Assessment UPdate
May–June 2014
Volume 26, Number 3
Progress, Trends, and Practices in Higher Education

Do Good Assessment Practices Measure Up to the Principles of Assessment?

Jillian Kinzie, Natasha Jankowski, Staci Provezis

THE AMERICAN ASSOCIATION FOR HIGHER EDUCATION'S (AAHE) *NINE PRINCIPLES OF GOOD PRACTICE FOR ASSESSING STUDENT LEARNING* (AAHE 1992) have appeared to stand the test of time, as evidenced by the fact that they are often referred to within the pages of *Assessment Update* and appear on various assessment websites and in texts (see Banta, Jones, and Black 2009). In fact, Hutchings, Ewell, and Banta (n.d.) reviewed the principles in 2010, declaring that they had "aged nicely." Looking back to 1992, the principles were conceived as a way to codify the responsible and effective conduct of assessment, advance assessment for educational improvement, and assist campuses to develop approaches that make a difference for students and their learning. The principles serve as a foundation for assessment practice.

Guidelines for assessment continue to be promulgated, such as the New Leadership Alliance's *Committing to Quality: Guidelines for Assessment and Accountability in Higher Education* (2012), which intended to help institutions evaluate their assessment practices and to establish shared commitments among sectors of higher education; and the *Principles for Effective Assessment of Student Achievement* (Western Association for Schools and Colleges 2013), endorsed in July 2013 by six higher education associations and all regional accreditors, which succinctly expressed the value of assessment. The newer statements share tenets of the AAHE principles and also reflect specific organizational commitments.

Similar to the AAHE Assessment Forum, the National Institute for Learning Outcomes Assessment (NILOA) has sought to move the needle on assessment efforts by surveying the landscape of assessment in higher education and by assisting institutions and others in discovering and adopting promising practices in the assessment of undergraduate student learning outcomes. Toward these ends, this article considers the most widely cited guidelines for effective assessment, namely, the AAHE Principles (Table 1, page 14), against the backdrop of NILOA's collection of accounts of good assessment practice. Simply put, how well do institutions' assessment activities align with stated principles for effectiveness?

CONTENTS

View this newsletter online at wileyonlinelibrary.com • DOI:10.1002/au.20005

JOSSEY-BASS™
A Wiley Brand

Assessment Update

Progress, Trends, and Practices in Higher Education

May–June 2014
Volume 26, Number 3

Assessment Update: Progress, Trends, and Practices in Higher Education (Print ISSN 1041-6099; online ISSN 1536-0725 at Wiley Online Library, wileyonlinelibrary.com) is published bimonthly by Wiley Subscription Services, Inc., A Wiley Company, at Jossey-Bass, One Montgomery St., Suite 1200, San Francisco, CA 94104-4594. Periodicals Postage Paid at San Francisco, CA, and additional mailing offices. Individual subscriptions are $135 per year (institutional $239). Back issues are available in limited supply at $29 per issue. To order, phone toll-free (888) 378-2537 or fax (888) 481-2665. Visit our Web site at www.josseybass.com. *Postmaster:* Send address changes to Assessment Update, Jossey-Bass, One Montgomery St., Suite 1200, San Francisco, CA 94104-4594.

JOSSEY-BASS™
A Wiley Brand

How the Principles Stack Up Against Good Assessment Practice

To understand what good work looks like related to assessment, NILOA embarked on nine institutional case studies (www.learningoutcomesassessment.org/CaseStudiesInstitutions.html), and synthesized findings in the report, *Using Assessment Results: Promising Practices of Institutions That Do it Well* (Baker, Jankowski, Provezis, and Kinzie 2012). The case-study institutions were selected for their robust assessment processes and their history of demonstrating the use of assessment evidence. We then thought it useful to consider assessment practices in relation to the AAHE principles.

In general, assessment across the nine case study sites demonstrated the following broad practices for effective assessment:

- Engaging faculty in and fostering ownership of assessment;
- Sharing assessment information with internal and external audiences;
- Embedding assessment in institutional processes;
- Ensuring administrative leadership that provides a vision for assessment and support for the professional development of faculty and staff.

These practices by and large reflect the practical wisdom described in the principles. Although our overall conclusion is that the assessment accounts from the case studies affirm the soundness of the nine principles, it is clear that some statements are less prominent, whereas others have taken on increased importance and/or have morphed to reflect advances in practice.

The key principles reflected in the case studies include the premises that assessment works best when the purpose is to improve learning and thus is embedded in the educational values of the institution (principle 1) and when programs have clear, explicitly stated purposes (principle 3). These two principles are reflected in all cases, and are exemplified in the activity at LaGuardia Community College, which emphasizes the assessment of individual student learning and demonstration of evidence in portfolios, and in Capella University's use of a backward-design approach. By

(continued on page 14)

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 aupdate@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. ■

Healthy Assessment: What Nursing Schools Can Teach Us about Effective Assessment of Student Learning

Douglas J. Eder

Twenty-five years of publications and presentations reveal continuing, widespread faculty disinclination for a "culture of assessment." Therefore, it is intriguing that nursing school faculties seem particularly adroit at doing assessment. They do it with more effectiveness and less consternation than faculties in most other disciplines do. I sensed this during my 19 years as a "pure" faculty member teaching anatomy and physiology to prenursing and premedical students. That impression grew during 15 more years of conversations with assessment scholars and my own encounters as assessment steward and visitor to approximately 140 other institutions. I wondered, "What insights do nursing faculties have about doing assessment *well*?"

Methods

I approached this question during late summer 2013 by interviewing knowledgeable people (e.g., nursing deans, associate deans, assessment coordinators) at 19 nursing schools and programs nationwide . . . large/small, public/private, 4-year and 2-year. Interviews consisted of seven open-ended questions (see Table 1). Multiple responses were allowed; content analysis permitted categorizing responses, which sometimes occupied more than one category. Interviews lasted 15–20 minutes. The lists of participants and questions appear in Table 2.

Table 1. The seven open-ended interview questions (multiple responses possible):

1. For whom do you assess—and why? In other words, why are you going through these assessment efforts?
2. What do you assess? (e.g., content, critical thinking, clinical judgment, writing, speaking…)
3. How do you manage the information and data flow? Do you use software to assist? Graduate assistants? Faculty assessment scholars? Something else?
4. What kinds of assessment do you do to capture the information you seek? (Student portfolios? Student projects? Simulations? Papers? Exams? Direct expert observation? Videos?…)
5. What feedback on the findings is received by professors and the program? What do they do with it?
6. What feedback on the findings is received by students in the program? How do they get it? What responses are expected from them as a consequence?
7. Especially in the arts and sciences (but not limited to those disciplines), students graduate by passing exams or projects of some sort that have grades of, say, "78%" or "89%." In flying an airplane or building a college residence hall, landing an airplane safely 78% of the time or making masonry walls of the residence hall 89% vertical is NOT acceptable. Similarly, nurses are expected to be virtually perfect in their work, too. Educating and training for near-perfection is often called "mastery-level learning." My final question of the interview asks, "Do you teach for mastery? If so, how is this assessed?"

