

Assessment UPdate

May–June 2009
Volume 21, Number 3



Progress, Trends, and Practices in Higher Education

Using the Academic Skills Inventory to Assess the Biology Major

Kyle Seifert, Carol A. Hurney, Clifton J. Wigtil,
Donna L. Sundre

THE ACADEMIC SKILLS INVENTORY (KRUGER AND ZECHMEISTER, 2001) WAS developed at Loyola University of Chicago and originally designed for use with psychology majors. It was later extended for use in a variety of academic programs. As Donna Sundre (2003) reported earlier in *Assessment Update*, the Academic Skills Inventory (ASI) assesses student self-reports of behaviors in ten skill areas: (1) written and oral communication; (2) information gathering; (3) groups, organization, and community; (4) interpersonal skills, counseling, interviewing, and mentoring; (5) behavior management, supervision, and teaching; (6) individual differences, special populations, and cultural diversity; (7) critical thinking and problem solving; (8) research methodology and statistics; (9) ethics and values; and (10) technology and computers. Sundre (2003) recommended broader use of the instrument, and here we review its use by the Biology Department at James Madison University (JMU) to explore the impact of curricular changes on student perceptions of their academic skill development. Not all of the ten skill areas would be expected to show differential impact for biology majors, but we certainly hoped to observe change in designated areas associated with science education.

James Madison University is a regional comprehensive university of about 18,000 students in Virginia. The Biology Department at JMU currently has 44 full-time faculty members and 10 part-time instructors, who serve over 700 undergraduate majors. Prior to 2002, biology majors were required to take five courses: Botany, Zoology, Cell Biology, Genetics, and Ecology. Three of these courses contained labs, and the faculty noted that many of the lab exercises involved observations of preserved specimens. Further, there was a developing consensus that a major overarching emphasis in classes was learning a substantial number of biological facts, with little or no emphasis on quantitative thinking or the process of doing science. We decided that the curriculum needed to be modified dramatically.

In 2002, the Biology Department, funded by a National Science Foundation Course, Curriculum, and Laboratory Improvement grant, implemented four new first-year and second-year laboratory courses: (1) Organisms, (2) Ecology and Evolution, (3) Cell and Molecular Biology, and (4) Genetics and Development.

CONTENTS

ARTICLES

Using the Academic Skills Inventory 1
to Assess the Biology Major
*Kyle Seifert, Carol A. Hurney,
Clifton J. Wigtil, Donna L. Sundre*

Expanding Professional Development 3
Opportunities to Enhance the
Assessment Process
Elizabeth A. Jones

Sustaining Assessment: A Post-
Epidemiological Approach Using 5
The Program Evaluation Standards
Rick D. Axelson, Arend Flick

Clarification of Assessment and 8
Evaluation
Gloria K. Goldman, Lori E. Zakel

Horizontal Assessment 10
W. James Bradley

COLUMN

From the States 12
Peter T. Ewell

FEATURE

Resources 15
Susan Kahn



Published online in Wiley InterScience
(www.interscience.wiley.com)
DOI 10.1002/au.213
© 2009 Wiley Periodicals, Inc.

Assessment Update

Progress, Trends, and Practices in Higher Education

May–June 2009

Volume 21, Number 3

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, *educational consultant*

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 Erwin, *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*

Assessment Update: Progress, Trends, and Practices in Higher Education (Print ISSN 1041-6099; online ISSN 1536-0725 at Wiley Interscience, www.interscience.wiley.com) is published bimonthly by Wiley Subscription Services, Inc., A Wiley Company, at Jossey-Bass, 989 Market St., San Francisco, CA 94103-1741. Periodicals Postage Paid at San Francisco, CA, and additional mailing offices. Individual subscriptions are \$127 per year (institutional \$175). 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, 989 Market St., San Francisco, CA 94103-1741.

Copyright © 2009 Wiley Periodicals, Inc., A Wiley Company. All rights reserved. Reproduction or translation of any part of this work beyond that permitted by Section 107 or 108 of the 1976 United States Copyright Act without permission of the copyright owner is unlawful. Requests for permission or further information should be addressed to the Permissions Department, c/o John Wiley & Sons, Inc., 111 River St., Hoboken, NJ 07030; (201) 748-8789, fax (201) 748-6326, www.wiley.com/go/permissions.

These new courses put greater emphasis on applied laboratory skills and experiences than on memorization of facts. The lab experiences were designed to engage students in the process of scientific inquiry. In addition, these courses were sequenced and integrated, with three lab themes (plants, animals, and microbes) to be investigated in each of the four new courses (Monroe and Hurney, 2002). The Biology Department annually collects assessment data to determine the extent to which we are achieving our stated student learning objectives. The implementation of the new curriculum represented an ideal occasion to explore the impact of an academic change on student learning.

The ASI is one of several assessment tools that our faculty administer each academic year to biology majors who are sophomores, juniors, and graduating seniors. An advantage of using the ASI is that students are not asked to report their perceived level of skill development. Rather, they are asked whether they have completed certain tasks or engaged in specified behaviors related to various academic skills (Kruger and Zechmeister, 2001; Sundre, 2003). Because the ASI was designed to assess specific academic skills and experiences, it could be used to determine whether the change in curriculum enhanced or provided more opportunities for the experiences that our faculty members value most, especially those related to experiencing the process of science. ASI data were collected from all graduating biology majors from 2004 to 2008. This time frame was deemed optimal because it would include three potential waves of students: (1) those who graduated prior to the implementation of the new curriculum (2004); (2) those who graduated during initial implementation of the new curriculum (2005, 2006, and, possibly, 2007); and (3) those who were likely to have completed the new curriculum (most of 2007 and 2008).

(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 would be 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). Short news items and content for the Book Review section may be 100–500 words in length. Annotations of recent publications for the Resources feature should be about 50–100 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. Future deadlines for submitting articles are August 1 (November–December 2009 issue), October 1 (January–February 2010 issue), and December 1 (March–April 2010 issue).

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

Expanding Professional Development Opportunities to Enhance the Assessment Process

Elizabeth A. Jones

HAVE BEEN WORKING WITH COLLEAGUES Trudy Banta and Karen Black on a new Jossey-Bass book, *Designing Effective Assessment: Principles and Profiles of Good Practice*, which focuses on assessment profiles from a variety of colleges and universities (Banta, Jones, and Black, forthcoming). These profiles also represent the spectrum of institutions from research universities to community colleges and public as well as private institutions.

We asked each of the authors of profiles to complete a common template that gave us insights about the purpose of their assessments, the methods used, required resources, the findings and impact of the findings, and success factors. We selected forty-nine of the best profiles to include in their entirety. Another 130 examples of

professional development practices that are portrayed in profiles in our new book. Faculty and staff at several institutions identified professional development as a major resource that was necessary in order to implement assessment and sustain it over time.

