Policymakers, accrediting bodies, and association leaders continue to focus on assessing student learning outcomes. But what is happening on the ground at colleges and universities? Where does student learning outcomes assessment rank in importance on an institution’s action agenda? To what extent are faculty involved in assessment activities and using the results for improving student learning?

The National Institute for Learning Outcomes Assessment (NILOA) is a multiyear effort to further the student learning outcomes agenda nationally. NILOA staff conducted four focus groups with academic deans, provosts, presidents, and directors of institutional research from a variety of two- and four-year institutions during 2009–2010 to discuss the state of assessment of student learning outcomes on campus. Roundtable discussions were conducted at meetings of the Association of American Colleges & Universities (AAC&U), American Council on Education (ACE), and the Association for Institutional Research (AIR). All told, we talked with about forty-five academic leaders, representing a range of institutional types and regions, to gain first-hand accounts of the state of efforts under way on campuses. This article summarizes what these leaders had to say and considers how the perceptions of academic leaders comport with findings from the 2009 NILOA Survey report “More Than You Think, Less Than We Need: Learning Outcomes Assessment in American Higher Education” (www.learningoutcomeassessment.org/NILOAsurveyresults09.htm), which describes what colleges and universities are doing to measure student learning. Four prominent themes cut across the focus group discussions and organize the main ideas in this paper:

1. Assessment has taken root on campus.
2. Accreditation is the major catalyst for student learning outcomes assessment.
3. Faculty involvement is central to meaningful assessment.
4. Best practices in assessment weave assessment into organizing structures.
Assessment Update
Progress, Trends, and Practices in Higher Education
September–October 2010
Volume 22, Number 5

Editor
Trudy W. Banta, professor of higher education and senior advisor to the chancellor, Indiana University–Purdue University Indianapolis (IUPUI)

Managing Editor
Karen Elaine Black, director of program review, IUPUI

Assistant Editor
Frances W. Oblander, director of institutional research & effectiveness, South University

Book Review Editor
Susan Kahn, director of institutional effectiveness, IUPUI

Consulting Editors
Peter T. Ewell, National Center for Higher Education Management Systems
Thomas Anthony Angelo, La Trobe University, Victoria, Australia
T. Dary Ervin, James Madison University
Cecilia L. Lopez, Harold Washington College
Marcia Mentkowski, Alverno College
Jeffrey A. Seybert, Johnson County Community College
Peter J. Gray, United States Naval Academy
Gary R. Pike, IUPUI


Call for Contributions
The editor welcomes short articles and news items for Assessment Update. Guidelines follow for those who would like to contribute articles on outcomes assessment in higher education.

• Content: Please send an account of your experience with assessment in higher education. Include concrete examples of practice and results.

• Audience: Assessment Update readers are academic administrators, campus assessment practitioners, institutional researchers, and faculty from a variety of fields. All types of institutions are represented in the readership.

• Style: A report, essay, news story, or letter to the editor is welcome. Limited references can be printed; however, extensive tables cannot be included.

• Format: In addition to standard manuscripts, news may be contributed via letter, telephone, or fax (317) 274-4651. The standard manuscript format is a 60-space line with 25 lines per page. Articles may be sent to <kblack@iupui.edu> as a Microsoft Word attachment. Please include your complete postal mailing address.

• Length: Articles should be four to eight typed, double-spaced pages (1,000–2,000 words). Annotations of recent publications for the Recommended Reading feature should be 200–500 words in length. Short news items and content for the Memos section should be about 50–200 words long.

• Copyright: Articles shall not have been registered for copyright or published elsewhere prior to publication in Assessment Update.

• Deadlines: Each issue is typically planned four months before its publication.

Please address mailed contributions and comments to Trudy W. Banta, Editor, Assessment Update, Suite 140 Administration Bldg., 355 N. Lansing St., Indianapolis, IN 46202–2896.

Assessment Has Generally Taken Root, and on Many Campuses Student Learning Outcomes Assessment Is Thriving

In the last decade, assessment has realized some important developmental steps. According to the dean at Drew University, “We’re in a different place than ten years ago. There is a core of people who believe in assessment and work with their colleagues on outcomes assessment.” Most campus leaders credited regional and professional accreditation with helping assessment gain ground on campus. Several leaders indicated that visible efforts like the Voluntary System of Accountability (VSA) and the independent college-sponsored site, U-CAN, had helped expand discussions about assessment and accountability among campus constituents and external stakeholders. Teagle Foundation-funded initiatives have furthered progress in developing assessment systems in the private liberal arts college sector since 2004 (see grantmaking activity www.teaglefoundation.org). The Luther College academic dean reported that learning outcomes assessment practices there had matured over the last several years, so that now at least half of the faculty are involved in assessment. Luther faculty are learning a lot about student learning outcomes through work with a consortium of colleges in the Midwest focused on the assessment of writing and critical thinking. At Oregon State University and Towson University all undergraduate programs now have learning outcomes. After a flawless accreditation review, Eastern Kentucky University embarked on an extensive analysis of course syllabi to examine the extent to which learning outcomes were common across sections of the same course. A president in the California State University system declared that assessment has been aggressively approached at the system level, and it seems that “we’ve been assessing everything that moves.”

Assessment at Westminster College grew out of the strategic planning process, and discussions about what it meant to be a Westminster graduate culminated in the articulation of learning goals specific to a Westminster education. Westminster since has developed rubrics in each academic program to assess learning goals.

(continued on page 14)
Some recent writing projects have reminded me of two truths that I discovered very early in my career in assessment. First, most faculty resist becoming involved in outcomes assessment when they first learn of this new demand on their time. Second, next to disciplinary accreditation, funding from an external source may be the second most powerful incentive for turning faculty angst and even anger about assessment to acceptance, and even appreciation.