Table 2. List of participants:

1. Blackhawk Technical College (Wisconsin)
2. SUNY Upstate University near Syracuse
3. Patty Hanks Shelton School of Nursing
4. Johnson County Community College
5. University of San Francisco
6. Michigan State University
7. Salish Kootenai College
8. Creighton University
9. Purdue University
10. Linfield College
11. University of Kentucky
12. Wichita State University
13. Abilene Christian University
14. New Mexico State University
15. East Tennessee State University
16. Stephen F. Austin State University
17. Laramie County Community College
18. Texas State Technical College of Nursing
19. Oklahoma State University–Oklahoma City

Findings

1. For Whom Do You Do Assessment? Why? Every respondent said, "We do this for the students, to assure that they are learning well." Indeed, the "for-the-students" response was unanimous (100%) and it was the *first* reason mentioned in all but two of the interviews. When I have asked this question in other disciplines, such as those in arts and sciences and business, the response has usually been, "For the accreditor."

Additional responses were: for ourselves, our program, and our university to assure that we are teaching well = 74%; for our accreditor = 58%; and for the public and for the profession = 26%.

Commentary: External accreditation pressures doubtlessly influence many nursing school practices. Nevertheless, part of nursing schools' successes with assessment might be because the faculties do it for a reason that actually matters, namely, to improve student learning. This reason might be abetted by the drive to have students pass the nursing board exams. Further commentary appears in the discussion section below.

2. What Do You Assess? Respondents said they assessed the following properties of student learning:

- Critical thinking/clinical judgment = 95%
- Communication skills = 53%
- Disciplinary/technical knowledge = 44%
- Disciplinary skills = 44%
- Levels of nursing essentials = 37%
- Professionalism = 26%
- Reflective thinking, science, aesthetics, leadership, evaluation of credible sources = 1 response each

Commentary: Nursing school respondents emphasized that they assess what matters to them, not what is prescribed by outsiders. Without doubt, critical thinking/judgment in the clinical environment stands out as a high-priority assessment because it matches what nursing faculties teach and what they expect students to learn. Therefore, actively monitoring and assessing student learning has merged into pedagogy.

3. How Do You Manage the Information/Data?

- We do it by hand, we collect it ourselves, we use our own spreadsheets = 63%
- We use course evaluations = 21%
- We use electronic surveys and an internal software system = 16%
- We receive assistance from a university software system = 11%
- Via alumni surveys = 11%

4. What Kinds of Assessment Artifacts Do You Collect?

- Exams = 95%
- Simulator observations = 84%
- Clinic observations = 63%
- Papers = 53%
- Standardized patients/nursing labs = 47%
- Projects = 47%
- Portfolios = 26%

Commentary: The following phrases often accompanied the "by hand" descriptions in Question 3: "We struggle with this," "painful," "we're taking baby steps," "it's a challenge," "we're limping along," "it's a mish-mash," "we're exploring for a good system and we're not there yet." Nursing schools and programs do not seem unusual in what they assess, as reported in Question 4, although the emphasis on exams is high. The requirement that nurses pass licensing board exams makes this emphasis understandable. Nevertheless, the assessed artifacts are common, and they present no data management problems unique to nursing. What seemed amazing (but not unusual) was that most data management was reported as done by hand.

5. How Does Feedback on Student Learning Get Back to the Faculty/Program So It Can Improve?

- By individual instructor reports to faculty . . . usually involving a self-improvement plan = 53%
- An established system of monthly or end-of-semester committee meetings = 47%
- Via written student, course, or site evaluations = 21%
- Through one main person = 11%
- Through larger institutional mechanisms = 5%
- Via e-portfolios = 5%

Commentary: Typical conversations that accompanied responses to this question included, "Good question," and "We have difficulty grappling with the data." Respondents stated emphatically that assessment findings garnered official attention at faculty meetings at least once every semester or, commonly, once a month. Corresponding adjustments to the curriculum—including changes in lecture materials and clinical exercises—took place within months, which is certainly warp speed in most academic environments!

Without doubt, critical thinking/judgment in the clinical environment stands out as a high-priority assessment because it matches what nursing faculties teach and what they expect students to learn.

6. How Does Feedback on Student Learning Get Back to the Students Themselves So They Can Improve?

- Students have direct, individual access to independent, electronic feedback systems = 63%
- Students receive individual feedback from professors = 47%
- Students serve on nursing school committees, receive information in group form = 44%
- Via published rubrics and individualized improvement plans = 26%

(continued on page 13)

Development of a National Survey for Secondary Mathematics Teacher Education Programs

Yukiko Maeda, Jill Newton, Vivian Alexander, Sharon L. Senk

The *Preparing to Teach Algebra (PTA)* study was designed to investigate opportunities that secondary preservice mathematics teachers have to learn about algebra; algebra teaching; equity issues related to algebra learning; and the algebra, functions, and modeling in the *Common Core State Standards for Mathematics (CCSSM)* (National Governor's Association Center for Best Practices and Council of Chief State School Officers 2010). As part of the study, a web-based national survey was developed to provide a snapshot of the current status of opportunities in these areas in secondary mathematics teacher education (SMTE) programs across the United States. The goals of this report are to share key survey findings and to highlight challenges encountered during the process of developing and administering the *PTA* survey. Throughout this report we hope to provide guidance for survey development and administration for researchers and practitioners who are interested in conducting large-scale program surveys across higher education institutions.

Survey Development

Survey development and the selection of survey sample programs occurred concurrently. To aid in the development of the survey, we referenced previous surveys that aimed to study the characteristics of mathematics teacher education programs and preservice teachers. However, because these surveys primarily focused on general program characteristics, which were not the only interest of our survey, a large number of additional items related to opportunities to learn (OTL) were developed.

Each item on the initial draft of the survey was evaluated by *PTA* researchers and an external advisory board specializing in mathematics, mathematics education, and survey methodology for its wording, alignment with the research questions, and ease of responding by potential participants. In addition, pilot testing at three institutions provided insight into the thought processes of respondents and the time needed to complete each part of the survey. More specifically, faculty members in these pilot programs participated in think-aloud sessions as well as interviews to share suggestions for revising the survey. As a result, the draft survey was significantly modified by reducing the number of items, changing item formats, and altering the language used in items. The final version of the survey included 13 multiple-choice and 7 open-ended questions related to (1) general characteristics of the largest SMTE program at the institution; (2) opportunities to learn algebra, how to teach algebra, equity issues in algebra, and algebra topics included in *CCSSM* in that program; and (3) additional information.

