National Institute for Learning Outcomes Assessment

October 2016

Tracing Assessment Practice as Reflected in *Assessment Update*

Trudy W. Banta, Peter T. Ewell, and Cynthia A. Cogswell

Foreword by George D. Kuh

Occasional Paper #28

learningoutcomesassessment.org

About the Authors

TRUDY W. BANTA is a professor and vice chancellor emerita at Indiana University-Purdue University Indianapolis. Dr. Banta is the recipient of several national awards for her work in outcomes assessment and has consulted with faculty and administrators in 47 states as well as given several invited addresses in other countries and organized conferences for domestic and international audiences. Dr. Banta has written or edited 20 published volumes on assessment, contributed 30 chapters to other published works, and written more than 350 articles and reports. She is the founding editor of *Assessment Update,* which won a national award for overall excellence in 2003 and the APEX Award for Editorial Writing in 2012.

PETER T. EWELL is president emeritus of the National Center for Higher Education Management Systems (NCHEMS) and NILOA Senior Scholar. Dr. Ewell's work focuses on assessing institutional and higher education system effectiveness and the outcomes of college, involving both research and direct consulting with institutions and state systems on collecting and using assessment information in planning, evaluation, and budgeting. He has consulted with over 375 colleges and universities and twenty-four state systems of higher education on topics related to performance indicators and the assessment of student learning as well as written commissioned papers for many agencies including the Education Commission of the States, the National Governors' Association, the National Conference of State Legislators, and the National Center for Public Policy in Higher Education.

CYNTHIA A. COGSWELL is a postdoctoral fellow for Assessment and Evaluation at Dartmouth College. At Dartmouth, her work focuses on assessing and evaluating campus efforts to enhance learning and classroom pedagogy. Her PhD is from Indiana University in Higher Education and Student Affairs. Through her research, Cynthia aims to enrich the national dialogue about student learning in higher education and regional accreditation in the United States. She is interested in understanding how institutions use the process of assessment to improve student learning; the relationship between institutions and accreditors; and supporting institutional improvement through regional accreditation.

Please cite this paper as:

Banta, T. W., Ewell, P. T., & Cogswell, C. A. (2016). *Tracing Assessment Practice as Reflected in Assessment Update* (NILOA Occasional Paper No. 28). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Contents

Foreword

At some future point, when a definitive history of the assessment movement is written, one of the most frequently cited, influential publications will be *Assessment Update* (*AU*). Since 1989, this bimonthly newsletter has been published by Jossey-Bass in partnership with Indiana University-Purdue University Indianapolis (IUPUI). It is no coincidence that the two most frequent contributors to *AU*, Trudy Banta—*AU's* founding editor and intellectual muse—and Peter Ewell, are also among the most prolific thinkers and writers shaping the scholarship and practice of student learning outcomes assessment.

In this featured NILOA occasional paper, Banta and Ewell with the assistance of Cynthia Cogswell mine the pages of *AU* between 2000 through 2015 to distill the major themes and advances that characterize the evolution of assessment as a field of professional practice. This 15-year period was especially fertile in terms of expanding the numbers of approaches and instruments available for documenting learning outcomes and institutional effectiveness. For example, Borden and Kernel (2013) reported that only 26 assessment tools were available in 2000, a number that mushroomed to 250 by 2009.

The litany of topics addressed in their respective *AU* articles integrated with recounting some of their scores of foundational contributions to the field since the early 1980s—the beginnings of the "assessment movement" when Banta and Ewell first met one another—provide a compelling albeit incomplete record of their visionary leadership in building a scholarly foundation for assessment work and advancing promising practices.

NILOA is pleased to disseminate this analysis of the challenges and opportunities that characterize the recent past of the assessment movement. It is especially timely, as both Banta and Ewell recently "retired" after decades of superlative service to their respective organizations—Banta at IUPUI and Ewell at the National Center for Higher Education Management Systems (NCHEMS). We also are fortunate they enlisted Cogswell to help memorialize what has been an especially formative period in the assessment movement, as she represents the best of what we expect to be the promising future of this work.