An important initial reason for resisting outcomes assessment—specifying learning outcomes, assessing student learning in these knowledge and skills areas, then using group findings to improve instruction and/or programming—is that taking the last step of aggregating assessment results and acting on the findings is not something that most faculty have been doing on their own. Thus, outcomes assessment is viewed initially as an external mandate—one that emanates from campus administrators, agents of state government, or a regional accrediting body. Mandates for assessment imposed by an accrediting agency in one’s own discipline may be viewed as an external threat initially. But since disciplinary associations are, after all, composed of one’s own colleagues, broadly speaking, their requirements lose the stigma of an external mandate much more quickly than those imposed by other sources. Access to sources of funding for assessment initiatives can give faculty a sense of power over an external threat—an ability to do it their own way—particularly faculty in fields such as the arts and sciences for which there is no opportunity for accreditation in the discipline.

While I was a faculty member at the University of Tennessee, Knoxville (UTK), in the 1980s, the chancellor asked me to take on the task of making palatable to my colleagues a conscientious response to the Tennessee Higher Education Commission’s performance funding initiative. Clearly, faculty had to be involved because over a five-year period seniors in every department were to undergo comprehensive evaluation in their major, and specified standardized tests of generic skills were to be administered to at least a sample of all seniors annually. Moreover, recent graduates had to be surveyed every year and an annual report compiled to provide evidence that test scores and survey data had been used to improve instruction, curricula, and/or student support programs. Although performance funding was called a voluntary program, no public college or university in Tennessee could afford to forego the 5.45 percent of the budget for instruction—about $6 million each year for UTK at that time—that was available to each institution on that basis.

Access to sources of funding for assessment initiatives can give faculty a sense of power over an external threat.

Not surprisingly, my attempts to engage my colleagues in performance funding-related projects aroused contempt in some departments. Reactions included: “This external mandate is an abridgment of academic freedom!” “We give students grades. Isn’t that assessment?” “If the state wants us to do something more than assign grades, that will take time away from our research. Isn’t the state interested in benefiting from our research?”

Amidst all the brickbats and anonymous threats (not really), one day a notice that the Kellogg Foundation planned to fund some projects designed to increase the use of student information to improve institutional effectiveness appeared in my mail. I wrote a grant, UTK was selected, and before long I was approaching colleagues to tell them about “an opportunity” they had to participate in: “a project funded by the Kellogg Foundation.” Suddenly I had friends again. We took an incredibly small amount of money—$10,000 (supplemented quietly with some institutional funds from the chancellor)—and developed three “Kellogg Task Forces,” each of which took on work that would help the institution respond to a component of the performance funding initiative. One group reviewed standardized tests available in major fields, but also offered faculty

Editor’s Notes

A Little Extra Funding Goes a Long Way

Trudy W. Banta
development workshops on methods for carrying out comprehensive evaluations of student achievement in the major that did not involve the use of standardized tests. A second task force studied the available standardized tests of generic skills, made recommendations about the most appropriate ones for UTK to use, and conducted important new research using the data collected. The third task force included survey research specialists. They created an alumni survey for UTK that eventually was adapted for use by all colleges and universities in the state. Within a few years we were able to compile a multipage bibliography of assessment-related studies by UTK faculty that had been published in disciplinary journals. And over those years we collected an average of 95 percent of the annual amount of performance funding for which UTK was eligible. The Kellogg grant enabled UTK faculty to make what they perceived as an external mandate their own. They did it their way.

For more than a decade, beginning in the mid-1980s, the Fund for the Improvement of Postsecondary Education (FIPSE) helped faculty at institutions across the country, as well as some national associations, do assessment their way. Alverno College faculty were enabled to assist colleagues in consortia of other institutions to develop learning outcomes and conduct “assessment as learning,” to use the now well-known Alverno term. Grants to the American Association for Higher Education (AAHE) and to UTK in 1985 enabled both to inaugurate a series of annual assessment conferences, the latter of which continues today in Indianapolis. The role of FIPSE in helping to engage faculty was so significant that from 1991 to 1997 we published stories about projects funded by that agency in an Assessment Update column called “With FIPSE Support.”

Ultimately FIPSE priorities shifted, and funding for innovation in assessment became difficult to find. Fortunately, many college and university deans and provosts learned the lesson of attracting bees with honey and began to use funds previously awarded for course and curriculum design and professional development/renewal to encourage involvement in assessment. Faculty used such funds to purchase and pilot-test standardized tests and surveys, to develop tests and surveys of their own, and to undertake systematic evaluations of new or existing academic programs or support services. A few states and some private foundations also have made funds for these purposes available to campuses.

As I noted in a 2009 column (21:4), today the availability of funds from external sources is increasing again. FIPSE funded the national VALUE rubric-development project of the Association of American Colleges and Universities. The Lumina Foundation for Education, the Teagle Foundation, and the Carnegie Corporation of New York are supporting the National Institute for Learning Outcomes Assessment (NILOA). As Jillian Kinzie notes in her article in this issue, Teagle has funded several national consortia that are helping faculty in liberal arts disciplines generate and use evidence about student learning. And Lumina is investing in “Tuning U.S.A.” as a component of its strategy to increase the percentage of Americans with postsecondary degrees and credentials to 60 percent by 2025.