The *Preparing to Teach Algebra (PTA)* study was designed to investigate opportunities that secondary preservice mathematics teachers have to learn about algebra; algebra teaching; equity issues related to algebra learning; and the algebra, functions, and modeling in the *Common Core State Standards for Mathematics (CCSSM).*

Survey Sampling

Based on our initial power analysis and to account for the probability of nonresponses, we planned to sample 400 SMTE programs stratified by institution type. A variety of pathways are currently available for preservice teachers to become secondary mathematics teachers. For example, some programs were entirely online, whereas others were jointly offered by two or more institutions. Some institutions offered multiple teacher education programs, for example, a 4-year undergraduate program and a 15-month postbaccalaureate program, both leading to initial licensure to teach mathematics in grades 7–12. However, because no

comprehensive list of SMTE programs existed when we began our work, we encountered difficulties in creating a representative sample of the target population of SMTEs. Accordingly, we had to develop our own sampling frame with the use of the Carnegie Foundation's Basic Classification of Baccalaureate, Master's, and Doctoral degree-granting institutions (Carnegie Foundation for the Advancement of Teaching n.d.) as the desired strata with the conjecture that the nature of programs and OTL may differ by institution type.

After potential institutions were identified through sampling, the websites of the selected schools were examined to determine the existence of an SMTE program and to locate a contact person for the program. Because of substantial variation in the information that could be obtained from the websites, it was often not possible to determine whether the institution had an SMTE program or to locate the contact person using information available online. In such cases, we called or sent e-mails to the academic departments associated with the program. If we determined that no program existed, we resampled institutions within the same stratum until three lists, one for each of baccalaureate (n = 176), master's (n = 160), and doctoral degree granting institutions (n = 64), totaling 400 institutions with STME programs, were compiled.

Survey Administration

A week before the survey was distributed, we sent a prenotice e-mail to inform respondents that the main survey would be sent within 1 week and also to provide them with a list of the information that they would need to complete the survey. About two weeks after the survey administration, we sent a reminder e-mail to the 334 nonrespondents to encourage them to take the survey. To encourage survey participation further, additional follow-up contacts were made via phone to further encourage survey participation before the survey was closed. The survey was officially closed six weeks after its original distribution. We received valid responses from 131 secondary mathematics programs. The response rate was 33%.

Key Survey Findings and Efforts for Maintaining Response Validity

Survey respondents indicated that SMTE programs provide more opportunities to learn algebra (56% of the programs reported that they provide opportunities to learn algebra to a great extent) than to learn how to teach algebra (35% reported that they provide opportunities to learn to teach algebra to a great extent). Opportunities to learn about equity issues related to algebra learning were particularly rare (11% of the programs reported providing opportunities to learn about equity issues related to algebra to a great extent). The survey responses also indicated that most SMTE programs are not providing opportunities for preservice teachers to take the specialized mathematics courses for teachers that are recommended by professional societies (Conference Board of the Mathematical Sciences 2012; Newton, Maeda, Alexander, and Senk 2014).

The results of the *PTA* national survey illustrated the diversity and complexity of SMTE programs for initial licensure in the United States.

While conducting the *PTA* survey, we also encountered several issues related to balancing the validity of the survey results and maintaining response rates. Our main concern was that potential respondents (i.e., mathematics or mathematics education faculty, program coordinators, or instructors) would experience difficulty in accessing the information needed to respond to the items regarding both program administration and the content of courses. For example, although information about the number of graduates in the last 3 years is likely available in a program or institutional database, it may not be readily available to the survey respondents. In addition, whereas a faculty member may know details about learning opportunities in a particular course, he or she may not have such information for *all* required courses in the program. In order to maximize the response rate, we wanted the time and effort that potential respondents would need to gather the information required to answer the survey questions to be reasonable (Dillman, Smyth, and Christian 2009). We addressed this issue by asking only key questions, reducing the depth of questions, and outlining the information needed to complete the survey in the survey invitation e-mail and at the beginning of the survey.

We also realized that some of the key terms have the potential for being interpreted differently across respondents; to minimize the impact of a range of definitions, we defined key terms at the beginning of the survey. In the PTA study, OTL is a key construct that has been defined and utilized in multiple ways. For example, Husén (1967) described OTL as "whether or not the students have had the opportunity to study a particular topic or learn how to solve a particular type of problem presented by the test" (162). More recently, OTL has been defined in terms of curricular coverage, exposure to content, and emphasis on the content (Stevens 1993). We considered only the mathematics, mathematics education, and general education courses required for all students as possible sites for providing OTL in a program.

Concluding Remarks

The results of the *PTA* national survey illustrated the diversity and complexity of SMTE programs for initial licensure in the United States. The process of conducting this survey revealed the challenges of balancing the richness of data, response rates, and validity of responses. The lack of a systematic database of SMTE programs not only makes it difficult for researchers to create a representative sample, but also limits the ability of mathematics

(continued on page 12)

From the States

New Looks in Specialized Accreditation

Peter T. Ewell

Specialized accreditors in the United States—those accrediting organizations that accredit specific academic programs in professional or vocational fields—have always fallen into two groups with respect to assessment. Those that accredit programs leading to practice in a licensed profession—like nursing, education, or one of the health professions—are fairly prescriptive about assessment and, of course, make prominent use of the licensure examination in their fields as well as structured observations of internships or field placements. The other specialized accreditors, in contrast, have gradually ratcheted up their requirements for evidence of student learning, but have been far less prescriptive. Evidence that this trend continues is apparent in two sets of specialized accreditation standards adopted last year and currently coming into force: the Council for the Accreditation of Educator Preparation (CAEP) and the Association to Advance Collegiate Schools of Business (AACSB) International.

CAEP is the result of an amalgam of two prior accreditors, the National Council for the Accreditation of Teacher Education (NCATE) and the Teacher Education Accrediting Council (TEAC), which merged about three years ago. Creation of a new accrediting organization, naturally, resulted in a need for a new set of standards, so CAEP chartered a Commission to create some. The new standards, unveiled in August 2013, significantly raised the profile of teacher education programs by establishing new levels of admissions selectivity; teacher education is commonly seen as enrolling less capable students. The standards also established unusually aggressive measures of program effectiveness embodied most prominently in a common set of annual indicators. CAEP's new standards also allow the accreditor to award a higher level of recognition for programs that excel; this action has been talked about in the accreditation community for years but has never been adopted by an accreditor until now.