Many administrators have created centers for teaching and learning that offer special workshops or programs to help faculty become better instructors. Research suggests that single-session workshops or interventions have little effect on behavior (Licklider, Schnelker, and Fulton, 1997). Sustained, structured professional development opportunities are more likely to help faculty understand how to assess student learning.

In beginning to create development opportunities for instructors, it is difficult to know

Faculty learning communities are an example of a sustained initiative that could have a great impact on instructors. According to Milton Cox (2004), learning communities typically help faculty become more actively engaged in collaborative work that extends for up to one year as they strive to enhance teaching and learning. At Texas Christian University, several campus units provided funding to support the creation of six faculty learning communities (FLCs), each representing a part of the core (general education) curriculum (Catherine Wehlburg, personal communication, April 1, 2008). Across these FLCs, faculty discussions centered on how to design and sustain assessments for each part of the core curriculum, how to disseminate the results from these assessments to other faculty, and how to enhance teaching.

At Ferris State University, professional development was the main initiative for helping faculty adopt a learner-centered agenda (Roxanne Cullen and Michael Harris, personal communication, April 22, 2008). Directors of two teaching and learning centers worked with faculty to develop a rubric to gauge the degree to which individual courses and syllabi demonstrated characteristics associated with learner-centeredness. One group of faculty did not make explicit connections between their learning outcomes and assessment activities. Professional staff from the Center for Teaching and Learning worked with these faculty to educate them about how to make these linkages more transparent.

Ideally, faculty should be provided with professional development opportunities

Faculty learning communities are an example of a sustained initiative that could have a great impact on instructors.



good practice, along with contact information, are listed in a resources section.

In 2007, Mary Deane Sorcinelli surveyed a group of faculty and administrators at a variety of institutions in order to identify the major challenges associated with professional development. The most important issue reported was that faculty needed to create student-centered learning experiences. Another of the top challenges was how to assess student learning in an effective and meaningful manner (Sorcinelli, 2007).

The intent of this article is to provide some examples of structured, ongoing

what types of help they need to be effective in their assessment efforts. At Widener University, a special task force was appointed in 2005 to develop and conduct a survey of faculty needs. The results indicated that instructors wanted to learn different strategies for assessing student learning and how to analyze the information and use the results to strengthen teaching and learning (Widener University, 2006, p. 124). Needs assessments like the one used at Widener can help academic leaders identify which assessment topics need attention and suggest how to offer educational opportunities for faculty.

during the entire assessment cycle—from the very beginning as plans are developed, to the implementation and interpretation of results, to understanding how to use the results to identify warranted changes. Faculty members at Northern Iowa University conceived a professional development plan that addressed the entire assessment process (Barry Wilson, personal communication, March 31, 2008). They initially offered curriculum mapping workshops during the 2006–07 and 2007–08 academic years so that they could identify which courses were designed to address certain program outcomes. The second series of workshops at Northern Iowa addressed how to assess student learning outcomes at the course level. In the final wave of professional development, an entire day was devoted to the study and interpretation of data and the development of action plans for change. The provost canceled classes for this professional development day so that faculty and administrators could spend the time necessary to review assessment results as well as curriculum mapping data. These

discussions led to the creation of action plans for change.

Formal training is required for the development and implementation of rubrics to assess student work.



discussions led to the creation of action plans for change.

Formal training is also required for the development and implementation of rubrics to assess student work. Typically, instructors collaborate as they develop specific rubrics to measure student learning. After they apply the rubrics to samples of student work, they examine how closely their ratings match in order to determine the degree of consistency across multiple raters. If there are large gaps in the ratings, faculty members discuss their reasons for their judgments and seek to reach a better consensus or revise the rubric so that it is more specific and clear. Senior MBA faculty at Johnson and Wales University mentor adjunct instructors, helping them understand how to teach courses and measure student learning in alignment with program learn-

ing outcomes. Workshops are held to assist faculty in designing and using rubrics, with a focus on teaching instructors to consistently apply the criteria so that there is greater inter-rater reliability. In addition, faculty have created portfolios to document their assignments and rubrics. They use this information to identify strengths and professional development needs.

In a similar manner, faculty members at St. Mary's College of Maryland provided extensive training for their colleagues who taught pilot first-year seminars so that they could learn strategies for promoting and assessing information literacy, critical thinking, and oral and written communication skills. The following year, an additional group of instructors received ongoing training, and some of the workshops were led by the faculty members who had taught the initial first-year seminars.

The development and revision of rubrics is crucial to assessing important general education outcomes. Teresa Flateby at the University of Florida discussed using

a special rubric designed to assess writing and thinking, while Patti Clayton and her colleagues discussed adopting rubrics to assess critical thinking and reasoning skills in service-learning courses at North Carolina State University.

While most of the profiles that address professional development in our new book are focused on academic affairs, it is important to provide similar programming for student affairs leaders and staff. Such programming can be designed as an integrated set of learning experiences that take place over several semesters. Aloi, Green, and Jones (2007) discuss a series of six professional development seminars that was offered to all student affairs leaders and staff at West Virginia University to help them develop learner-centered assessment plans for their units.

Providing systematic, formal professional development opportunities is vital in order to adequately prepare and motivate faculty and student affairs staff to engage in meaningful assessment practices. If leaders thoughtfully plan and develop a series of ongoing professional development experiences, participating instructors and staff can learn how to conceptualize new ideas and receive constructive feedback from their peers about suggested improvements. Such feedback can help faculty become better instructors and more knowledgeable about the quality of student learning. It can also help student affairs staff understand the contributions that their programs and services make to student development. ■

References

- Aloi, S. L., Green, A. S., and Jones, E. A. "Creating a Culture of Assessment Within West Virginia University's Student Affairs Division." *Assessment Update*, 2007, 19(2), 7–9.
- Banta, T. W., Black, K. E., and Jones, E. A. *Designing Effective Assessment: Principles and Profiles of Good Practice*. San Francisco: Jossey-Bass, forthcoming.
- Cox, M. D. "Introduction to Faculty Learning Communities." In M. D. Cox and L. Richlin (eds.), *Building Faculty Learning Communities*. New Directions for Teaching and Learning, no. 97. San Francisco: Jossey-Bass, 2004.
- Licklider, B. L., Schnelker, D. L., and Fulton, C. "Revisioning Faculty Development for Changing Times: The Foundation and Framework." *Journal of Staff, Program, and Organizational Development*, 1997, 15(3), 121–133.
- Sorcinelli, M. D. "Faculty Development: The Challenge Going Forward." *Peer Review*, 2007, 9(4), 4–8.
- Widener University. *Self-Study Report for Re-Accreditation by the Middle States Commission on Higher Education*. Chester, PA: Widener University, 2006.