Lumina’s tuning project has had the same transformative impact in some instances as the Kellogg Project had at UTK so long ago. For years a colleague here at Indiana University-Purdue University Indianapolis fairly snarled at me, “The liberal arts are different. We don’t have specific learning outcomes because we are not preparing our students for specific jobs, as the professional schools are. In fact, the outcomes of a liberal arts education are not apparent until years after a student graduates!” As a participant in “Tuning U.S.A.,” this colleague has had the opportunity to work with ten history faculty members from Indiana’s public two- and four-year institutions to develop integrated sequential learning outcomes for associate, baccalaureate, master’s, and doctoral degree recipients in the field of history. I was in the audience at a recent national meeting where this colleague participated on a panel of liberal arts faculty from institutions in other states where “Tuning U.S.A.” support has encouraged similar work. Each panel member in turn expressed thanks to Lumina for giving them the chance to take the first step in assessment—setting learning outcomes—their way!
At Marquette University, a Jesuit, Catholic research-extensive institution, construction of a comprehensive student learning assessment system began in 2005. Under the Marquette assessment framework, student learning outcomes are specified, and an annual assessment cycle is conducted at three levels: in courses, in programs and units, (undergraduate and graduate academic, co-curricular, and the Core of Common Studies) and institution-wide.

An important principle of the Marquette Assessment System is that leadership for assessment is shared among faculty, co-curricular staff, and university administrators. Each program or unit has a designated Program Assessment Leader responsible for both the coordination of the program’s assessment plan and activities and serving as the program’s link with the institutional assessment system. The Vice Provost for Undergraduate Programs and Teaching and the University Assessment Committee, composed of faculty and staff, share the responsibility for decisions regarding the development, support, and maintenance of the overall Marquette Assessment System and for monitoring the quality of assessment conducted by the programs.

Every program faculty completes an assessment cycle annually and files a report that includes information about five components of the cycle: learning outcomes, assessment measures, results, faculty/staff conclusions about student learning, and actions planned to improve learning (program reports can be viewed at www.marquette.edu/assessment). The first annual assessment cycle was completed during the 2006–2007 academic year.

Peer Review Working Seminars

Once the first assessment reports were submitted in September 2007, the University Assessment Committee (UAC) members faced a pressing question: How can the assessment findings of 116 programs (104 academic and 12 co-curricular) be examined, shared, and constructive feedback provided to help programs improve assessment and, ultimately, student learning? The UAC members explored various approaches used at other universities. A practice that seemed a good fit with Marquette’s principle of faculty and staff ownership of assessment was identified at Cleveland State University (CSU). At CSU, program assessment reports are reviewed in a three-day faculty review session by paid, volunteer faculty members working in pairs. The Marquette UAC expanded on this approach and over the semester developed a peer-review process that involved all 100-plus program assessment leaders and gave each an opportunity for face-to-face interaction with peers as the program’s assessment report was reviewed.

Having clarified that the goal was to provide feedback and a formative evaluation of each program’s assessment activities rather than a summative statement, the use of a rating form and/or computing a total “score” was rejected. Instead, an assessment cycle rubric (Figure 1) was developed to guide the peer reviewers as they considered the extent to which a given program incorporated five components of the Marquette Assessment Cycle. This rubric has four levels of completion for each component: “none of the com-

The three-hour peer review portion involves each participant working in an assigned peer-review group of three or four program assessment leaders per table.
UAC serve as table facilitators to keep discussions constructive and focused but not as assessment consultants or peer reviewers. Each participant receives a packet containing a sheet of instructions and ground rules, copies of the annual assessment reports of each program assigned to the table, and blank copies of the peer-review report form, which uses the assessment cycle rubric. Thus, by design, over 100 faculty and co-curricular staff are actively participating in an immediate peer-feedback process, engaging in dialogue about a variety of assessment methods and practices and, through both giving and receiving peer feedback, learning more about outcomes assessment.

### Figure 1. Assessment Rubric for Marquette Program Assessment Systems

<table>
<thead>
<tr>
<th>Assessment Component</th>
<th>Beginning Assessment System</th>
<th>Meets Expectations for Assessment System</th>
<th>Assessment System Reflects Best Practices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Outcomes</td>
<td>Program learning outcomes have been identified and are generally measurable.</td>
<td>Measurable program learning outcomes. Learning outcomes are posted on the program website.</td>
<td>Posted measurable program learning outcomes are routinely shared with students and faculty.</td>
</tr>
<tr>
<td>Assessment Measures</td>
<td>General measures are identified (e.g., student written assignment).</td>
<td>Specific measures are clearly identified (student global case study in the capstone course). Measures relate to the program learning outcomes. Measures can provide useful information about student learning.</td>
<td>Multiple measures are used to assess a student learning outcome. Emphasis on specific direct measures. Rubrics or guides are used for the measures. Measures are created to assess the impact on student performance of prior actions to improve student learning.</td>
</tr>
<tr>
<td>Assessment Results</td>
<td>Data collected and aggregated for at least one learning outcome.</td>
<td>A majority of learning outcomes assessed annually. Data collected and aggregated are linked to specific learning outcome(s). Data are aggregated in a meaningful way that the average reader can understand.</td>
<td>If not all learning outcomes are assessed annually, a rotation schedule is established to assess all learning outcomes within a reasonable timeframe. Data are aggregated and analyzed in a systematic manner. Data are collected and analyzed to evaluate prior actions to improve student learning.</td>
</tr>
<tr>
<td>Faculty Analysis and Conclusions</td>
<td>All program faculty receive annual assessment results. Faculty input about the results is sought.</td>
<td>All program faculty receive annual assessment results and designate program or department faculty to meet to discuss assessment results in depth. Specific conclusions about student learning are made based on the available assessment results.</td>
<td>All of previous level and faculty synthesize the results from various assessment measures to form specific conclusions about each performance indicator for a learning outcome.</td>
</tr>
<tr>
<td>Actions to Improve Learning and Assessment</td>
<td>At least one action to improve learning or improve assessment is identified. The proposed action(s) relates to faculty conclusions about areas for improvement.</td>
<td>Description of the action to improve learning or assessment is specific and relates directly to faculty conclusions about areas for improvement. Description of action includes a timetable for implementation and identifies who is responsible for the action. Actions are realistic, with a good probability of improving learning or assessment.</td>
<td>All of previous level and assessment methods and timetable for assessing and evaluating the effectiveness of the action are included in the planned action.</td>
</tr>
</tbody>
</table>
Participants are preassigned to tables to ensure that each set of peers includes individuals from a mix of academic colleges and/or co-curricular units. During the first year of the reviews, assignment to a widely diverse group seemed counterintuitive to many of the program assessment leaders, who asked to be switched to a table with similar disciplines. However, the disciplines in each peer group are intentionally mixed based on prior experiences with assessment faculty development events. I have observed that individuals from similar disciplines, for example, chemistry, biology, and biomedical sciences, tend to share and reinforce similar views about student learning and use similar assessment methods. This sets the stage for a quite narrow peer discussion and potentially competitive assessment review, and yields fewer opportunities to learn alternative assessment approaches.