Classroom observation is labor intensive and requires considerable observer training for ratings to be reliable. And only a few states at the moment collect value-added data and none of them do so for the full range of subjects and grade levels.

The 10 annual indicators that CAEP is putting in place are intended to play a prominent role in the accreditation process. Traditionally, such indicators have been used by accreditors to keep track of institutional or program conditions as background for a review. In addition to this, CAEP intends to establish threshold levels of performance on many of them to help determine a program's accredited status and to award "accreditation with distinction." The list of annual indicators itself, moreover, contains familiar measures like retention/graduation rates and licensure test scores that have been used by specialized accreditors for years. But it also has a couple of new and highly controversial measures that look at the performance of a program's graduates after they have become classroom teachers. The first is based on classroom observations in which trained raters observe a graduate's behaviors in an actual classroom setting after she or he has been employed and rate it against a protocol listing teaching practices that are known to be effective. The second is value-added scores for the pupils of graduates of the teacher preparation programs collected soon after these graduates were hired as teachers. (The same measure was proposed by the US Department of Education two years ago as a measure of the effectiveness of teacher education programs under Title II, but was subsequently withdrawn.) Proponents of these measures argue that they are far more valid as quality indicators than more traditional indicators like preservice test scores or grades in education school courses because they directly reflect the phenomenon of interest: performance in the classroom after candidates have become teachers. Detractors decry them because they remain largely untested and, in the case of value-added measures, rely on a causal inference that many find dubious. They also suffer from the drawback that data are not readily available. Classroom observation is labor intensive and requires considerable observer training for ratings to be reliable. And only a few states at the moment collect value-added data and none of them do so for the full range of subjects and grade levels.

Indeed, problems with data are challenging the CAEP indicators in many ways, and they affect the more traditional measures as well. Graduation rates, for example, are conceptually straightforward at the institutional level, where it is clear at what point a student enters and therefore becomes part of a tracking cohort. On a program level, though, things become a good deal less manageable because students may become teacher education majors at any point in their first two years. Further complicating matters is the fact that many teacher education programs are graduate programs. The good news here is that they all enter at the same time, but there is no national standard for the completion of a two-year master's program against which to benchmark performance. As at the institutional level, moreover, the presence of large numbers of part-time attenders considerably complicates the calculation of uniform graduation rates. The news is potentially better for employment outcomes because the necessary data can be obtained from unemployment insurance (UI) wage record databases held by each state's Department of Labor. UI wage record databases are unsuitable for tracking the graduates of many professional programs because most of them only report the industry in which the graduate is employed, not the actual occupation she or he has, and none of them contains data on individuals who are self-employed. But this is not a drawback for education because the industry and occupation are the same, and almost nobody is self-employed.

Potential use of UI wage record data underlines another important condition of the CAEP measures, however. Many of the most important sources of data are held by states. This is true of licensure test scores and value-added data as well as UI wage record data. This unusual level of dependence on state data, rather than data reported by teacher education programs themselves, means that CAEP is forced to pay considerable attention to cultivating state departments of education. Fortunately, relationships with state agencies are already strong because of state licensure of teacher education and the ready availability of licensure examination pass-rate data through Title II reporting. But these relationships in many states are not yet mature. What all of this means is that very few eligible teacher education programs will be in a position to meet CAEP's new standards in the first few years they are in effect, and the organization will have to continue to accredit on the basis of measures of graduate effectiveness provided by the programs themselves for some time to come.

Despite its different, and far less prescriptive, stance with respect to assessment, AACSB International's requirements are clearly having an effect on the practice of assessment in business programs.

AACSB International, meanwhile, is representative of the other type of specialized accreditor. Within its purview, only accounting is a licensed profession with test scores available for use in accreditation through the Certified Public Accountant (CPA) examination. Accordingly, AACSB International's requirements for assessment, termed "Assurance of Learning," are decentralized and determined by the program itself. Much like the requirements of regional accreditors, business programs are told to set learning outcomes goals consistent with their institution's mission, containing a mix of generic competences like communication and field-specific competences like evaluating a company's business plan. They are then free to choose or develop appropriate assessment methods to collect evidence that students are achieving these learning goals. And, like regional accreditors, AACSB places much of the weight of "Assurance of Learning" on use of the resulting information in continuous program improvement.

Despite its different, and far less prescriptive, stance with respect to assessment, AACSB International's requirements are clearly having an effect on the practice of assessment in business programs. When the National Institute for Learning Outcomes Assessment (NILOA) surveyed program leaders on their assessment practices in 2010, all programs that were accredited reported that the requirements of their specialized accreditor were among the top reasons they were doing assessment at all (Ewell, Paulson, and Kinzie 2011). Partly stimulated by AACSB International, programs in business also showed a distinctive pattern of activity with respect to assessment. Of the 12 programmatic areas addressed by the survey, business programs were fourth in overall level of assessment activity, second in using capstone courses as a principal method of assessment, and second in reporting high levels of faculty involvement in assessment. But the most interesting finding is that business programs were the most likely of all to report that programmatic assessment needed better instruments. This last finding is ironic in light of the fact that AACSB made major efforts in the mid-1980s to develop a standard assessment in business under the leadership of its Managing Director for Accreditation Services, Milton Blood, who then abandoned the effort. All in all, however, the leaders of AACSB International have a

(continued on page 12)

Community College Strategies

Curriculum-Level Assessment: Finding Meaning and Focus for Real Results

Dianna Renz

Addressing Student Learning

Why address student learning? Almost all accrediting agencies require some evidence of student learning; taxpayers and donors alike expect that student learning occurs at an institution of higher education; parents fervently hope their children are learning something; students themselves would admit they want to learn—if for no other reason than improving their employment odds. But really, why should higher education institutions care about student learning? Isn't it sufficient to increase student numbers and rely on FTE funding? No. Most educators will willingly share that they did not enter the field of education for any reason other than that they DO care about student learning. We must address student learning because it's the right thing to do.

Here at Western Wyoming Community College (Western), student learning is tied to our Guiding Principles, a set of institutional values derived from our more lengthy mission statement. The first two of six principles are directly related to student learning: "learning is our purpose" and "students are our focus." Evidence of student learning also is a requirement of our accrediting body, the Higher Learning Commission of the North Central Association (HLC). Student learning fulfills the HLC criterion of "Academic Program Evaluation & Improvement." After test-driving the process from 2006, Western Wyoming Community College was accepted into the Academic Quality Improvement Program (AQIP Pathway) of the Higher Learning Commission in 2008. Student learning evaluation addresses four of AQIP's nine categories: student learning, understanding students' (and other stakeholders') needs, measuring effectiveness, and planning continuous improvement.