Elizabeth A. Jones is professor of higher education leadership in the College of Human Resources and Education at West Virginia University in Morgantown, West Virginia.

Sustaining Assessment: A Post-Epidemiological Approach Using *The Program Evaluation Standards*

Rick D. Axelson, Arend Flick

AT A RECENT ASSESSMENT WORKSHOP, one of our former colleagues, Andrew Stuart Bergerson, offered a novel metaphor for building our university's assessment program. "We need to spread the assessment virus on campus," Drew remarked, half seriously. Although we had a good laugh over this image at the time, it encapsulates the faculty-driven approach many of us seek to develop when launching new assessment initiatives. We attempt to infect prominent faculty with the assessment virus and support these carriers as they spread it campuswide. But while such an epidemiological approach may be effective at creating initial enthusiasm, it is often dif-

icult to sustain when the early adopters' interest in assessment wanes in the face of their colleagues' immunity to the virus.

Educational Evaluation (1994). Although these standards were developed to guide evaluation practice related to education and training programs, projects, and materials, they also identify the essential issues that institutions need to address in developing sustainable assessment systems.

Why aren't the standards already part of the assessment practitioner's tool kit? This deficit is probably attributable to the disciplinary wall erected between assessment and evaluation in the 1980s. As Banta (2002) observes, concern that anxious faculty would conflate program evaluation activities with the personnel evaluation processes related to promotion and tenure

cational Evaluation (JCSEE) wrestled with this issue shortly after its inception and concluded that personnel evaluation and program evaluation should be addressed separately as distinct types of evaluation. As its first task, the JCSEE chose to address program evaluation issues and published a set of evaluation standards in 1981. The standards were developed through a rigorous and systematic public process involving many evaluation experts, users of evaluation, and others concerned about the quality of education.

The second edition of *The Program Evaluation Standards*, released in 1994, provides thirty standards, organized along four dimensions: utility, feasibility, propriety, and accuracy. (The current standards, new draft standards, and more background information are available at <<http://www.wmich.edu/evalctr/jc/>>.) Applying these standards to assessment, we believe, could help structure a campus dialogue about sustainable assessment programs. Besides sensitizing discussants to the breadth of the issues requiring resolution, the standards are helpful in framing our responses to assessment discontents, who often misperceive bad assessment practices as problems inherent in all forms of assessment. Thus, to return to our original metaphor, campus dialogue informed by the standards can help us design, not merely hope for, the kinds of institutional conditions that would make assessment chronic—and, in contrast to almost all other organizational

Besides sensitizing discussants to the breadth of the issues requiring resolution, the standards are helpful in framing our responses to assessment discontents.



How can we translate the momentum from early-adopting faculty leaders into a sustainable assessment program for our institution? What organizational processes and conditions are needed to support these efforts? How should we initiate and structure campus dialogue on these matters? As we reflected on these questions, they seemed strikingly similar to those addressed in *The Program Evaluation Standards*, developed by the Joint Committee on Standards for

led to a new use of the term *assessment*. This term helped to allay fears because it signified a faculty-driven brand of program evaluation that focused on the improvement of student learning rather than the evaluation of individual instructors.

Evaluators have also struggled to differentiate personnel evaluation from program evaluation. (For a description of the evaluation profession and the full range of inquiries and activities undertaken by evaluators, see the American Evaluation Association's Web site at <<http://www.eval.org/>>.) The Joint Committee on Standards for Edu-

contagions, curative. How? Following are some examples.

One charge often leveled at assessment is that it's *unnecessary* or even *useless*. This charge is predicated on the assumption that higher education today has failed or that collecting evidence about student learning will magically improve it. Assessment discontents like to remind us that American education is the envy of the world, so why try to fix what isn't broken? They argue that a well-trained and conscientious professoriate should simply be left alone to do the job it has always done well. To the extent that improvement is possible, they suggest that our problems are mostly the fault of un-

would have much less to criticize if institutions consistently met these standards in developing their assessment processes. Simply ensuring that assessment activities arise from the needs and interests of faculty and lead to meaningful (and well-publicized) institutional change would go a long way toward answering their criticisms. We can also convince naysayers that assessment is useful by directing them to the mounting evidence that assessment has already been useful for many colleges and universities in addressing real problems in student learning.

Critics of assessment also frequently attack it for being *reductive*, for distorting or

possible sources of bias and error (A5, A6), and to report results with sufficient contextual information for informed interpretation (A6, A7, A8). It also reminds us that the chain of reasoning that leads from data and analysis to conclusions must be "clearly documented without omission or flaws" (A2). Thoughtful practitioners agree that assessment is a messy science, that learning is not always (easily) measurable, that standardized tests often sacrifice content and construct validity for (illusory) reliability, that treating knowledge as a simple commodity distorts it beyond recognition. Reductive assessment (and such forms are widespread, employed by those who would trade reliability and validity for speed and comfort) cannot be supported, nor should it be. Accurate and useful assessment is much more likely to be sustained.

Yet another charge frequently made against outcomes assessment, particularly against the expectation that faculty members work collaboratively and systematically to gather evidence about patterns of student learning in their courses and programs, is that this expectation is *unrealistic*. Instructors at four-year institutions already have heavy teaching and advising loads and are expected to engage in active scholarship. Instructors at two-year colleges are ordinarily not required to do research but have commensurately higher teaching loads. How can faculty be expected to find the time to work together on rubric development, artifact scoring, test item mapping to program or institutional outcomes, or analysis and interpretation of assessment data? The problem, according to assessment skeptics, is compounded by the lack of meaningful institutional support for assessment efforts. Assessment, they say, is yet another unfunded (or underfunded) mandate—perhaps a good idea, but impractical, given how busy faculty already are and how unable administrators seem to be to acknowledge and support the time required for faculty to do this work.

The feasibility standards offer suggestions for making assessment responsive, efficient, and effectively managed (F1, F3, F4), balancing the interests of all

Simply ensuring that assessment activities arise from the needs and interests of faculty and lead to meaningful (and well-publicized) institutional change would go a long way toward answering criticism.



derprepared or undermotivated students, a condition that assessment, with its focus on pedagogy, curriculum, and student support as sites of institutional change, can do nothing about. Some discontents concede that outcomes assessment could play a useful role in improving institutional effectiveness, but they are skeptical about the willingness or ability of institutions to act on assessment results. For all of these naysayers, assessment is neither necessary nor even particularly worthwhile.