After the assessment review seminar, each program receives a written copy of the assessment cycle peer review form completed by the peer group during the seminar. Copies of each program peer review are also sent to the program’s department chair and college dean for their information and follow-up. Marquette has conducted the Peer Review Assessment Seminars for three years. Feedback from participants has been quite positive. After each review seminar program, assessment leaders have reported feeling more knowledgeable about quality learning assessment, and that they had gained new ideas from peers about assessment methods and found reassurance in shared frustrations and experiences.

Multiple Gains

Marquette’s peer review assessment process, initially developed as a strategy to help the University Assessment Committee complete reviews for over 100 annual program assessment reports, has had multiple uses and outcomes that will improve and help sustain our assessment system. Peer reviews can provide valuable data for evaluation of the assessment system itself. By aggregating across programs to determine the number of programs with assessment practices at each level of performance on the five components of the assessment cycle, I gained data about the overall quality of program assessment activities and the percentage of programs successfully completing a full-assessment cycle, that is, “closing the loop.”

Understanding of collective program assessment cycle strengths and weaknesses can be gained by examining the differences between the percentage of programs meeting or exceeding the expectations and the percentage not meeting expectations for each assessment cycle component. Such comparisons provided helpful data when setting priorities for faculty development about learning assessment. In addition, by comparing these percentages across the academic years for each assessment component, I had an impacted measure of the actions we undertook during the past year to improve program assessment processes.

An example of this would be our actions to improve the assessment measures being used by each program. In 2007 fewer than half of the programs met or exceeded the Marquette Assessment System expectations for the component “assessment measures.” This was critical information. Without adequate measures, the student learning data collected would not be useful, reinforcing some faculty members’ sentiments that assessment is useless busy work. During the next academic year the director of Institutional Research and Assessment and I visited each department to assist program faculty in developing and selecting more meaningful assessment measures. In the data from the peer reviews the following year, the percentage of programs meeting or exceeding expectations for the assessment measures had grown from 47 to 82 percent.

Banta (2004) has identified thirteen “hallmarks of effective assessment practice.” The annual peer review of program assessment efforts helps the Marquette Assessment System to accomplish seven of these. Requiring the 100-plus program assessment leaders to participate together in peer reviews reinforced the fact that student learning assessment is the responsibility of program faculty and staff and kept the source for the evaluation and improvement of assessment among the faculty and staff. The table discussions at the peer-review assessment seminar, guided by a clear assessment rubric, provided faculty

(continued on page 16)
While a substantial national discourse on assessing student achievement in higher education is under way, much of it is focused at the undergraduate level. Less considered are assessments designed to measure the competences that graduate students, particularly doctoral students, need to acquire to complete their degrees. Literature that does exist on this subject highlights assessments occurring near doctoral program completion (e.g., Lovitts, 2007). Scant attention has been given to the level of thinking, reasoning, and writing skills that students need at the start of a doctoral program to be able to matriculate successfully. To address this, we report here on the development and implementation of an assessment strategy and accompanying rubric created to provide a baseline measure of these critical skills for recent applicants to a doctoral program in higher education administration.

The decision to use this strategy and create an accompanying rubric stemmed from our need to comply with reporting requirements of the state’s coordinating agency for higher education, the South Carolina Commission on Higher Education (CHE). The CHE required programs within the College of Education at the University of South Carolina (USC) to submit a report containing a minimum of five programmatic assessments, each with accompanying rubrics and data. As the coordinator of the USC higher education administration doctoral program, the first author identified rubrics currently used in the program, but realized a need to develop at least one more. The development of this rubric was informed by a review of Shulman’s (2007) work advocating that the assessments selected for inclusion in the report be viewed not as a disjointed patchwork but as a comprehensive story, a persuasive narrative accounting of our doctoral students’ development. But, what story did we want to tell the CHE about our students and our program?

One clear story the USC doctoral program faculty could tell, as could many doctoral faculty across disciplines, is about our desire to facilitate students’ growth of scholarly inquiry skills. Our doctoral students come to us as midcareer practitioners, usually higher education administrators with demonstrated expertise in a specific area, such as financial aid or admissions. They want to advance their careers with a broadened understanding of higher education administration functions and the leadership capabilities to guide those functions. They rarely come with a burning desire to sharpen their skills and identities as scholars. Thus we strive to create “scholar practitioners” skilled in approaching a question with the mindset and tools of a scholar to arrive at insights, and ultimately solutions, that might elude a less-schooled practitioner.