Student Learning Goals

If we accept the fact that institutions should, and in most cases, must evaluate student learning, then we arrive at the natural next question: How can we evaluate student learning? The most important tool here is to choose a lens through which to focus. An institution can gather data from a variety of different sources, and can collect those data, and even look at the possible interpretations of individual pieces of data, but no solid conclusions can be drawn from this haphazard approach. Unless there is a scaffold, or framework, on which to hang the data, no meaning can be derived from individual data points. In 1994, Western faculty developed five goals for student success. These goals exist at the curriculum level, and apply to all Western graduates. Whether the student completes a degree in art, nursing, history, or welding, we expect all those majors to be able to communicate competently, retrieve information, solve problems, see issues from multiple perspectives, and develop life skills. These goals are more clearly defined with various objectives. Although individual institutions will have different learning goals based on their unique audiences, geographic locations, cultural expectations, and areas of expertise, agreeing to a set of curriculum-level learning goals is an essential first step in assessing student learning.

A Cross-Institutional Team

Accepting a student learning framework is logical, but who does all that work? Who should evaluate student learning? It is easy to say that this evaluation occurs with the faculty members who are teaching in the classroom. However, each faculty member is charged with accomplishing the goals of the individual courses he or she teaches, in agreement with the expectations of that particular department or division, in accordance with any articulations with other institutions, and in collaboration with other content-area instructors. The instructor's role is not enviable: scoring work from 80

If we accept the fact that institutions should, and in most cases, must evaluate student learning, then we arrive at the natural next question: How can we evaluate student learning?

to 500 individual students, trying to assess whether the mass of those persons have, in fact, grasped the essential concepts of the particular content area. No, institutional assessment of student learning cannot rest with the faculty alone; they are working hard to ensure student learning at the course level—not the curriculum level. (A side note: at our institution, the Program Review process allows programs to tie specific courses—and their accompanying course goals—to the program's specific learning goals, and then show how the program goals support the goals for student success. However, when the individual instructor is teaching a content-specific course, he or she is focused in a very narrow view—and necessarily so! Evaluation of student learning at the curriculum level cannot, and should not, rest with the faculty members alone.)

A cross-institutional team is the necessary tool for evaluation of student learning. Here at Western, our Assessment of Student Learning Team is composed of 13 voting members. Seven of those are full-time faculty members, with one representative from each academic division. Six additional members are professional and paraprofessional employees, representing the Student Success Services area, the Administrative Services area, Outreach coordinators (we have 13 Outreach sites), and the Distance Learning program (in our rural service area, we have a significant number of course offerings online), for example. Coincidentally, four of our six professional/paraprofessional team members also serve as adjunct instructors. Additional team participants are nonvoting members; these include the Associate Vice President for Planning & Improvement (Chair), the Assessment Coordinator, and the Planning & Improvement Assistant (Recorder). Other administrative staff members—including our president, three vice presidents, and two associate vice presidents—have a permanent guest invitation.

Gathering and analyzing individual data components is like trying to see the big picture by looking only at one puzzle piece. Looking at the big picture—the mosaic, if you will—is feasible only with the scaffold of the goals for student success (or a similar framework).

Membership Rotation

One aspect that makes Western's A-Team unique is the specified membership rotation. Voting members serve three-year terms. As with many campus committees, the work we do is complex, and it takes most members a full year of participation before they conceptualize the framework, the assessment measures, and the analysis of results. Two-year terms are common on other committees, but we find that the third year really allows individuals to feel competent about discussing student learning outcomes. At the end of the three years, two-thirds of the voting members rotate off the team, and are replaced by new representatives from their respective areas. One-third of the voting members remain for a second three-year term. New membership provides fresh perspective, and also ensures that the various departments, divisions, and programs receive A-Team reports from a diversity of voices and faces. The second three-year term also ensures shared leadership and team memory; more experienced members impart their knowledge to newer members.

Data Analysis

Once the institution has a team to do the work, and a framework for the results, it's time to gather data. Gathering and analyzing individual data components is like trying to see the big picture by looking only at one puzzle piece. Looking at the big picture—the mosaic, if you will—is feasible only with the scaffold of the goals for student success (or a similar framework). At Western, our A-team uses a wide variety of external, internal, direct, and indirect measures to evaluate student learning. We use external direct data sources, such as nationally standardized tests, and transfer data from our local university. Internal direct data sources include actual student work samples; our Planning and Improvement office gathers research writing samples from all graduates in the current academic year, removes identifying information, and randomly selects 50 essays, which are then scored by A-Team members with a standard rubric. External indirect sources of student learning assessment include specific questions on nationally standardized surveys, such as the Community College Survey of Student Engagement. We round out this effort with some internal indirect data sources: a faculty survey within the first three weeks of each term, and a student survey in the fifth or sixth week of each term. This variety of sources allows us to examine the scope and depth of student learning at Western. We do not look at any one source alone, but attach each measure to one (or more) of the goals for student success. We use reading and writing test scores from the ETS Proficiency Profile, for example,

to inform ourselves about how well students learn to communicate competently, but we use the mathematics and critical thinking portion of the test to help define how well students solve problems, for example.

The complexity of this mosaic is challenging to conceive, and certainly difficult to track longitudinally. We first created a measures table to help us see student progress in the various goals for student success. The table includes the specific measure, the benchmarked goal, and several years of results (for the purpose of comparison).

Influencing Student Learning

Regardless of the framework you build, or the team you construct, or the various data points you track, unless you can use these tools to influence student learning, your efforts have been wasted. The goal of this work, then, is to help the team make data-informed (not data-driven) decisions about how to direct student learning. At Western, our A-Team meets for 2 days in June each summer. Prior to this meeting, A-Team members have already done the work of scoring the student work samples mentioned previously for research writing, reflective writing, and oral presentation, using agreed-upon rubrics and a three-person scoring panel to ensure interrater reliability. We discuss and evaluate all measures and results for the goals for student success. From these results, we use an AQIP-style strategy forum to determine the direction of a new Student Learning Action Project. The team chooses just a vital few areas on which to focus. The objective is to locate those critical learning components that will have a systemic impact on all other learning goals, and in the end, achieve the largest effect with a very focused use of resources and time.

The Action Project approach is encouraged by the AQIP, and is a component of the annual strategy forum training for member institutions. We have modified the process for our summer A-Team workshop. Although we struggled at first, after three years of designing Action Projects, the process has become quite efficient and very effective. First, we identify the challenge with a clear statement. Second, we examine the roadblocks that prevent success in the targeted area. Third, we brainstorm solutions in small groups, then share with the larger team. Last, we determine our vital few areas of focus; for the past several years these have been variations of Writing Across the Curriculum (this year's project specifically focuses on research writing), and Active Reading in the Content Areas.