The utility standards address these objections and suggest some good practices for sustaining assessment. Good assessment processes should not merely be a bureaucratic exercise but should address issues of interest to faculty (standards U1 and U2). Good assessment processes have clearly defined purposes and are informed by explicit values (standards U3 and U4). Good assessment processes have clearly defined products or outcomes that aid participants in achieving those purposes (standards U5 and U6). Good assessment processes result in products that are clearly communicated to all participants (standard U7) and that inform decision making at the institution (standard U8). Assessment discontents

oversimplifying the nature of higher education, thereby reducing all learning to only that which can readily be measured quantitatively. Discontents who speak darkly of assessment as an imposition of corporate culture upon academia in which students are seen as customers and education as a commodity belong in this category of complaint. Such critics often speak wistfully, as Laurie Fendrich (2007) does, of their own college experiences with idiosyncratic professors whose indirect and enigmatic lectures contained ideas that "took years to really settle in." How, she wonders, can outcomes assessment ever capture that kind of influence? Such assessment discontents worry that assessment will undermine the richness of the educational process itself, as it seems to have done under No Child Left Behind (NCLB) in the K–12 system, forcing teachers to teach mostly to outcomes measured by standardized tests. Assessment, in these critics' view, is quackery masquerading as science.

The accuracy standards speak to these objections by encouraging us to develop assessment processes that are both valid and reliable (A1, A3, A4, A5), to collect data with sufficient rigor and attention to

the individuals and groups involved (F2). Meeting these standards would require that assessment processes be designed to produce information of sufficient value that the resources expended on them can be justified. In practical terms, this might mean that institutions focus (at least initially) on modest assessment projects, frugal ones that are more likely to show immediate results. An electronic portfolio system for assessing general education outcomes for all students might be a legitimate long-term approach, but gathering and scoring sample paper artifacts in a specific program or two might pay off in usable data more quickly. Overly complex reporting matrices with multiple levels of cells to fill in or detailed questions to answer might be simplified. Feasibility is also likely to be enhanced and assessment more likely to be sustained when institutions commit to a scholarship of teaching that conceives of assessment as a vital form of faculty research that is acknowledged appropriately in promotion and tenure deci-

In practical terms, meeting the feasibility standards might mean that institutions focus (at least initially) on modest assessment projects, frugal ones that are more likely to show immediate results.



sions. Finally, administrators can do more to encourage faculty not to see themselves as working in isolation from each other and from the institution as a whole. Faculty need to be made to feel that simply cultivating their own garden is bad husbandry because that garden is dependent on so many others for its yield and productivity.

Perhaps the most repeated criticism of outcomes assessment is that it is *intrusive*. Assessment is said to subvert academic freedom and introduce a layer of oversight that has a chilling effect on teaching and learning. Assessment discontents often express the fear that assessment processes will be used to punish teachers or even discontinue whole programs that are determined to be underperforming, although critics typically fail to identify evidence that this has ever happened. NCLB and the recommendations of the Spellings Commission

on the Future of Higher Education are seen as indications that a higher education dystopia looms, one in which no teacher is ever free from observation and measurement. Such concern about how assessment results might be misused, sometimes bordering on paranoia, also hinders efforts to sustain assessment.

The propriety standards suggest that good assessment should be legal and ethical, conducted “with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.” Agreements need to be developed to outline the legitimate uses of assessment data and practices with respect to instructors and departments (P2, P3, P4, P5, P6). The issue of who owns assessment data—and what the appropriate uses of such data are—needs to be addressed in consultation with all stakeholders at the institution (P1, P2). It’s easy enough to develop a campus policy that prohibits the use of assessment data for the evaluation of individual faculty or the discontinu-

ation of any program. After all, unless such policies exist, the greatest danger is that only the most innocuous of data will ever be generated. Faculty need to be assured that data will be aggregated and sufficiently protected from illicit use to make targeting individuals impossible.

Because much of the early debate in the assessment literature centered on utility and accuracy issues, matters related to propriety and feasibility tend to be overlooked. For example, the article “Principles of Good Practice for Assessing Student Learning” (Hutchings, 1993) provides insightful guidelines for conducting useful and accurate assessment but does not explicitly address propriety or feasibility issues. Consequently, *The Program Evaluation Standards* can serve a valuable function in expanding campus dialogue to address nuts-and-bolts implementation is-

ues—issues that, if unattended, can torpedo even the most useful and accurate assessment programs.

Framing our campus assessment dialogue in terms of *The Program Evaluation Standards* would enable us to have more productive, comprehensive, inclusive, and issue-focused dialogue about what we can accomplish with a high-quality assessment effort on campus. Such a framework would help us move past the distracting critiques of bad assessment as reasons for not doing assessment at all. In addition to facilitating our internal discussions on campus, the standards provide a well-respected framework for communicating with external audiences, helping us articulate the shortcomings of heavy-handed efforts to impose false forms of accountability on higher education. In sum, adherence to professional evaluation standards can help us design high-quality assessment systems and, at the same time, communicate well-reasoned explanations of our approaches to critics within and outside higher education. ■

References

- Banta, T. W. “A Call for Transformation.” In T. W. Banta (ed.), *Building a Scholarship of Assessment*. San Francisco: Jossey-Bass, 2002.
- Fendrich, L. “A Pedagogical Straitjacket.” *Chronicle of Higher Education*, June 8, 2007, p. B6.
- Hutchings, P. “Principles of Good Practice for Assessing Student Learning.” *Assessment Update*, 1993, 5(1), 6–7.
- Joint Committee on Standards for Educational Evaluation. (James R. Sanders, chair.) *The Program Evaluation Standards: How to Assess Evaluations of Educational Programs*. (2nd ed.) Thousand Oaks, Calif.: Sage, 1994.

Rick D. Axelson is assistant professor in the University of Iowa’s Carver College of Medicine in Iowa City. Arend Flick is associate professor of English at Riverside Community College in Norco, California.

Clarification of Assessment and Evaluation

Gloria K. Goldman, Lori E. Zakel

ASSESSMENT PRACTICES AT SINCLAIR Community College have evolved since the early 1980s to a highly sophisticated set of processes that begin with the collection, analysis, and documentation of general education and program outcomes data and result in course and program improvements. A continuing challenge has been the development of new full-time and adjunct faculty with little or no assessment experience so that they become full participants in these critical college processes. A persistent

question from new faculty has been “What is the difference between assessment and evaluation?” The confusion is not easily resolved by referring new faculty to the literature, for there are still many different definitions of both terms.

Even if the definitions were aligned, confusion might not be eradicated. Language is a funny thing. We all know what *we* mean when we are trying to convey something to others, so why don’t they always understand us? Why do messages sometimes get confused?