Another reason the publication month of this "Tracing Trends" paper is noteworthy is that it coincides with the 25th annual Assessment Institute in Indianapolis, the leading international conference, a forerunner of which Banta started in Tennessee in 1986. It is hard to imagine assessment work making the progress it has without the ideas addressed and best practices shared at this annual signature event.

On behalf of the thousands of scholars and practitioners past and present engaged in student learning outcomes assessment, NILOA salutes the contributions of *Assessment Update* to the art and science of learning outcomes assessment and the legendary careers of Trudy Banta and Peter Ewell.

George D. Kuh
NILOA Senior Scholar and Founding Director

References

Borden, V. M. H., & Kernel, B. (2013). *Measuring quality in higher education: An inventory of instruments, tools, and resources* [Online inventory]. Retrieved from http://apps.airweb.org/surveys/Default.aspx

Prologue

In 1980, Trudy Banta was an administrative intern in the office of then-Chancellor Jack Reese when he learned that his campus, the University of Tennessee at Knoxville (UTK), would be subject to Tennessee's new performance funding initiative. To be eligible for up to two percent of the university's budget for instruction, an annual report would have to be submitted to the Tennessee Higher Education Commission demonstrating that, among other things, alumni had been surveyed and that graduating seniors had been tested in general education and in their major field. Chancellor Reese said, "This looks like an abridgement of academic freedom to me." Banta said, "It looks like program evaluation is coming to higher education to me." Reese said, "Well then you can handle this."

So began Banta's career in outcomes assessment—working with deans, department chairs, and faculty to choose or develop measures that would demonstrate student learning.

In 1981, the Kellogg Foundation funded a national project aimed at increasing the use of information on student outcomes in institutional decision-making. Banta proposed that UTK be one of the institutions involved in this project, and UTK was selected to participate. Peter Ewell was chosen to manage the national project, which gave Banta and Ewell their first opportunity to work together. This collaboration has taken several forms over the years, including service on project advisory boards, presentations at conferences, and even a trip to China for that nation's first conference on evaluation in higher education.

In 1989, Jossey-Bass named Banta the founding editor of its bi-monthly periodical, *Assessment Update: Progress, Trends, and Practices in Higher Education*, and she invited Ewell to contribute a continuing column, which he titled "From the States." Now Banta has written over 100 "Editor's Notes" columns, and Ewell has drafted almost 70 in his series.

In 1992, Banta moved to Indianapolis to coordinate assessment activities at Indiana University-Purdue University Indianapolis (IUPUI) and to serve on the faculty of Indiana University's bi-campus Higher Education and Student Affairs program. By 2015, Banta had worked with then-doctoral student Cynthia Cogswell on several projects and invited her to conduct a content analysis of *Assessment Update* (*AU*) articles since 2000 in a beginning effort to identify trends in outcomes assessment from 2000-2015.

The identification of trends in outcomes assessment began by examining the frequencies of key words, which had been assigned over the years to describe *AU* articles, occurring in a dozen primary document families. With almost 1,000 *AU* articles falling into one or more of these categories, we needed to pare that number for further study. From the initial analysis we selected documents from a half-dozen primary document families, and we decided that the series of Ewell and Banta columns could provide the best insights into the evolution of the field in the period 2000-2015. Cogswell then conducted an exploratory content analysis, a sense-making effort that allows themes to emerge (Patton, 2002), using both borrowed codes from the initial frequency analysis and open coding, assisted by the qualitative computer software program, ATLAS.ti. This analysis was woven together by Banta and Ewell to form the categorical and chronological summaries of their columns in the paper that follows.