To facilitate students’ development of thinking, reasoning, and writing skills that undergird scholarly inquiry, the USC program offers several required research courses and a dissertation preparation retreat to jump-start students’ dissertation progress. We also stress during the interview process that this is a research degree and those accepted into the program are required to undertake independent inquiry. However, we had made little attempt previously to measure applicants’ readiness to engage in inquiry. Instead, we relied on scores from standardized examinations (i.e., the GRE or MAT), grades earned from methods classes completed perhaps years or even decades earlier, and applicants’ written responses to take-home essay questions about a current writing chronicling an issue or concern in higher education. While their essay response allowed us to assess applicants’ writing skills and, to some extent, their critical thinking skills, in reality it was never heavily considered in acceptance decisions unless an applicant performed poorly.

Thus a confluence of events set the stage for the development and implementation of a new assessment strategy and rubric to measure applicants’ baseline research skills. The story we decided to tell through our CHE report was one of...
the USC faculty regularly monitoring the growth of students’ capacity to engage in inquiry during their time in our doctoral program. Already in place was a research skill assessment administered during a required course offered in doctoral students’ first year, designed to introduce them to paradigms of inquiry, as well as assessments to measure skill attainment on their dissertation proposal and dissertation study. Secondary assessments measured the extent to which students participated in activities based on scholarly engagement, such as conference presentations and publications from their developing doctoral studies. However, a baseline assessment of newly admitted students’ readiness to engage in scholarly thinking, reasoning, and writing was absent; the story was missing its introduction.

Concurrent to work on the doctoral program’s CHE report, the first author was also involved in a large-scale multidisciplinary study investigating the effect of graduate students’ engagement as teaching assistants on the development of their research skills. This study required participants to submit a research proposal, which was then evaluated using the Universal Lab Rubric (Timmerman, Johnson, and Payne, 2007). The rationale behind the use of this rubric was that commonalities existed in defined research competence across disciplines. Whether a scholar came from a biological or a civic engineering background, his or her work could be measured against a common metric: To what extent were the questions posed significant, the work undertaken anchored within current contextual knowledge, the method used rigorous, and the interpretations drawn illuminative of new understandings?

The idea of a common template to assess research skills appeared fruitful to apply to the development of a rubric to assess doctoral applicants’ readiness to engage in scholarly thinking, reasoning, and writing. This common template was also represented in journal articles in our field in the use of headings such as introduction, theoretical framework, literature review, methods, findings, and conclusion. So it seemed natural to assess applicants’ readiness to engage with scholarly literature through their critiques of a current empirically based article guided by questions designed to assess baseline capacity to engage in scholarly thinking, reasoning, and writing.

The rubric was applied to applicants’ critiques of a current scholarly article. An additional twist was to select an article directly relevant to the situation at hand, that of doctoral education. The first author selected a recent article on doctoral advising (Barnes & Austin, 2008) and developed questions to guide applicants’ critiques of this article. These questions were designed to reveal applicants’ readiness to engage with and critically analyze key parts of the common template mentioned above. Applicants were asked to characterize the strength of the literature review and the study’s theoretical framework, identify methodological strengths and weaknesses, articulate what, if anything, surprised them about the study results, and identify what, if anything, they would add to the article’s discussion section. Additionally, applicants were asked to consider how they might apply the findings to a future doctoral advisor-advisee relationship if accepted into the program. The rubric for assessing the applicant’s characterization of the article’s strengths and weaknesses of the study method across a three-point scale is presented in Table 1.

Of the twenty-nine applicants who applied to the USC program, fifteen submitted an application file judged independently by at least two faculty evaluators as containing sufficiently high standardized test scores, graduate and undergraduate grade point averages, solid letters of reference, and a notable personal statement and, thus, were deemed to have made the “paper cut” and were invited to participate in the article critique activity and interview with at least two faculty evaluators. Prior to their campus interview, applicants were instructed to evaluate the selected article using the questions outlined above. Applicants were advised to use APA guidelines; suggested paper length was between five and seven pages.

All fifteen applicants asked to submit article critiques did so at the time of their campus interviews. Each critique was stripped of identifying information and then rated independently by two faculty evaluators. One faculty evaluator rated all fifteen critiques, a second faculty member rated eight critiques, and a third faculty member rated seven critiques. Inter-rater reliability between the first and second faculty raters was \( r = .77 \); inter-rater reliability between the first and the third rater

<table>
<thead>
<tr>
<th>Table 1: Characterization of Article’s Strengths and Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1 point</strong></td>
</tr>
<tr>
<td>Writer identifies strengths and weaknesses of study method and provides evidence to support claims.</td>
</tr>
</tbody>
</table>

5 = Exceeds
3.5 = Meets
<3 = Below

Assessment Update • September–October 2010 • Volume 22, Number 5 • © 2010 Wiley Periodicals, Inc. • DOI 10.1002/au
was $r = .98$. Based on total rubric scores and interview performance, nine applicants were offered program admission.

In subsequent discussions of the admission process, faculty noted that use of the article critique and accompanying rubric added a measure of precision to doctoral selection that had been missing. Faculty are now assured that incoming doctoral students demonstrate a readiness to engage in doctoral-level work through critical analysis of a current empirical article. Additionally, the story of the developmental trajectory of our doctoral students’ growth as research practitioners can be told fully, either on an individual student basis or in the aggregate, to determine the overall programmatic effectiveness in developing competent scholar practitioners. Nevertheless, the development of the rubric continues, as it has attracted the interest of higher education faculty for use in other programs.

The second author and her colleagues are in the higher education administration concentration at the University of Massachusetts Amherst (UMASS-Amherst), and also sought to assess doctoral applicants’ readiness for doctoral-level work. Although many of our applicants work full time as higher education administrators, increasingly we are attracting, admitting, and enrolling a greater number of applicants who are interested in pursuing doctoral studies on a full-time basis and who want to enter the professoriate after degree completion. Therefore, it is increasingly important to find ways to ascertain these applicants’ readiness to engage in scholarly thinking, reasoning, and writing prior to beginning their doctoral studies.