Some very basic examples can shed light on the difficulty people have when communicating. Take the simple word *next*. What does it mean when we use the word *next* in a sentence? If we are on a highway, and I tell you, “Take the next exit,” that’s pretty clear, isn’t it? Most people would take the first exit that appears after hearing that directive. Now I say to you, “I’ll pick you up next Saturday.” Most people will be ready not on the first Saturday that occurs after today but on the Saturday after

Table 1. Examples of Assessment and Evaluation from Sinclair Community College

Assessment	Evaluation
<p>A faculty member provides feedback to a student regarding performance on an examination. The student uses that feedback to study differently in order to improve learning and performance.</p> <p>A team of faculty members analyzes examination results of all students in a course and discovers that 65% of the students did not demonstrate understanding of an important concept. Faculty members investigate possible causes and plan changes in teaching/learning strategies to improve student understanding.</p> <p>A student delivers an oral presentation in class. The faculty member provides a critique of delivery and content so that improvements may be made in the student’s subsequent presentations.</p> <p>A faculty member analyzes the results of oral communication checklists completed for all students in the course section who delivered oral presentations in class, in order to determine opportunities for improving teaching and learning.</p> <p>The class attendance record indicates that a student has been absent multiple times. The faculty member advises the student in order to facilitate improved attendance, as studies suggest that regular class attendance contributes to student success.</p>	<p>A faculty member corrects an examination and assigns a grade of 82% to a student.</p> <p>Pop quizzes are given in a class to determine if students have read sections of the text that cover important concepts. Simple Pass/Fail grades are assigned and tallied at the end of the quarter. The quizzes count for 5% of the total course grade.</p> <p>A student delivers an oral presentation in class. The faculty member provides a critique of delivery and content accompanied by a grade for the assignment.</p> <p>A health professions faculty member uses a rating scale to assign numbers (1–4) that indicate the level of achievement of clinical criteria based on observation of a student’s performance of patient care.</p> <p>Points are deducted from a student’s grade for each class absence, in accordance with a department policy.</p>

the nearest one. Confusing? Well, our individual definition for any given word depends on the context (for example, where we are when we say it, to whom we are talking, how much history we have with that person). But that's not the only problem that can arise from our choice of words.

Words can also cause problems because some arouse strong feelings when we hear them. When a faculty member is told that he or she is to be evaluated or needs to engage in assessment, defensiveness is often a natural reaction. Clarifying terminology can help to reduce the defensiveness that the words *evaluation* and *assessment* can generate.

It clearly benefits students, faculty, and the assessment process to ensure the existence of clear definitions, definitions that allow everyone in the institution to understand the sometimes complex and often contradictory terminology pertaining to assessment. Because a universal set of definitions for assessment terms in higher education does not exist, individual institutions must agree on terms and ensure that their own terminology is widely available and utilized.

Members of the Sinclair Community College Assessment Committee have

developed a set of definitions, complete with examples, to help put all faculty on the same page. Although this may seem fundamental, we have found that we can move quickly to assisting faculty with data collection and analysis and to improved practice once the confusion over terminology has been resolved. This ultimately results in better understanding and implementation of assessment practices at Sinclair.

The following terms and examples were published in Sinclair Community College's quarterly newsletter, *Building the Collective Responsibility for Student Learning*, and are posted on the assessment Web site at <<http://www.sinclair.edu/about/learning/assessment/>>.

Evaluation is the analysis and use of data by faculty to make judgments about student performance. Evaluation includes the determination of a grade or a decision regarding pass/fail for an individual assignment or for a course.

Assessment is the analysis and use of data by students, faculty, and/or departments to make decisions about improvements in teaching and learning. This description is congruent with the College's definition: "Assessment at Sinclair is the shared process of purposeful, systematic meas-

urement used to document, reflect upon, and improve student learning."

Examples also are included to clarify the college's distinctions between the terms (see Table 1).

An added benefit of offering these examples is the opportunity they provide for faculty to understand that although the end results of assessment and evaluation differ, the instrument (or instruments) used for each may actually be the same. For example, a test might determine part of a student's course grade (evaluation) and the results of the same test for all students in the course might help the instructor determine what course topics or areas may require additional coverage (assessment). New faculty have indicated that this clarification has been most helpful in alleviating anxiety about their role in implementing assessment and evaluation practices that result in improved student learning. ■

Gloria K. Goldman is director of Curriculum and Assessment and Lori E. Zakel is professor and chair in the Department of Communication Arts at Sinclair Community College in Dayton, Ohio.

Horizontal Assessment

W. James Bradley

IN NOVEMBER 2004, CALVIN COLLEGE was reviewed for reaccreditation by the Higher Learning Commission (HLC) of the North Central Association of Colleges and Schools. In its final report, the review team, as it has in evaluating many other colleges, faulted us for insufficient progress in assessment, made some recommendations, and asked for a follow-up report in two years.

The HLC team's recommendations dealt with

- Developing measurable learning outcomes
- Obtaining quantitative data that will permit measurement
- Creating feedback loops that will permit faculty to enhance educational effectiveness through curricular and pedagogical improvement
- Centralizing the administration of academic assessment
- Implanting a formal protocol for program review

One statement particularly concerned us: "The College has not yet succeeded in establishing a coherent approach to assessment." In fact, Calvin faculty and administrators had worked hard for several years to encourage development of assessment plans and clearly state anticipated student learning outcomes in academic departments. But the work was distributed over several offices and it was not the principal responsibility of anyone. Thus, efforts were uneven and the results were not consistent.

Our first step in responding to the recommendations was to establish a new

position—director of assessment and institutional research (DAIR)—with overall responsibility for assessment. Next, we produced a philosophy of assessment that was compatible with our institutional culture. We strongly affirmed assessment and framed it within the concept of reflection. That is, the standard procedure for meeting the assessment requirement of accrediting agencies is a three-step cyclical process: clearly identifying educational objectives, measuring their achievement, and using the results to improve the quality of education. A reflective person or community engages in this same process—asking where they

We produced a philosophy of assessment that was compatible with our institutional culture, framing it within the concept of reflection.



are going, how they are progressing, and making adjustments along the way. Thus, for us, developing a culture of assessment came to mean fostering a climate in which we are encouraged to be reflective people, not just as individuals but as members of a department and of the community as a whole. The role of assessment data is to provide content for that reflective process.

We have developed what we call a horizontal approach to assessment. Rather than asking faculty in an academic department to identify every desired outcome of their program, we ask them to identify one or more outcomes that are especially important and about which they have a concern. Our DAIR then works with departments to clarify these outcomes

and identify an appropriate measure for them. The measure is applied, data are collected, and faculty members discuss them in departmental meetings. At this point, departments typically make policy or pedagogical adjustments and ask for more data. The entire cycle occupies a few months to a year.