Tracing Assessment Practice as Reflected in *Assessment Update*

Trudy W. Banta, Peter T. Ewell, and Cynthia A. Cogswell

Introduction

Outcomes assessment in higher education may be defined as a "process of providing credible evidence of resources, implementation actions, and outcomes undertaken for the purpose of improving the effectiveness of instruction, programs, and services in higher education" (Banta & Palomba, 2015, p. 2). At mid-decade there has been no shortage of studies aimed at taking stock of developing trends in higher education, including equity in access and degree attainment (Cahalan & Perna, 2015), state funding trends and affordability of college (Emrey-Arras, 2014), and return on the investment in postsecondary education (Baum, 2014). There have even been recent studies on trends in outcomes assessment in the United Kingdom (Medland, 2016) and internationally (Coates, 2016; Pereira, Flores, & Niklasson, 2015). But we are not aware of a recent attempt to chronicle developments in higher education assessment with a primary focus on the United States. This is the task that the current authors have undertaken.

__Few have been more engaged participants or closer observers of the development of the field of outcomes assessment in higher education since its inception in the early 1980s than Peter Ewell and Trudy Banta, whose collaboration began in 1981.__

Few have been more engaged participants or closer observers of the development of the field of outcomes assessment in higher education since its inception in the early 1980s than Peter Ewell and Trudy Banta, whose collaboration began in 1981. Since 1989, Ewell and Banta have had the privilege of contributing their thinking about assessment issues and practice in periodic columns in the bimonthly *Assessment Update (AU),* published and distributed by Jossey-Bass, a division of Wiley. With research assistance from Cynthia Cogswell, Banta and Ewell have drawn on material from their columns to identify and describe trends, primarily about outcomes assessment in higher education in the U.S., with some references to international developments. The two columnists have looked almost exclusively at their contributions to *AU* between 2000 and 2015, but occasionally they situate their themes in the broader context of the prior two decades. The four themes that serve to organize this paper include

- international, national, and state developments;
- accountability issues;
- assessment methods and strategies; and
- encouraging developments.

An overall summary concludes the paper.

International, National, and State Developments Shaping Assessment

Similar Responses in the U.S. and Europe

In Europe, higher education policies and procedures are determined primarily by central governments. So the first assessment directives for European universities emanated during the 1980s and 1990s from national education ministries. In the United States, the individual states provide primary direction for colleges and universities; thus the first assessment mandates that emerged in the early and mid-1980s originated in state legislatures and governing boards. Initally, Peter Ewell's "From the States" columns for *Assessment Update* were focused on state policies concerning assessment and institutional effectiveness. By 2000, Ewell had published seventeen columns focused on various states. He had also begun to branch out somewhat, including columns on accreditation, performance indicators, and international developments. This was timely, because by the year 2000 the main actors in U.S. assessment policy were shifting from the states toward regional and specialized accreditors with an occasional, and frequently disruptive, foray by the federal government. This topical shift was reflected in the fact that both "From the States" columns published in 2000 were about individual states—Utah (2000, January-February) and Colorado (2000, September-October)—while the next year featured a column about accreditation in teacher education (2001, May-June).

__Initially, Peter Ewell's "From the States" columns for Assessment Update were focused on state policies concerning assessment and institutional effectiveness. By 2000, Ewell had published seventeen columns focused on various states.__

In 2001 (January-February), Ewell also wrote about the National Center for Public Policy and Higher Education's (2000) first fifty-state "report card" on higher education, an effort in which he was substantially involved. Entitled *Measuring Up*, this report graded all fifty states every two years in five areas—preparation, participation, affordability, completion, and benefits. A sixth category, learning, was awarded a grade of "Incomplete" in the report's early years, but in 2003 (May-June) this finding furnished the impetus for a five-state demonstration project led by Ewell and Margaret Miller. This involved administering direct measures of student achievement in five states—Kentucky, Illinois, Oklahoma, Nevada, and South Carolina—including the brand-new Collegiate Learning Assessment and four tests from ACT's Work Keys battery. Secondary source data on each state's literacy performance on the National Assessment of Adult Literacy (NAAL), as well as graduate admissions and professional licensure test pass-rates were also assembled. *Measuring Up*'s last report was issued in 2008 and the initiative had by that time significantly shaped the national dialogue about higher education performance.