For the first time, during the UMASS-Amherst 2010–11 doctoral admissions process, we added the article critique, as outlined above, to our application procedure. We believe that this assessment tool allows us to accomplish three goals that will contribute to the betterment of our doctoral program. First, the article critique provides additional evidence about applicants who may not have the baseline skills that are needed at the onset of doctoral study. Second, it helps us as stewards of our field to continue to assess and readdress our definition of doctoral “readiness.” Finally, the article critique makes us more cognizant of the courses we are offering that continually advance our students’ scholarly thinking, reasoning, and writing skills, and thus enable them to complete their doctorates, particularly the independent research necessary to complete a dissertation (Lovitts, 2008).

As preparation for including the article critique in the UMASS-Amherst admission process, we selected an article centered on doctoral education, thus making both the guiding questions as well as the rubric originally designed by the USC higher education faculty particularly relevant for our purpose. In this way, UMASS-Amherst can tell its own story of how our program cultivates our doctoral students’ development.

### References


Michelle A. Maher is associate professor of higher education administration at the University of South Carolina in Columbia, and Benita J. Barnes is assistant professor of higher education administration at the University of Massachusetts Amherst.
From the States

The Voluntary Framework of Accountability (VFA)

Peter T. Ewell

In the wake of the Spellings Commission hearings and subsequent report, most of the major Washington higher education associations scrambled to demonstrate that they were constructively on board the accountability bandwagon by creating performance reporting templates for their member institutions. Probably the most widely recognized of these was the Voluntary System of Accountability (VSA) developed jointly by the Association of Public Land-Grant Universities (APLU) and the American Association of State Colleges and Universities (AASCU). Among the VSA’s several “cousins” were the University and College Accountability Network (UCAN) produced by the National Association of Independent Colleges and Universities (NAICU) and Transparency by Design (TBD) developed by a consortium of adult-serving institutions. All of these are voluntary and all contain at least some common comparative measures of performance such as graduation rates. For two of them, these common measures include standardized tests. In the case of the VSA, institutions are offered a choice of the Collegiate Learning Assessment (CLA), the ACT Collegiate Assessment of Academic Proficiency (CAAP), or the ETS Proficiency Profile. For TBD, the ETS Proficiency Profile is the probable choice.

Prominently missing from this array of voluntary reporting templates were the nation’s public two-year institutions. Three years later, this absence is being addressed by the Voluntary Framework of Accountability (VFA), currently under development by the American Association of Community Colleges (AACC), the Association of Community College Trustees, and the College Board.

Community colleges are among the most distinctive types of institutions in American postsecondary education, and this affects their basic attitude toward accountability and public reporting. First, they serve a variety of different functions simultaneously including providing a) the first two years of a baccalaureate degree, b) associate degree instruction in many vocational fields that also carries transfer credit, c) terminal occupational certification that has immediate workplace value (both associate level and certification) but does not carry transfer credit, d) remedial and developmental instruction to render students college-ready, e) noncredit instruction such as literacy training and English as a Second Language (ESL), and f) contract training for employers and local businesses. This multimission character poses significant challenges to the application of traditional conceptions of institutional effectiveness because institutional effectiveness is usually predicated on a unitary institutional mission that defines what it means to be “effective.”

Second, community college leaders have long claimed that established measures of student progression like the Graduation Rate Survey (GRS) required by the Integrated Postsecondary Educational Data System (IPEDS) are not appropriate to their institutions because they are based on full-time first-time students—a fraction of the entering student population in most two-year college enrollments. More-

These areas of distinctiveness have frequently induced community college leaders to be wary of traditional performance measures.

Over, they do not recognize that many students come to these institutions not intending to earn a degree. These areas of distinctiveness have frequently induced community college leaders to be wary of traditional performance measures or, indeed, any common set of performance measures at all. And this wariness has been quite visible in the development of the VFA.

To begin to develop the reporting framework, the sponsors established an advisory committee of community college presidents and researchers. Initial meetings of this group revealed substantial differences among its members with respect to the basic purpose of the effort. Consistent with the VSA, many felt that the reason to develop such an
The working groups were asked to go back to the drawing board to incorporate true cohort-based graduation rate measures as well as externally benchmarked measures of student learning outcomes.

The initiative was to respond proactively to stakeholder expectations about accountability. As a result, they felt that the effort should be centered on a limited number of comparative benchmarks of performance. But just as many—largely drawn from the ranks of presidents and senior administrators—believed that the primary purpose was to guide institutional improvement. As a result, they wanted to avoid measures that looked at comparative performance and believed that the report should be customized for individual institutions. Early drafts of the Statement of Purpose for the effort reflected this tension, and it was not until its main features were presented at the American Association of Community Colleges (AACC) conference in April that the choice of accountability was clear. As Eileen Baccus, the chair of one of the working groups, succinctly put it at this meeting, “The VFA is designed to show responsiveness to ‘those who are on our backs.’”

This initial tension has also been apparent in the technical design of the template’s measures. This work was assigned to three working groups—Communications and College Engagement charged with developing ways to get large numbers of institutions to participate; Workforce, Economic, and Community Development charged with examining workforce and community impact measures; and Student Persistence and Outcomes charged with developing measures of student progression and learning. Many of the indicators these last two working groups initially suggested were familiar, including college readiness, success in completing remedial and college-level courses, various “credit accumulation” milestones (e.g., earning fifteen hours of college-level work), and degree or certificate attainment. But conspicuously missing from this initial list were any externally benchmarked measures of student learning outcomes—the most prominent ingredient of VSA. Instead, the working group on Student Persistence and Outcomes proposed a reporting method through which institutions would describe their own learning outcomes, followed by a depiction of the methods used to gather evidence of the achievement of these outcomes, without reporting specific results at all. To help guard against graduation rates being misconstrued, the working group also proposed an overall success indicator based on the extent to which students reported having achieved the goals they had in attending.