We have found this approach to be satisfactory in three ways:

- Prior to adoption of the horizontal approach, assessment at Calvin tended to bog down and the feedback loop was rarely closed. The effort involved in a vertical approach—developing a comprehensive list of outcomes and choosing metrics for each—produced resentment because no fruit was realized from that effort and it required so much time. Assessment was viewed as a bureaucratic imposition from without, not as something organic to the college's life. The resentment was counterproductive in fostering a culture of assessment.
- The horizontal approach produced notable results. In many cases, our faculty members have found the results very helpful; this experience has gone a long way toward fostering that culture of assessment.
- Our faculty members report that this approach has eased their sense of being pressured about assessment.

The horizontal approach to assessment is not original to Calvin College. The concept originated in a software development method called *agile programming*, which quickly yields small amounts of software that are useful to a customer and helpful in building a larger system. It's an alternative to the classical *waterfall model* that was based on extensive detailed design first and that often produced incomplete or cumbersome systems. (See, for example, <http://agilemanifesto.org>.)

The horizontal approach has parallels in other disciplines. It's similar to a social work method known as the "strengths approach," which focuses on understanding clients' individual decision-making processes and helping them make incremental constructive changes rather than beginning with a comprehensive treatment plan. (There are many sources for information about the strengths approach. For example, see <http://www.practicebasedevidence.com/strengths/strengths/principles.html>.)

The horizontal approach is also modeled on a growing movement in science education—undergraduate research. Unlike the graduate education model—developing an encyclopedic knowledge of a discipline before starting research—undergraduate research aims to give students the chance to probe a small area in sufficient depth to enable them to experience the joy of discovery. Having this experience will motivate students to learn more and to expand their research into new areas. Similarly, we are seeing that faculty move toward assessing more objectives after they have the experience of learning from assessment results.

We recognize the principal limitation of the horizontal approach—that it is not as comprehensive as the vertical approach. However, studies of software developed using the waterfall model have shown that 64 percent of the capability built into such software is rarely or never used (Larman, 2005), and we suspect that a similar situation holds for assessment data collected via the vertical approach. Thus, we believe that this limitation is more than compensated for by the fact that the data collected through the horizontal approach are of critical importance to department faculty and hence will be used, that faculty get the results quickly, and that as faculty complete the assessment of one area, they will move on to assess other areas.

Within the first year after we received the HLC report, a set of student learning outcomes (SLOs) for the curriculum as a whole was adopted. Faculty in individual departments also adopted their own SLOs, using the horizontal approach. Within the next year, of our twenty-six departments, fifteen had closed the loop for at least one assessment task and six additional departments had put structures in place for reviewing assessment data and responding to them. Several modes of assessment are currently being used: fourteen departments are using grading rubrics; six are using standardized exams; three are using student interviews; and eight are using various other assessment tools, including embedded questions in tests, student oral presentations, activity and placement records, periodic curricular review, advisory boards, disciplinary accreditation criteria, and locally developed tests. Several changes have been made in response

to assessment data: six departments have modified course syllabi; three have introduced new courses; and three have modified departmental requirements. Other changes include development of a new student handbook, selection of a new text, increased student feedback, increased use of service learning, and addition of an in-service training program for adjunct faculty. Assessment procedures were also put in place for the core curriculum. Based on assessment data, the entry course for the core curriculum that all Calvin students take was significantly modified, and its subsequent evaluations improved markedly. A rubric was developed for assessing student work in that beginning core course, and similar rubrics are being developed for senior capstone courses; use of these will provide a rough before and after measure of student academic development during their years at Calvin.

The HLC's response was positive as well. In a subsequent review, the college's reaccreditation was awarded without qualifications or request for further action.

In summary, then, use of the horizontal approach has dramatically changed the tone of the conversation about assessment on our campus. No longer seen as a bureaucratic imposition, assessment has moved to something that departments find doable, helpful, and not burdensome. ■

Reference

Larman, C. *Applying UML and Patterns*. Englewood Cliffs, N.J.: Prentice Hall, 2005.

W. James Bradley is emeritus professor of mathematics and statistics at Calvin College in Grand Rapids, Michigan.



From the States

Well, Um, We Actually Have a National System for Tracking Student Progress

Peter T. Ewell

Some four years ago, the National Center for Education Statistics (NCES) wanted to build a student “unit record” database that would enable researchers and policymakers to track student progress through postsecondary education to graduation on a national basis. Such a database, if created, could finally address the fact that most students these days attend more than one institution before earning a degree. True, 42 of the 50 states maintain unit record databases and are thus able to keep track of folks within their own boundaries. But few of them have established data exchange arrangements with neighboring states and most do not contain information about private college enrollment. The federal proposal therefore made a great deal of sense. But it hit a wall. Amid arguments advanced primarily by the National Association of Independent Colleges and Universities (NAICU) on the grounds of privacy protection but motivated in reality by a less-than-admirable desire to remain unaccountable, the proposal was dead on arrival.

Overlooked in the midst of this political Sturm und Drang is the fact that for all practical purposes, the nation already has a student unit record database. The National Student Clearinghouse (NSC), a nonprofit corporation located in the suburbs of Washington, D.C., was established about fifteen years ago to track student loan recipients. Over the years, its main business has morphed into a service for verifying college

enrollments and degrees. Looking over the résumé of a prospective job candidate, for example, an employer can pay a small fee and determine whether said candidate actually attended the college claimed and earned a degree. This service has proved so useful that colleges began sending records on all their students. So by the early part of this decade, the NSC could boast data files containing more than 91 percent of the nation’s undergraduate college enrollment. It now maintains records for nearly 15 million currently enrolled students and more than 60 million former students.

To be sure, these records have signifi-

By the early part of this decade, the National Student Clearinghouse could boast data files containing more than 91 percent of the nation’s undergraduate college enrollment.



cant limitations for the kind of longitudinal student tracking that the feds had in mind. Until recently, for example, it was impossible to distinguish a graduate student from an undergraduate and there was no record of a student’s first term of academic history to mark the beginning of a cohort. Besides, the leaders of the NSC had a perfectly good business model going and up to then, it had not occurred to them to mine their millions of records in search of a true national graduation rate. We at NCHEMS got the opportunity to do this with the support of a generous grant from the Lumina Foundation for

Education last year. To be honest, because of the obvious data limitations, I did not have a lot of hope for this effort. But the data turned out to be surprisingly robust and returned statistics that were remarkably consistent with similar results from national longitudinal studies like High School and Beyond or the National Longitudinal Study that Cliff Adelman has so ably led.