Despite the important influence of governments, by 2000 faculty presentations at national and international meetings were more likely to focus internally on institutional effectiveness and teaching and learning issues than on national or state assessment initiatives, as they had during the two prior decades (Banta, 2000, November-December). This is not surprising since a nation-wide survey conducted in 1998 (Peterson et al., 1999) revealed that while accreditation was an important influence in motivating college and university faculty and administrators to begin work on outcomes assessment, subsequent faculty engagement in assessment was undertaken in their own ways for their own purposes (see www.umich.edu/~ncpi/52/52.html).

In terms of conference presentations in the 1990s, some of the questions asked at assessment meetings in the U.S. included: Why do we have to use assessment to prove ourselves to the public? What is it that state agencies and accreditors really want of us? How can we meet their demands with minimal effort and disruption? But in 2000 (November-December), when Banta compared the programs of the annual meetings of the European Association for Institutional Research (EAIR) and the American Association for Higher Education (AAHE), she found parallel emphases on such topics as institutional portfolios; quality management; academic audits; performance indicators; league tables; and surveys for students, faculty, graduates, and employers. In Banta's 2001 (September-October) column, she compared presentation topics at the 2001 AAHE Assessment Conference in Denver, Colorado and the 13th International Conference on Assessing Quality in Higher Education in Glasgow, Scotland and again found similarities. Using the Denver agenda as the model, organizational structures for getting started in assessment were of most interest, followed by uses of technology and surveys in assessment, assessment approaches in general education and the major, and electronic portfolio applications.

In Banta's 2002 (November-December) column, she summarized an article entitled "The End of Quality," written earlier that year by Lee Harvey (2002), then at the University of Central England in Birmingham. The article was based on proceedings of a conference that Harvey had convened in Birmingham a year earlier addressing the key question: Has control of quality in higher education been usurped by market forces and information technology? That is, with higher education becoming a big business with alternative providers extending their services worldwide, will the world market ultimately be the arbiter of quality? Delegates to the Birmingham conference were quick to point out that their governments' quality monitoring mechanisms created extra work on campuses, added to bureaucracy, asked the wrong questions, focused on accountability rather than improvement, and produced only short-term responses with no lasting impact on student learning. Banta noted the contrast between the tenor of the conference in England and that at the 2002 AAHE conference in Boston. Since U.S. institutions enjoy more autonomy, the principal theme in Boston was: How can we use assessment to improve teaching and learning? Sessions described methods for assessing writing, critical thinking, speaking, and information literacy, with particular emphasis on electronic portfolios. At about the same time, Ewell (2003, January-February) noted strong parallels between government-sponsored institutional review in England and counterpart assessment activities in the U.S. stimulated largely by regional accreditation. Both featured strong institutional resistance to accountability-centered attempts to assess institutional effectiveness. Clearly academics on both sides of the Atlantic prize their independence and want to control how pressures for external scrutiny unfold on their home turf.

Proceedings of a conference convened in Birmingham a year earlier addressed the key question: Has control of quality in higher education been usurped by market forces and information technology? That is, with higher education becoming a big business with alternative providers extending their services worldwide, will the world market ultimately be the arbiter of quality?

Sour Notes in Europe

In August 2009, Banta (2009, November-December) attended the annual EAIR meeting held that year in Vilnius, Lithuania. She had just read on the ASSESS listserv that several assessment coordinators in the U.S. had experienced personal public verbal attacks by faculty colleagues angered by attempts to engage them in assessing student learning outcomes and using the findings to improve their teaching. In Vilnius, Banta attended several sessions that featured similar negative reactions to accountability requirements in Europe. Two faculty members from the United Kingdom argued that quality assurance was naturally resented by academics because it was imposed by external authorities; a proposed solution was to provide administrative support to faculty in individual departments so that they could determine their own approaches to assessment. Another prominent British scholar lamented the fact that attempts to build quality-oriented cultures at universities had resulted in far too few benefits to justify the time and money invested. He asserted that more value could have been derived from a similar investment in faculty development. Then-current leader of the European University Association said in a plenary address that the involvement of academics in the Bologna Process,[1] had been relatively poor and that the process known as tuning had produced as much competition among universities as tuning's intended subject-specific convergence. Five years earlier, in commentary on the relatively poor attention paid to learning outcomes in the Bologna Process, Ewell (2004, November-December) had made similar points. Other EAIR speakers warned that specifying learning outcomes for degree programs would narrow the curriculum and would not account for learning occurring outside the classroom. Four undergraduate students who had won the EAIR best paper award for scholars under the age of 35 were so disillusioned by their experiences under the Bologna Process that they had adopted "Obama-style community organizing tactics" to attract faculty attention to their concerns. The students wrote their own learning outcomes and reported that faculty were considering their list. They said that they had little difficulty gaining faculty support "because Bologna is also viewed negatively by faculty."