Given the apparent resolution of the purpose question, this recommendation stimulated considerable push back by some members of the Advisory Committee, who believed that the VFA should set the accountability bar at least as high as the public four-year institutions had done in the VSA. Accordingly, the working groups were asked to go back to the drawing board to incorporate true cohort-based graduation rate measures as well as externally benchmarked measures of student learning outcomes. Among the testing measures to be considered were the CLA and the ACT CAAP, as well as generic skills examinations in the ACT WorkKeys battery. But unlike the design of the VSA, nonstandardized assessment methods were also to be encouraged—for example, electronic portfolios or student work samples evaluated using a common scoring scheme—so long as they could support comparative analysis across institutions. The entire scheme will be pilot tested by a diverse group of community colleges this fall.

As of this writing, the final decision about whether or not to include comparative learning outcomes measures in VFA has not been made. If the answer is affirmative, the VFA will emerge as a strong counterpart to the VSA to help demonstrate the accountability and responsiveness of the nation’s public colleges and universities. But whatever the VFA’s eventual force and content, the higher education sector that arguably will be most critical to achieving the nation’s future goals with respect to regaining global competitiveness in educational attainment—the community college sector—is assuming collective responsibility for performance.

Peter T. Ewell is vice president of the National Center for Higher Education Management Systems in Boulder, Colorado.
Barbara E. Walvoord and Virginia Johnson Anderson’s classic volume, *Effective Grading: A Tool for Learning and Assessment in College*, first published in 1998, is about much more than grading; it uses grading as a window into the wider world of pedagogy and assessment, insisting that grading can be a powerful tool for student learning. The stated purpose of that work is “to help instructors in college classrooms use the grading process effectively for learning and to explore how it can be used for assessment in the classroom and in broader contexts, such as the department or the general education program” (p. ix). The newly issued second edition incorporates new knowledge about teaching, learning, and assessment and addresses changes in the higher education landscape over the past decade, including accrediting organizations’ increasing emphasis on assessment.

The second edition, like the first, is divided into two sections: grading in the classroom and using grading to serve broader assessment purposes. Throughout both sections, the authors argue forcefully for the integration of grading “with everything else that happens in the classroom. The grade is not an isolated artifact slapped on at the end; it is part of a system that includes shaping goals and assignments, communicating with students, helping them learn what they need, responding to them, and evaluating the quality of their work” (p. 61). This “system” focuses on the higher-order abilities and skills that college graduates need to master, asking faculty to move beyond the content “coverage” mentality. Efficiency is another key theme. The chapters on “Managing Time for Teaching, Learning, and Responding” and “Making Grading More Time Efficient” offer a wealth of suggestions on how faculty members can maximize learning while minimizing expenditure of their own time.

Throughout the volume, the authors upend conventional assumptions. To faculty who complain that students are motivated to learn only for the sake of grades, they say that “trying to keep students from caring about grades is futile. Trying to pretend that grades are not important is unrealistic. . . . Grades are the elephant in the classroom. Instead of ignoring the elephant, we want to use its power for student learning” (p. 1). To those who turn up their noses at “teaching to the test,” they answer that “if the test or assignment is right—if it really tests the central learning goals of the course—then we should teach to it. In fact, it seems criminal not to” (pp. 61–62).

Part two of the book, on using grades to serve program and general education assessment purposes, has been completely rewritten for this second edition and will be especially useful to faculty members new to assessment. Included there is a helpful chapter on developing assessment plans for grant proposals.

**“Grades are the elephant in the classroom. Instead of ignoring the elephant, we want to use its power for student learning” (p. 1).**

Susan Kahn is director of institutional effectiveness at Indiana University-Purdue University Indianapolis.
The few presidents and deans who did not view their efforts as very far along indicated that the lack of progress was tied to faculty asserting that assigning grades was sufficient evidence of student achievement. Most leaders commented that the key to advancing assessment is for it to flow more directly out of existing processes for learning. For most campuses, the next big challenge is how to use effectively an astounding amount of data to improve student learning.

Accreditation as Catalyst for the Assessment of Student Learning.

The 2009 NILOA Survey revealed that across all institutional types regional and specialized accreditation were the primary drivers for student learning outcome assessment activity. Our focus group discussions were consistent with the survey results, providing specific examples of the strong role that accreditation plays in the use of assessment results. One administrator commented that although it is acceptable for accreditation to drive assessment, the problem is that student learning outcomes assessment results are rarely being used to influence institutional improvement. Another administrator lamented that because accreditation is motivated by a compliance mentality, little attention is paid to the assessment interests and questions about student learning that are important to educational effectiveness.

Although the general complaint from campus leaders was that the compliance mentality can make assessment less meaningful, several campus leaders described using accreditation as a lever for assessment. When Roosevelt University started planning for their Higher Learning Commission (HLC) reaffirmation process, university leaders and faculty were intent on achieving a ten-year approval and suggested that the best way to achieve this would be to develop meaningful assessments while ensuring that activities and evidence satisfied the HLC requirements. The president, two associate deans, and several senior faculty leaders participated in the HLC Assessment Academy to provide further education and structured support, and devoted time to translating HLC standards into institutional purposes.