Clearly, student learning is an essential component of our shared efforts in higher education. Taking the time to create a framework for evaluation, form a functional team to focus the analysis, and devise a solid assessment plan is worth the effort.

The Action Project proposals required by AQIP are very detailed and assist the team to identify an area of focus and plan specific dates, tasks, and events, complete with appropriate goals for completion. When we saw that our graduates' reading scores continued to decline for several consecutive years, we took action. We solicited best practices in active reading from our faculty members and aligned those ideas with research-based practices. We used this information to create a dynamic brochure for students, which was then distributed in targeted entry-level classes (we did not target English classes, as active reading in the content area is a content-specific task—not just the job of English teachers). We used this same information to host on-campus workshops to assist students in learning and practicing the active reading strategies. Faculty members served on roundtable forums regarding active reading in the content areas during our fall and spring in-service sessions.

Our efforts have evolved since this Active Reading Action Project was initiated in 2009. After workshop attendance dropped off, for example, the team elected to create brief video recordings of the techniques for posting online; this effort is now in progress. Our best evidence of success is in the data: reading (and writing) scores have leveled out and are once again above the national average, and our transfer data show that, after a few years of poor performance, our transfer students are again performing with a first-semester GPA above the "native students" at our local university. Internal surveys show that faculty report increasing and significant efforts to incorporate the strategies into their curricula, and students report that faculty members are assigning the techniques.

Clearly, student learning is an essential component of our shared efforts in higher education. Taking the time to create a framework for evaluation, form a functional team to focus the analysis, and devise a solid assessment plan is worth the effort. As one of our student survey participants said of the active reading techniques, "It takes more time, but it works, Dude!" ■

Dianna Renz is director of planning and improvement at Western Wyoming Community College.

From the States

(continued from page 8)

right to feel pleased about the impact that "Assurance of Learning" is having on assessment activities among business programs in the United States.

So in very different ways, CAEP and AACSB International are shaping assessment practice among the programs they accredit. When data availability improves, as CAEP confidently expects it will, teacher education programs will be subject to a powerful and creative set of performance measures that attempt to tap directly the professional practice of their graduates in the ultimate performance setting—the K–12 classroom. Meanwhile, AACSB International's "Assurance of Learning" is steadily stimulating the ability of business programs to improve continuously on the basis of evidence of effectiveness. Both are moving assessment practice forward for specialized accreditors and we should expect to see similar efforts on the part of others in the coming years. ■

Reference

Ewell, P., Paulson, K., and Kinzie, J. 2011. *Down and In: Assessment Practices at the Program Level*. Champaign, IL: National Institute for Learning Outcomes Assessment (NILOA).

Peter T. Ewell is vice president of the National Center for Higher Education Management Systems (NCHEMS).

Development of a National Survey for Secondary Mathematics Teacher Education Programs

(continued from page 6)

educators, prospective students, or school counselors to identify the most appropriate program for their research and/or academic interests.

Asking more detailed questions in our study about each mathematics and mathematics education course in the program, while providing important information, might have affected the survey participants' motivation to finish the survey. Even with our efforts to reduce respondents' cognitive load by only requesting the number of required courses and credits instead of additional details about each course, 31 respondents stopped taking the survey at the point at which the course information was requested. The processes we used to overcome those challenges, including receiving input from multiple scholars with diverse academic backgrounds, conducting pilot studies, and understanding the limitations of survey methodology, are worthwhile to share with other researchers and practitioners interested in conducting a program survey to help them envision issues that may occur while designing such surveys and find their remedies.

Because the original purpose of the *PTA* survey was to create a snapshot of the current status of SMTE programs, it was not intended to illustrate the full extent of the unique efforts made by each program to prepare preservice teachers to teach algebra. Of course, each program provides a variety of distinct opportunities that contribute to preservice teachers' preparation. The National Research Council (2010) highlights the need to capture both the breadth and depth of teacher education programs through the combined use of both qualitative and quantitative data. To respond to this need, along with the national program survey described herein, the *PTA* study is now in the process of analyzing data from interviews with faculty and pre-service teachers from five programs with diverse characteristics to understand particular instances of SMTE programs in the United States.

Preparing to Teach Algebra is a three-year Collaborative Project at Michigan State University and Purdue University, funded through the National Science Foundation's Research and Evaluation on Education in Science and Engineering (REESE) program (MSU 1109256; Purdue 1109239). Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the National Science Foundation. ■

References

Carnegie Foundation for the Advancement of Teaching. n.d. Classification Description. Retrieved from http://classifications.carnegiefoundation.org/descriptions/basic.php

Conference Board of the Mathematical Sciences. 2012. *The Mathematical Education of Teachers II*. Providence, RI and Washington, DC: American Mathematical Society and Mathematical Association of America.

Dillman, D. A., Smyth, J. D., and Christian, L. M. 2009. *Internet, Mail and Mixed-Mode Surveys: The Tailored Design Method*, 3rd ed. Hoboken, NJ: John Wiley and Sons.

Husén, T., editor. 1967. *International Study of Achievement in Mathematics: A Comparison of Twelve Countries*. New York, NY: John Wiley and Sons.

National Governor's Association Center for Best Practices & Council of Chief State School Officers. 2010. *Common Core State Standards for Mathematics*. Washington, DC: Authors.

National Research Council. 2010. *Preparing Teachers: Building Evidence for*

Sound Policy. Committee on the Study of Teacher Preparation Programs in the United States, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Newton, J., Maeda, Y., Alexander, V., and Senk, S. L. 2014. How well are secondary mathematics teacher education programs aligned with the recommendations made in MET II? *Notices of the American Mathematical Society* 61(3):292–295.

Stevens, F. I. 1993. *Opportunity to Learn: Issues of Equity for Poor and Minority Students*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

Yukiko Maeda is assistant professor of Educational Studies, Jill Newton is assistant professor of Curriculum and Instructions, and Vivian Alexander is a graduate student in Applied Measurement and Research Methodology at Purdue University. Sharon L. Senk is professor in the Program in Mathematics Education and Department of Mathematics at Michigan State University.