In 2006 the work of the Commission on the Future of Higher Education was a topic for columns by both Banta and Ewell. Ewell covered reports of the Business Higher Education Forum, which for the first time in that organization's history addressed accountability for higher education, and the State Higher Education Executive Officers Commission on Accountability—both of which foreshadowed conclusions of the Spellings Commission.

Accountability Issues Influencing Assessment

The Press to Assess with a Test

In the immediate aftermath of the turn of the millennium, many states were experimenting with a new round of required standardized testing. This included Utah with the ACT Collegiate Assessment of Academic Proficiency (CAAP) (Ewell, 2000, January-February) and Colorado with the ETS Proficiency Profile (Ewell, 2000, September-October).

Actions affecting accountability on the part of the federal government were also an intermittent topic of "From the States." In 2001 (May-June), Ewell reported critically on the then-new federal practice of public reporting of teacher examination pass rates by state—a policy that bordered on the irresponsible because of the differing methodologies of the states in calculating this statistic.

In 2005 (May-June), Ewell reported on Missouri's experimental use of the Collegiate Learning Assessment (CLA), funded for both public and private institutions by the state's higher education coordinating board. The CLA was also featured in the University of Texas' new accountability system developed under the leadership of Charles Miller, soon to be chair of U.S. Secretary of Education Margaret Spellings' Commission on the Future of Higher Education, as Ewell reported in 2006 (May-June).

In 2006, the work of the Commission on the Future of Higher Education was a topic for columns by both Banta and Ewell. Ewell (2005, September-October) covered reports of the Business Higher Education Forum (BHEF), which for the first time in that organization's history addressed accountability for higher education, and the State Higher Education Executive Officers (SHEEO) Commission on Accountability—both of which foreshadowed conclusions of the Spellings Commission. Ewell also testified on assessment and accountability before the Spellings Commission in 2005 and reported in 2009 (January-February) on the higher education community's reactions to the Commission in the form of the Voluntary System of Accountability (VSA) and the promising, but unfortunately short-lived, Alliance for New Leadership on Student Learning and Accountability.

[1] The Bologna Process was and continues to be an iterative series of policy agreements among European countries to move toward comparability in student learning outcomes across institutions. To learn more see Adelman (2008) or Gaston (2010).

Banta served on the Learning Outcomes Working Group that advised the National Association of State Universities and Land-Grant Colleges (now APLU) and the American Association of State Colleges and Universities (AASCU) as their staff developed the VSA. In 2008 (March-April) she outlined eight reasons for arguing during the working group's deliberations against the use of standardized tests of generic skills like writing and critical thinking in comparing student learning outcomes across institutions. Then she noted that the very large California State University and University of Texas systems had by presidential fiat adopted a common test for freshmen and seniors on all campuses. Moreover, the Paris-based Organisation for Economic Cooperation and Development (OECD) had convened representatives of Korea, Japan, the European Union, and the U.S. to plan a pilot project aimed at developing a test of culturally relevant tasks to measure college students' critical thinking abilities in countries around the globe. In the face of the imminent international flirtation with standardized testing, Banta offered six suggestions for institutions faced with the need to administer a standardized test of generic skills:

Banta described the difficulties of designing one test that could, in a few hours on a single day, measure even a small part of what a student knows and can do.