Specifically, we were in search of state-level graduation and transfer performance that could be reported in the “Completion” category of Measuring Up, the fifty-state report card on higher education produced every two years by

the California-based National Center for Public Policy in Higher Education. For each previous issue, we had relied on numbers assembled from the NCES Graduation Rate Survey (GRS), compiled on the basis of first-time full-time students examined six years later. Frustrated by the inconsistencies in reporting on the GRS, we wondered whether the data held by the NSC could help us do better.

Using a methodology designed to get around the substantial limitations of the NSC data, we selected a total of 1,002,300 first-time full-time starters at four-year institutions and 986,439

first-time half-time-or-more starters at two-year institutions in the 2002–03 academic year. The cohorts were tracked through a six-year period, noting the state where enrollments took place each year and where any degrees were earned. If a student earned a baccalaureate degree, he or she was not tracked further. But if a student earned an associate degree, tracking was continued to see whether a baccalaureate degree was ultimately earned.

On a national basis, some 65.2 percent of four-year starters had completed a baccalaureate degree at some institution six years later—quite close to the rate reported by Adelman. It turns out that six-year graduation rates for four-year starters are boosted just over six percentage points if all institutions in a state are included and an additional four percentage points if all states are

substantially across states. For example, states like Minnesota, Vermont, and the Dakotas—whose two-year colleges are oriented primarily toward providing vocational and technical programs—show unusually high rates of associate degree completion. Meanwhile, states like California, Florida, Georgia, and Utah—whose public higher education systems are deliberately designed to promote transfer—show relatively high rates of four-year enrollment among two-year starters at the expected time.

Although the ultimate graduation rate estimates we generated were quite close to those that Adelman and others have reported using national samples, we wanted to apply a few other external checks to our results. An easy one was to compare same-institution six-year graduation rates derived from NSC data to those reported publicly

both full-time and part-time students were included in our NSC tracking, whereas GRS only tracks full-time students.

These results confirm the fact that despite NAICU's best efforts to prevent it, we already *have* a national database of student unit records. The latest application of NSC data in the accountability follies is that institution graduation rate statistics drawn from the NSC are being included in the Voluntary System of Accountability. This is the public accountability reporting initiative that is being promulgated by the National Association of State Universities and Land-Grant Colleges and the American Association of State Colleges and Universities in reaction to Margaret Spellings's assault on accreditation last year. The NSC has also just been awarded a \$2.9 million grant from the Bill and Melinda Gates Foundation to expand its capabilities. The most important technical challenges associated with the NSC data have now been fixed, and more and more institutions are providing data. This means that the database will only get better. With respect to student tracking on a national basis, in short, we are in business. And taken together, this story provides one more striking example of a good old American nongovernmental solution.

It is, of course, far too early to predict the outcome of this venture. But I think I can say with confidence that, at minimum, the quality of conversation about this important topic will rise in the year to come.

Peter T. Ewell is vice president of the National Center for Higher Education Management Systems in Boulder, Colorado.

On a national basis, some 65.2 percent of four-year starters had completed a baccalaureate degree at some institution six years later.



included. These patterns vary substantially by state. For example, states like Arkansas and Texas show relatively low out-of-state graduation rates, while states like North and South Dakota show relatively high rates.

Results for two-year starters were calculated similarly. On a national basis, 21.2 percent of two-year starters had earned an associate degree within six years and 40.3 percent had taken at least one credit from an in-state four-year institution. Ultimately, 10.3 percent of these two-year starters earned a baccalaureate degree from an in-state institution and an additional 1.8 percent did so from an out-of-state institution. These rates also vary sub-

stantially across states. For example, states like Minnesota, Vermont, and the Dakotas—whose two-year colleges are oriented primarily toward providing vocational and technical programs—show unusually high rates of associate degree completion. Meanwhile, states like California, Florida, Georgia, and Utah—whose public higher education systems are deliberately designed to promote transfer—show relatively high rates of four-year enrollment among two-year starters at the expected time. Although the ultimate graduation rate estimates we generated were quite close to those that Adelman and others have reported using national samples, we wanted to apply a few other external checks to our results. An easy one was to compare same-institution six-year graduation rates derived from NSC data to those reported publicly on the GRS. Reassuringly, we found that in most states, these rates were nearly identical for four-year starters. The only exceptions were Utah (where students going on religious missions are excluded from the GRS cohort but appear in the NSC data) and six states whose institutional participation in the NSC is comparatively low. For two-year institutions, however, NSC results are substantially lower than those reported publicly on the GRS. This is dismaying but probably accurate for a couple of reasons. First, many community college registrars exclude “non-degree-seeking” students from their official GRS cohorts, whereas all students are included in the NSC. Second,

Using the Academic Skills Inventory to Assess the Biology Major

(continued from page 2)

Based on the data, we did not observe significant differences in the performances of our seniors in the various graduating classes. Students graduating after the implementation of the new curriculum reported skills and experiences similar to those of students graduating before 2002. The reported experiences of biology majors did not show differential effects of the revised and improved curriculum; response patterns were remarkably similar across graduating classes. However, though the differences were not statistically significant, we did observe consistently higher student perceptions of their skills on the scales for written and oral communication; research methodology and statistics; ethics and values; and technology and computers. These findings led to some important discussions concerning the fidelity of our instruction to our curricular goals. Was our new curriculum fully implemented in the way we had designed it? Another important consideration is that the ASI asks students to report only whether they have performed an activity. No information is provided about the quality of the performance. We were pleased to have other methods that address quality and level of learning.

We continued our ASI skill area analyses by examining differences by academic stages. We observed quite significant ($p < .05$) changes in skill acquisition between entering biology majors and seniors for several of the ASI skill areas: written and oral communication; critical thinking and problem solving; research methodology and statistics; and ethics and values. We questioned why our students did not show significant gains in technology and computers, but review of the pretest scores clarified the reason for this result. JMU is a fairly selective institution and is also

listed as one of the top 100 most wired institutions by *Wired* magazine. Our students enter the university and begin the major with advanced technological skills. The results in technology and computers may also be a function of the four basic technology competences (information-seeking skills; word processing; spreadsheet use; and PowerPoint usage) that all first-year JMU students are required to demonstrate as a part of the general education program prior to moving to sophomore status.

Using our four years of biology assessment data, we attempted to conduct repeated measures of individual students over time. We were able to conduct paired samples tests for the biology majors

Written and Oral Communication (.47), Groups/Organization/Community (-.18), Critical Thinking and Problem Solving (.26), Research Methodology and Statistics (.41), Ethics and Values (.36), and Technology and Computers (.20). The biology faculty noted that the areas in which students reported the highest gains are those most aligned with the program's learning objectives.