Healthy Assessment

(continued from page 4)

Grades provide the feedback = 26%

Only after the fact via exit, alumni, and employer surveys = 11%

Newsletter or student orientation = 11%

Commentary: Two primary routes convey corrective feedback to students. One route is the recognizable pathway involving classroom professors and clinical supervisors. Weekly conferences with groups of students are common. The other route is independent of the faculty and typically involves an electronic system that grades multiple-choice exams. Use of sophisticated item analysis and analytics provides individualized, corrective feedback that is often accompanied by tutorial material and practice exams—followed by more feedback. According to the respondents in these interviews, nursing students are expected to use the tutorials. These expectations and these uses of assessment differ from what has generally taken place in most other disciplines.

7. Regardless of Student Grades, How Do You Use Assessment To Assure Competence When Students Graduate and Leave?

Summation of multiple methods of assessment = 79%

Capstone or critical element or essential traits that must be shown as competent = 53%

Exit exam = 11%

There were three additional responses (paraphrased):

"We have a special day each week for discussing individual student competence" = 1

"We apply Bloom's taxonomy to analyze student competence" = 1

"There are licensed nurses in our program, and they have already been certified as competent" = 1

Commentary: One interview took this path: "Mastery can't be assured because the nursing discipline is a *practice* that evolves with circumstance and time. All we can do is do our best, and assessment helps us do that." Nevertheless, nursing programs cannot allow themselves to produce graduates who are only "82% accurate in their calculations of medical dosages." Accordingly, nursing faculties either triangulated several assessment methods, or they established "essential elements"—such as medical dosage calculations and high-intensity clinical simulations—that students must pass to perfection before they progress further in the program. Moreover, nursing graduates cannot practice even as novices until they pass the nursing board exam. Several respondents affirmed that assessment and its feedback helped nursing students achieve high pass rates on board exams, even if real mastery is, in principle, not achievable.

Discussion

At least seven generalizable lessons (labeled a–g below) became apparent in the nursing school interviews. Part of nursing schools' successes with assessment is that (a) they align assessment operations with what they want to accomplish, namely, assuring student learning at high levels . . . and improving. This is the primary lesson about assessment that is usable in other disciplines, because high expectations are not unique to nursing. Moreover, (b) nursing faculties have regular, intentional conversations about curriculum and pedagogy, which tend to make visible those curricular areas that are candidates for improvement. Thus, (c) curricular visibility allows assessments to align with priorities. Because roles and goals are identified, (d) nursing schools tend to focus on assessing what matters to them so that assessment processes become embedded in teaching and learning. Associated with the curricular and pedagogical efforts, (e) feedback exists for students (with opportunity to remediate), and (f) feedback exists for faculty members. The presence of these feedback loops appears critical to assessment effectiveness. Significantly, *students are expected to take advantage of assessment feedback as part of their own responsibility*. It quickly becomes apparent that those who take advantage do better than those who do not. Thus, assessment makes teaching efforts more effective by obliging students to *engage*—and thus, students learn more! One can debate the influence of the board exam, but even that influence is learning-related.

Despite their successes with assessment, nursing schools reported struggling with manipulating data by hand. (g) Respondents who reported using turn-key

electronic services such as ExamSoft, the Assessment Technologies Institute (ATI), and Health Education Systems, Inc. (HESI) also reported that their students received targeted feedback that improved their learning. Simultaneously, these services saved faculty members time and effort, all while improving teaching efficiency and effectiveness.

Nursing faculty members are exceptionally busy, probably more than other academics, because they attend not only to their students and their scholarly duties but also to their patients. They don't have time to wallow in trivial administrative tasks. Thus software systems that assemble data for convenient feedback to faculty, while giving targeted feedback to students *that they themselves can access without occupying faculty time*, are highly valued by the institutions that use them. ■

Recommended Reading

Assessment Technologies Institute (ATI). n.d. *Nursing Education*. https://www.atitesting.com. Accessed January 5, 2014.

Banta, T. W. 1993. *Making a Difference: Outcomes of a Decade of Assessment in Higher Education*. San Francisco, CA: Jossey-Bass.

ExamSoft. n.d. *Program Focus*. http://learn.examsoft.com/exam-programs. Accessed January 5, 2014.

Health Education Systems, Inc. (HESI). n.d. http://www.elsevieradvantage.com/article.jsp?pageid=10748. Accessed January 6, 2014.

McDonald, M. E. 2007. *The Nurse Educator's Guide to Assessing Learning Outcomes*, 2nd ed. Sudbury, MA: Jones and Bartlett.

Oermann, M. E., and Gaberson, K. B. 2013. *Evaluation and Testing in Nursing Education*, 4th ed. New York, NY: Springer.

Weiner, W. F. 2009. *Establishing a Culture of Assessment*. Academe (AAUP). http://www.aaup.org/article/establishing-culture-assessment#.UTlTet_NLnZa5. Accessed January 17, 2014.

Douglas Eder is emeritus at Southern Illinois University Edwardsville and also served in higher administration positions at Arizona State University, the University of North Florida, and most recently as interim "assessment coach" at Purdue University.

Do Good Assessment Practices Measure Up to the Principles of Assessment?

(continued from page 2)

starting with outcomes and building an aligned curriculum and program, with embedded measures to achieve those outcomes from course assignments through to program completion, faculty better understood and were positioned to use assessment results to make changes throughout the curriculum to enhance learning. Assessment work that lasts, and becomes meaningful and useful to institutions, is work that is rooted in valued educational outcomes.

Table 1. AAHE Principles of Good Practice for Assessing Student Learning (AAHE 1992)

1. The assessment of student learning begins with educational values.
2. Assessment is most effective when it reflects an understanding of learning as multidimensional, integrated, and revealed in performance over time.
3. Assessment works best when the programs it seeks to improve have clear, explicitly stated purposes.
4. Assessment requires attention to outcomes but also and equally to the experiences that lead to those outcomes.
5. Assessment works best when it is ongoing not episodic.
6. Assessment fosters wider improvement when representatives from across the educational community are involved.
7. Assessment makes a difference when it begins with issues of use and illuminates questions that people really care about.
8. Assessment is most likely to lead to improvement when it is part of a larger set of conditions that promote change.
9. Through assessment, educators meet responsibilities to students and to the public.

The case studies also exemplify principle 2, which specifies that assessment should employ a diverse array of methods that aim for a more complete, integrated picture of learning, and therefore firmer bases for improving students' educational experiences. Carnegie Mellon University (CMU), for example, developed strong departmental assessment featuring diverse methods. Faculty were supported in the development of varied approaches, including embedded assessments, assignment rubrics, and performance measures, through the leadership of CMU's Eberly Center for Teaching Excellence. Assessment at CMU begins with faculty questions about learning and effective teaching and is informed by departmental curricular interests, program goals, and the particular discipline.