Several academic leaders pointed to a negative accreditation review as stimulating activity. As the dean of Centre College put it, “there had been lots of talk [up to then)...now we have to do it.” A president whose institution is accredited by the Middle States Association (MSA) reported that the MSA review was sobering, as it concluded that the institution had made progress, but needed to do more. This critique stimulated immediate action on his campus. While accreditation is clearly a driver for assessment, it is worrisome that the focus of assessment is too often for responding to accreditation demands, and less so for improving student learning, allocating resources, or guiding strategic planning. In this sense, accreditation may devalue assessment for improvement. As one dean noted, “We were advised by our accreditors not to post ‘directions for future study’ on our Web site, and to only post the glowing aspects of our self-study,” underscoring the long-standing tension that Peter Ewell (2009) described between assessment for improvement and assessment for accountability.

Faculty Involvement Is Key to Meaningful Assessment.

As with our focus group participants, provosts responding to the 2009 NILOA Survey indicated that engaging more faculty is the major challenge to advancing assessment. A dean commented that faculty view assessment as a “distraction from the important job of teaching” and grades as sufficient information about how well students are learning. Compounding the general complaint is that several of the current criteria for educational effectiveness, such as increased student retention and graduation rates, are considered outside the purview and interest of faculty.

On a positive note, faculty are interested in assessment evidence from authentic student work that is directly linked to teaching and learning. According to a liberal arts college dean, “faculty want to evaluate student work, and want to talk about what it demonstrates in terms of student learning.” The dean at Hobart and William Smith Colleges indicated that

Most leaders commented that the key to advancing assessment is for it to flow more directly out of existing processes for learning.
work in e-portfolios and using rubrics captured the interest of faculty and students at LaGuardia. One approach to involving faculty in assessment advocated by some campus leaders is to simply avoid the term “assessment” wherever possible. As one dean put it, “there is lots of interest in the topic, but not the technique or superstructure.” Helping faculty improve their evaluation of student work is one alternative approach. Another is to use every opportunity to ask faculty: “How do you make academic decisions about what to teach, and how do you know what your students are learning?”

Assessment Is Furthered When It Is Woven into Institutional Structures.

Nearly all campus leaders reported that their progress in assessment involved creating structures and mechanisms to support and sustain assessment activities and making assessment part of standard institutional policies and procedures. At Ohio State University, for example, departments must include student learning outcomes to modify curricular requirements and must describe methods to assess these goals. This policy helped to facilitate the gradual phasing in of required learning goals and plans for assessment across a variety of departments. Albany State University adopted a similar approach by streamlining its required reports. Now, every required report—program reviews, annual reports, and assessment reports—must include information about student learning outcomes.

Developing assessment expertise is also important. Several presidents and deans reported encountering difficulty in making the right choices about assessment tools and approaches. One dean noted, “We are confronted with a bewildering array of techniques and instruments.” Albany State developed a cadre of assessment experts by creating a rotating, two-year appointment with course release time for faculty members. Assessment directors serve their two-year term and then, armed with this experience, return to their departments. According to the director of institutional research, “The faculty member’s colleagues are now going to him [the faculty member who rotated out of the term] whenever they have assessment questions. . . . The director that we have now is from the College of Sciences and Health Professions, so we’ll work with her and then she’ll rotate back to her full college duties and we’ll pick somebody else, hopefully from the College of Business, . . . until we have faculty experts in assessment in every college.”

Deans and directors of institutional research at several institutions emphasized the importance of focusing on what faculty members are already doing in their classrooms in terms of learning outcomes assessment as the natural place to begin to advance assessment. At the University of Missouri, the director of the office of assessment focused on working with faculty in programs who signaled they wanted to work with him. His first step was asking program faculty, “What do you want your undergraduates to be able to do?”

In some institutions, faculty needed help bringing their assessment practices into the spotlight. Some faculty members found welcome connections between assessment activities and the scholarship of teaching and learning, or within centers for teaching and learning, while others benefited from being involved in cross-campus and in some cases cross-institution initiatives.

Final Thoughts

Assessment efforts are growing and deepening on campuses, primarily propelled by accreditation and other national accountability initiatives. While support from campus leadership is essential, real progress requires that faculty members take ownership of assessment processes and outcomes, particularly at small colleges and universities. In addition, infra-

Faculty are interested in assessment evidence from authentic student work that is directly linked to teaching and learning.

References:

Jillian Kinzie is associate research scientist at the Indiana University Center for Postsecondary Research in Bloomington, Indiana.
and staff hands-on training about multiple assessment methods and best practices in a collegial environment. Likewise, the focus in the reviews on examining assessment processes, as well as the outcome data, helped the program assessment leaders understand a more integrated view of assessment as an ongoing process.

We have noted that the presentation and discussion of a program’s assessment processes and findings in the small group review provided opportunities for recognition by peers of individual efforts and successes. Finally, the results from the annual peer reviews provided a vehicle for quantitatively demonstrating accountability to stakeholders and identifying areas for improvement of the Marquette assessment process itself.

After each peer-review assessment seminar, the next meeting of the University Assessment Committee is devoted to evaluation of the seminar. Feedback from the participants is examined, but equally valuable is the sharing by committee members of their observations and experiences as table facilitators. Our goal is to reach a shared conclusion about what worked and what didn’t, and identify actions to improve the next peer-review process.

In conclusion, we have found using a peer-review approach to be an effective strategy for the maintenance and further development of quality assessment at Marquette University. Faculty members understand the value of a peer-review approach, as they often participate in peer reviews of written scholarship, research presentations, and teaching evaluations. Many co-curricular staff members were not familiar with peer review. The mock role-play of a review and the table facilitator’s discussion of the ground rules for participating in a peer review usually relieved their initial caution and any fears of exposure.

Reference

Margaret Fong Bloom is vice provost for undergraduate programs and teaching at Marquette University in Milwaukee, Wisconsin.