While faculty members were pleased to see some expected changes in students' perceptions over their academic career, they had hoped that the new curriculum would produce more readily observable changes across the cohorts. One possible interpretation and solution may be that the instrument needs to be revised to more

Our findings led to some important discussions concerning the fidelity of our instruction to our curricular goals. Was our new curriculum fully implemented in the way we had designed it?



who entered in fall 2004 and graduated in spring 2008. We observed significant ($p < .05$) differences for the results on several ASI scales. We also calculated effect sizes (Cohen's d) to aid in the interpretation of these findings. An effect size standardizes the observed differences to a common metric—in this case, the standard deviation (Cohen, 1988). Effect sizes are often used to move beyond statistical significance testing to make judgments about the practical meaning of study results. Cohen provided general guidelines for interpreting these values. Effect sizes around .20 are considered small, .50 medium, and .80 large. The scales and the effect sizes observed are as follows:

closely mirror the skill sets intended to be developed by the biology curriculum. Another interpretation suggests the need to assess the quality of student performances directly rather than rely on their reports of completion of identified activities. This idea points to an oft-cited good practice in assessment: using multiple measures to assess each objective. Fortunately, the Biology Department's faculty members employ several other measures that map to their goals and objectives.

Overall, consideration of our ASI data and our analyses has promoted several very useful conversations among faculty members and students. A frequent observation made by students after

completion of the instrument is that they have a renewed appreciation of the many achievements and activities they have completed. It appears that when surrounded by peers who are all capable of doing senior-level laboratory work, data analysis, and write-ups, students tend to take for granted that all students can engage in these activities effectively. They do not appreciate how much they have developed. The ASI provides a means for students to gauge their maturity as students and emerging professionals. As for the faculty's point of view, we would like to see greater gains in students' completion of tasks for many of the skill areas. It may be that the ASI items do not directly tap the intended outcomes. For example, observing a statistically significant decline (but small effect size) in student scores related to groups, organizations,

and community may not be an unexpected outcome for students in this major. We certainly do not see objectives associated with this skill area emphasized in biology courses. We may follow the lead of the Integrated Science and Technology baccalaureate program and the Technical and Scientific Communications master's program at JMU by modifying the ASI to align more closely with our desired student transformations. Reducing the number of skill area categories and items may also enhance student interest as they complete the instrument. We hope to implement the new Biology ASI instrument in spring 2009. Stay tuned. ■

References

Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. (2nd ed.) Hillsdale, N.J.: Erlbaum, 1988.

Kruger, D. J., and Zechmeister, E. B. "A Skills-Experience Inventory for the Undergraduate Psychology Major." *Teaching of Psychology*, 2001, 28, 249–253.

Monroe, J. D., and Hurney, C. A. "CCLI and Curriculum Change in Biology." *CUR Quarterly*, 2002, 22(3), 122–125.

Sundre, D. L. "The Academic Skills Inventory: A Measure of Skills Experience." *Assessment Update*, 2003, 15(4), 8–9.

Kyle Seifert is assistant professor of biology, Carol A. Hurney is associate professor of biology, Clifton J. Wigtil is a graduate assistant, and Donna L. Sundre is professor of graduate psychology and executive director at the Center for Assessment and Research Studies at James Madison University in Harrisonburg, Virginia.

RESOURCES

Susan Kahn

Spurlin, Joni E.; Rajala, Sarah A.; & Lavelle, Jerome P. *Designing Better Engineering Education Through Assessment*. Sterling, VA: Stylus, 2008.

In 2000, the Accreditation Board for Engineering and Technology (now ABET, Inc.) issued a revised framework for accrediting engineering and technology programs that shifted the focus of accreditation from what is taught in these programs to what is learned. Intended as a new paradigm for engineering education, the so-called EC2000 criteria required programs to engage in systematic processes for assessing, documenting, and improving student learning. Nine years later, many engineering programs continue to struggle with issues of assessment methodology, curriculum, and faculty buy-in.

This new volume, *Designing Better Engineering Education Through Assessment*, aims to cut through the education jargon that too frequently characterizes assessment discussion and provide a practical guide to assessment theory and practice as applied to engineering.

Parts of the book are applicable to any discipline. Early chapters provide an introduction to the basics of assessment and include a useful discussion of assessment of graduate-level programs, a topic that receives scant attention in the assessment literature. The volume is filled with helpful examples, illustrations, and tables, including a sample curriculum map (p. 31), a succinct summary of the pros and cons of various assessment methods (pp. 55–58), and a step-by-step exercise in interpreting assessment data (pp. 115–116). More specific to

engineering are the examples of "embedded" or "authentic" assessment for each ABET-mandated learning outcome (pp. 69–74) and the sample rubrics for assessing senior design projects, team engineering projects, lab reports, and other common components of engineering programs (pp. 86–103). Chapters on assessing freshman engineering programs, capstone engineering experiences, and course and program-level assessment in engineering maintain the focus on translating assessment theory into practical applications.

A substantial section of the book on barriers and challenges to assessment and approaches to overcoming these obstacles might easily be applied across academic disciplines. J. Joseph Hoey, IV and Eleanor W. Nault, the authors of Chapter 6, argue that trust is "the most basic and fundamental actor involved in

Resources

(continued from page 15)

assessment design and implementation” (p. 174). In addition to overcoming this “assessment anxiety” (p. 181), assessment leaders must address the reluctance of some faculty members to engage in intro-

last idea in the following chapter, arguing that assessment “requires a pervasive, cooperative effort with attributes distinct from the individual teaching, research, and service activities otherwise required of faculty

[A]ssessment leaders must address the reluctance of some faculty members to engage in introspection or admit the possibility that change may be needed.



spection or admit the possibility that change may be needed, as well as the propensity for faculty to view themselves as “solo contractors,” rather than as members of a larger community with collective responsibilities (pp. 177–178). Authors Sherra Kern and Karan Watson expand this

members” (p. 190). These authors posit three levels of resistance to change (pp. 192–194), noting that faculty members are “highly trained to be skeptical” and that “for complex issues there are always debatable questions and trade-offs that must be made” (p. 191).

Throughout the book, the authors emphasize such fundamentals as using multiple assessment methods, basing assessment on observable student behavior, and incorporating assessment into the curriculum, rather than treating it as an add-on.

Equally useful to assessment professionals working with engineering programs and to engineering faculty new to assessment, this volume is likely to be a fixture on assessment leaders’ bookshelves for some time to come.

Susan Kahn is director of the Office of Institutional Effectiveness at Indiana University–Purdue University Indianapolis.