Assessment approaches that employed an array of methods also emphasized the inclusion of process measures to help understand which students learn best under what conditions (principle 4). The consideration of process and outcome measures that were linked to courses and

programs contributed to holistic improvements in the student experience. However, although all the case-study institutions were employing multiple methods and some included unique performance and process measures, they all indicated a hope for richer methods that provided opportunities for students to demonstrate their best work and to use assessment approaches particular to their institution and discipline that would inform program and curricular improvements.

As Hutchings et al. (n.d.) noted, the emphasis on assessment for improvement that is ongoing rather than episodic (principle 5) and involves multiple stakeholders (principle 6), has taken on greater emphasis in assessment than when the principles were first developed. These were also strong themes across the nine case studies. Advances have clearly been made to involve more faculty and staff in assessment activities. For instance, Augustana College shares assessment information at faculty retreats and through an on-line, easily accessible website. Yet even the institutions with advanced assessment continue to strive for more involvement, as does St. Olaf College, where the faculty governance structure includes an assessment committee to encourage more widespread involvement.

Principle 7, which specifies that assessment should focus on questions that people care about, was a prominent theme across the case studies. When assessment practices matched what faculty and staff cared about, institutions reported greater investment in assessment processes and results. This principle is particularly evident in the Colorado State University (CSU) case study, which depicts CSU's use of PRISM, an on-line planning infrastructure to support continuous improvement efforts by providing the university community and external constituents access to assessment data. PRISM allowed for entry points customized by audience to highlight assessment results of specific interest to groups including alumni, faculty, students, families, and employers. This finding about the relationship between what is meaningful to people and institutional investment in assessment is particularly salient given concerns about the proliferation of assessment data and reports, but little use (Blaich and Wise, 2011). Ensuring that assessment questions and results relate to matters people care about has value and deserves promotion.

The importance of principle 8, which asserts that assessment is most likely to lead to improvement when it is part of a larger set of conditions that promote change, was also a strong theme. Texas A&M International University's (TAMIU) involvement with the Building Engagement and Attainment for Minority Students (BEAMS) project provided the campus with data and the opportunity to develop an action plan for the improvement of student engagement, learning, and success. This project provided crucial information to address concerns about student writing, and for the campus to develop this further as part of its *Write-On, TAMIU!*, the institution's Quality Enhancement Plan (QEP) for reaffirmation of accreditation. Assessment alone changes little. The case studies illustrate that assessment's greatest contribution comes when the quality of teaching and learning is visibly valued and examined with the use of a variety of efforts to foster improvement.

Conclusion: Relevance of the Principles

Just as Hutchings et al. (n.d.) concluded that some of the principles have even greater resonance in our current context, this analysis points out that in relation to the case studies, principles 1–4, 7, and 8 are essential to the framing and the doing of assessment, and are more or less embedded in current practice. However, the mixed findings for principles 5 and 6 suggest that these aspects still demand attention. In addition, the broader moral responsibility stated in principle 9, although not expressed as a reason for doing assessment, is the rationale for efforts to increase transparency and communication of assessment results that is gaining prominence in the field and occurring at the case study sites.

In general, practices across the case study sites appear to be consistent with the AAHE principles. However, our exploration also exposed principles that were less often observed. In particular, although good practice statements emphasize multiple measures to recognize the multidimensionality of learning outcomes, it remains challenging to employ all the measures necessary to capture the full range of learning (i.e., institution-level learning outcomes; and program-level, course, and cocurricular outcomes). This finding suggests that more needs to be done to support, validate, and demonstrate diversity in assessment methods to make this principle a reality. Ensuring that assessment work is ongoing and not episodic appears to be a persistent struggle, and is related to efforts to shift from assessment as an add-on to being embedded in good educational practice. Engaging more stakeholders was another less-observed practice, with institutions reporting that they were still in the preliminary stages of fully engaging a variety of stakeholders in assessment.

Ensuring that assessment work is ongoing and not episodic appears to be a persistent struggle, and is related to efforts to shift from assessment as an add-on to being embedded in good educational practice.

Each case study site was committed to continuing to improve its assessment efforts through ongoing fine tuning and making results more useful for planning and decision making. These case studies demonstrate that some institutions can simultaneously answer the call for information on student learning and take action to improve by committing to

JOSSEY-BASS™
A Wiley Brand
One Montgomery Street
San Francisco, CA 94104-4594
www.josseybass.com

assessment and sharing results. Moreover, the findings from these cases affirm that assessment is most likely to last and make a meaningful difference to educational effectiveness and improvement when it illuminates questions that people really care about. However, the challenge of using information for improvement persists. The problem, as assessment principle 7 suggests, is not that data are not good enough, or that we do not have enough data, but that our questions are not necessarily the right ones. Consequentiality is critical to using information both in terms of considering the evidence but also in using this to make changes to teaching and learning. Considering good assessment practice against the principles of assessment illustrates the strengths of current practice, provides empirical evidence of the principles, and illuminates areas that still require attention. ■

References

American Association for Higher Education (AAHE). 1992. *Nine Principles of Good Practice for Assessing Student Learning*. North Kansas City, MO: AAHE. www.learningoutcomesassessment.org/PrinciplesofAssessment.html#AAHE

Baker, G. R., Jankowski, N., Provezis, S., and Kinzie, J. 2012. *Using Assessment Results: Promising Practices of Institutions That Do It Well*. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

Banta, T. W., Jones, E. A., and Black, K. E. 2009. *Designing Effective Assessment: Principles and Profiles of Good Practice*. San Francisco, CA: Jossey-Bass.

Blaich, C. F., and Wise, K. S. 2011. *From Gathering to Using Assessment Results: Lessons from the Wabash National Study*. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Hutchings, P., Ewell, P., and Banta, T. n.d. *AAHE Principles of Good Practice: Aging Nicely*. www.learningoutcomeassessment.org/PrinciplesofAssessment.html

New Leadership Alliance for Student Learning and Accountability. 2012. *Committing to quality: Guidelines for Assessment and Accountability in Higher Education*. Washington, DC: New Leadership Alliance for Student Learning and Accountability. http://www.chea.org/alliance_publications/committing%20to%20quality-3rd%20edition.pdf

Western Association for Schools and Colleges [WASC], Senior College and University Commission. n.d. *Principles for Effective Assessment of Student Achievement*. 2013. Alameda, CA: WASC. www.wascsenior.org/files/principles_for_effective_assessment_of_student_achievement.pdf

Jillian Kinzie is associate director at the Indiana University Center for Postsecondary Research and NILOA senior scholar; Natasha Jankowski is assistant director of NILOA at the University of Illinois; and Staci Provezis, formerly with NILOA, is now senior communication and evaluation coordinator at the University of Illinois Urbana-Champaign.