1. Involve respected faculty, staff, and student leaders in studying the tests being considered.
2. Ask this group to compare the names and descriptions of scales in each test to stated institutional goals for general education.
3. Give members an opportunity to take the test they perceive to be the best match and to review their scores. Ask them to predict their highest and lowest areas on each test.
4. Ask a few students to take the test alone with a faculty observer, describing aloud their thought processes as they approach each item.
5. Encourage warranted use of the test data in improving curriculum, instruction, and student services by publicizing the work of the selection committee and advocating that attention be focused on the findings.
6. Conduct institution-specific studies to demonstrate the usefulness of the test.

In 2006 (July-August), Banta made an effort to keep higher education from repeating past failures by providing a brief history of the search for a simple solution to the problem of demonstrating how much students learn in college. She recalled research that she and Gary Pike had conducted at the University of Tennessee, Knoxville beginning in 1986 based on testing of freshmen and seniors required as part of the state's performance funding initiative. They found that as much as 80 percent of the variance in institution-level scores was due to the knowledge and skills students brought to college (high correlations between college students' test scores and their entering ACT/SAT scores) as opposed to what they had learned there. They also discovered that there were various methods for calculating value added, or gain (or loss) from freshman to senior years, each of which could yield a different result, and all of which were unreliable. Banta described the difficulties of designing one test that could, in a few hours on a single day, measure even a small part of what a student knows and can do. A strength of higher education in the U.S. is its diversity, and the students attracted to a small private religiously-affiliated college are themselves different in some ways and will certainly have a different learning experience than students who choose a large public research university—and both will receive a valuable education. In addition, no test is truly content-free, so students majoring in some disciplines will be advantaged while others are disadvantaged.

At about the same time in 2008, Ewell (2008, July-August) was serving as chair of the Technical Advisory Group of the multinational Assessment of Higher Education Learning Outcomes (AHELO) Feasibility Study sponsored by the OECD and he reported on the experience. The Feasibility Study ultimately involved some seventeen countries or jurisdictions in administering assessments in Engineering, Economics, and "generic skills" (the instrument here was a version of the CLA). All three instruments employed both multiple choice and constructed response items. While the demonstration was technically successful, formidable challenges to generalizability were associated with translation and context in such differing countries as the U.S. and Egypt. Professional contacts made during this experience led Ewell (2013, May-June) to follow up and report on Colombia's national assessment program that involves testing almost half a million students in postsecondary institutions in multiple fields annually.

Banta has long believed that the most valuable outcomes assessment methods provide direction for improvement internally but also demonstrate accountability.

Banta has long believed that the most valuable outcomes assessment methods provide direction for improvement internally but also demonstrate accountability. The Voluntary System of Accountability (VSA), introduced in 2007, defined accountability in part on demonstrating value-added based on scores on standardized tests of generic skills and on comparing institutions on this metric. At the annual Assessment Institute in Indianapolis in 2009, Banta (2010, January-February) took an informal poll by asking the approximately 900 people attending the first session if they believed standardized tests of generic skills could provide direction for campus improvement. Very few indicated agreement. A larger number, but not all, agreed that these tests could even be helpful in demonstrating accountability. On the other hand, most participants agreed that information yielded by student portfolios could do both. Banta then presented some encouraging work on several campuses designed to establish inter-rater reliability for the rubrics needed to evaluate student work in portfolios and other authentic products. But she also raised some of the psychometric issues that were just beginning to surface regarding the reliability and validity of rubrics. Until 2008 Banta had assumed that a rubric could be created to assess virtually anything. But in that year she read several articles suggesting that many faculty resent the process of trying to reduce a complex concept to a common definition (Banta, 2008, July-August). Imposing the requirement of inter-rater agreement to establish the reliability of rubrics provokes further ire. Banta concluded that she would continue to use and to recommend rubrics, but with the understanding that not everyone would embrace them. She noted that ETS had been working to improve the SAT for 80 years and wondered if it would take that long to convey on rubrics even the credibility currently accorded the SAT!

