning information, along with other data sources, to inform program improvement—is the essential "final" step in the outcomes assessment process. ("Final" in the sense that it is the last step in a process that is ongoing and cyclical as well as iterative and recursive.)

A variety of information about students should be collected to inform program improvement. The information could include data such as profiles of newly admitted students, student retention and graduation rates, and interviews with employers of graduates as captured in part though assessments such as the National Survey of Student Engagement. Though valuable, this information is only part of the picture. A system based on whether students have achieved the learning outcomes of the program or college requires direct measures of evaluative learning such as exams, projects, models, term papers, internships, studio demonstrations, artistic recitals or recitations, and the like.

Once the data are aggregated and analyzed, it is essential that faculty within a program (or school or college) come together to review and interpret the information. Various formal or informal processes are possible, but what is necessary is that faculty members in a program take the time to discuss what the information means. Data alone have no meaning; it is through analysis and discussion of the data that meaning is created.

One approach taken by the Colorado Consortium for Data-Driven Decisions (J. O'Brian, December 15, 2006, personal communication) in guiding school teachers in interpreting and acting on the data they collect about the students in their schools, consists of four steps—predicting, exploring, explaining, and acting—which have been adapted from Wellman and Lipton (2004).

Asking faculty first to "predict" what they expect to find before they begin a comprehensive review of the data allows for conjectures to be made and, in some cases, for assumptions to surface—about the students, about their learning, about the assessment process itself. The analysis of the data can sometimes confirm these conjectures and assumptions and at other times contradict them. Literacy research has pointed to the value of asking readers to predict before they engage with a text as a way of activating their schema and stimulating a deeper involvement with the content. People are more invested in finding out the results when they have first speculated about what those results might be.

"Exploring" the data prevents faculty from rushing to judgment and keeps open the possibility of finding the unexpected. Playing around with the data and juxtaposing various data sets can sometimes reveal patterns that might not have been evident otherwise. As John Dewey reminds us, "A problem well-framed is half-solved."

After the faculty has taken some time to consider various possible trends in the data, they next settle on the key findings and offer "explanations" for these findings. For example, in reviewing research proposals from graduate students, the faculty might have noted that there were often weak alignments between the students' research questions and their data collection strategies. Many students were either proposing to collect information that had no bearing on their research questions, or were not collecting data they needed to address their questions.

Finally, after identifying patterns such as these, faculty can begin to suggest cures in the "acting" phase. In the problem of a lack of alignment between research questions and data collection methods, the faculty might assign several small-scale studies for students to conduct, with an emphasis on aligning research questions and data collection. Possibly what is called for is a new course, or maybe the content could be strengthened across three different courses. Once the modifications to the program have been put in place, ongoing data collection and monitoring help to ensure that the changes are having their desired effect.

Summary

Table 2 presents a rubric that programs can use to rate their overall level of success with outcomes assessment across the four key components of the program improvement cycle. A program has a solid outcomes assessment system in place when it has meaningful and measurable learning outcomes, provides engaging and challenging learning experiences for students, draws on sound and diverse assessments, and puts in place a feedback loop that uses assessment data to inform program improvement for the purpose of advancing student learning.

Table 2. Rubric for Rating the Overall Quality of Outcomes Assessment in Higher Education Programs			
	Absent or Emerging	Designed & Implemented	Aligned & Integrated
Program Goals & Student Learning Outcomes	Program goals are absent or minimally identified. Student learning outcomes are absent, narrow, or vague.	Program goals are sound, and aligned with professional organizations, accrediting agencies, and university mission. Student learning outcomes have been identified, and are meaningful and measurable.	Program goals are sound and embedded in all features of the program (e.g., syllabi, course descriptions, assessments, field work, research) as well as aligned at all levels. Student learning outcomes are aligned with and integrated across all elements of the program.
Instructional Program	The instructional program is disjointed and/or not clearly connected to program goals or learning outcomes.	The instructional program engages students in sound and varied learning experiences that are linked to the program goals and outcomes.	The instructional program engages students in rich and diverse learning experiences that are clearly linked to the program goals and outcomes. The instructional program accommodates student needs and backgrounds.
Assessments of Student Learning	Direct measures of student learning are limited in number or format or not clearly connected to learning outcomes.	A variety of direct and indirect measures of student learning are in place, and assessment processes and criteria, such as rubrics, have been developed.	A variety of direct and indirect measures of student learning with sound scoring criteria and processes, including performance-based assessments, have been developed and are aligned with and integrated across all elements of the program.
Feedback Loop For Program Improvement	The program does not use student learning data to guide program improvement decisions.	The program uses both indirect and direct measures of student learning, along with other sources of information, to guide program improvement.	The program systematically draws on a variety of direct and indirect measures of student learning, along with other sources of information, and makes productive changes in the program in light of what it has learned.

References

Arter, J., and J. McTighe. 2001. *Scoring rubrics in the classroom.* Thousand Oaks, CA: Corwin Press.

Banta, T., J. Lund, K. Black, and F. Oblander. 1996. *Assessment in practice: Putting principles to work on college campuses.* San Francisco: Jossey-Bass.

Cooperman, B. 2005. Outcomes assessment. Report for the Finance Program in the Business School, University of Colorado at Denver and Health Sciences Center.

Federation Internationale de Natation. n.d. *Rules and regulations: FINA diving rules 2005-2009* http://www.fina.org (accessed January 27, 2006).

McTighe, J., and K. O'Connor. 2005. Seven practices for effective learning. *Educational Leadership* 63 (3): 10-17.

Middle States Commission on Higher Education. 2003. *Student learning assessment: Options and resources.* Philadelphia, PA.

Palomba, C., and T. Banta. 1999. *Assessment essentials: Planning, implementing, and improving assessment in higher education.* San Francisco: Jossey-Bass.

Stiggins, R. 2001. *Student-involved classroom assessment,* 3rd ed. New York: Merrill.

Wellman, B., and L. Lipton. 2004. *Data-driven dialogue: A facilitator's guide to collaborative inquiry.* Sherman, CT: MiraVia.

Author Information

Kenneth Wolf, the Interim Director of Assessment and an associate professor in educational psychology in the School of Education and Human Development, received his Ph.D. in curriculum and teacher education from Stanford University. He has published a number of books and articles on the topic of education assessment and on the schoolteacher's portfolio in particular.

Laura Goodwin, the Associate Vice Chancellor for Faculty Affairs, oversees the offices and centers for faculty development and outcomes assessment at the University of Colorado at Denver and Health Sciences Center. She also is a professor of educational psychology and research methods. Her academic areas of expertise include statistics, psychometrics, applied measurement, and research methods. She is the author of books and articles on topics in early childhood education, psychometric and statistical techniques, and faculty mentoring.

Kenneth Wolf
Interim Director of Assessment
University of Colorado at Denver and Health Sciences Center
5480 Ptarmigan Circle
Boulder, CO 80301
E-mail: Kenneth.Wolf@cudenver.edu
Telephone: 303-556-6284

Laura Goodwin
Associate Vice Chancellor for Faculty Affairs
University of Colorado at Denver and Health Sciences Center
1380 Lawrence Street, Suite 1400
Campus Box 137, P.S. Box 173364
Denver, CO 80217
E-mail: Laura.Goodwin@cudenver.edu
Telephone: 303-556-2550