During the first two years of the VSA's history, the subject of student motivation to do well on the low-stakes tests it recommended were talked about on campuses, but the testing companies expressed little concern about the effects of motivation on scores and on the value-added calculation. In 2009 this began to change as scholars turned their attention to the psychometric problems posed by unmotivated examinees in low-stakes testing situations. In 2010 (July-August), Banta called attention to a special issue of *The Journal of General Education* (Sundre & Colleagues, 2009) to which faculty associated with the Center for Assessment and Research Studies at James Madison University (JMU) contributed articles on this topic. They introduced their Student Opinion Scale (SOS), which can be given to students after they take a test. The instrument presents five questions about the level of importance students attach to doing well on the test and five questions about the levels of effort and persistence they have expended in completing the test. The authors convey the importance of using SOS scores or item response times to remove the scores of unmotivated examinees when calculating institutional averages. They also document significant improvements in these averages when test proctors are trained to stress the importance of conscientious performance and encourage engaged test-taking behavior.

Pushing Back

In 2012 (November-December), Banta described her reluctance to sign on to the recommendation to administer standardized tests of generic skills like writing and critical thinking when in 2007 she served on the task force on assessing student learning that contributed to the development of the VSA. Between 2007 and 2012 several studies produced evidence to support Banta's contention that the recommended tests and the value-added measure being used to report and compare institutions' scores on the tests were not valid for the purpose of assessing and comparing the quality of colleges and universities. First, research financed by the Fund for the Improvement of Postsecondary Education (FIPSE) demonstrated that while the three tests specified in the VSA were related, they were not equivalent measures; thus scores earned on one test could not appropriately be compared with scores derived from either of the other two tests (Klein, Liu, & Sconing, 2009). Then a 3-year study involving 47 institutional members of the Council of Independent Colleges in administering the CLA to their students on multiple occasions was concluded with the finding that no specific improvements on any of the 47 campuses could be attributed to use of the CLA (Paris, 2011). And finally, a formal evaluation of the VSA conducted by principals associated with the National Institute on Learning Outcomes Assessment (NILOA) revealed that the requirement to report the value-added statistics based on scores derived from the standardized tests of generic skills had undermined credibility and acceptance of the VSA and had discouraged other institutions from participating (Jankowski et al., 2012). A new task force assembled by APLU and AASCU modified the student learning assessment component of the VSA to include, in addition to the three tests, ratings of student performance on the Written Communication and Critical Thinking Valid Assessment of Learning in Undergraduate Education (VALUE) rubrics developed by staff of the Association of American Colleges and Universities (AACU) (Rhodes, 2009). Meanwhile, as Ewell (2010, September-October) reported on the development of the community college version of the VSA—the Voluntary Framework of Accountability (VFA)—the designers of this initiative had eschewed the use of standardized tests altogether.

Another theme of "From the States" in this fifteen-year period was the growing salience of state-level unit record data in creating institutional performance measures for accountability purposes. Two columns in 2009 addressed this issue. Ewell (2009, September-October) described the significant sums of federal stimulus package funds being invested in upgrading state-level databases. He also examined the data holdings of the National Student Clearinghouse—an organization that collects unit-record enrollment and graduation data directly from colleges and universities on a voluntary basis, thus constituting, in effect, a national unit-record database (Ewell, 2009, May-June).

Another theme of "From the States" in this fifteen-year period was the growing salience of state-level unit record data in creating institutional performance measures for accountability purposes.

Although regional accreditation was the topic of just three of Ewell's columns before 2000, this topic received a good deal of attention in the decade and a half that followed. In 2002 (January-February), he examined some significant changes in regional accreditation practice that were occurring, stimulated in part by grants made by the Pew Charitable Trusts. The most extensive of these changes were made by the Western Association of Schools and Colleges (WASC) Senior College Commission, which issued a brand-new Handbook in 2001 featuring a two-visit accreditation cycle and a much more focused look at student learning outcomes. By 2007, reflecting conversations occurring among members of the Spellings Commission and the Secretary's subsequent attempt to change its role through negotiated rulemaking, accreditation was coming increasingly under fire, so Ewell's (2007, January-February) column was fittingly titled "Accreditation in the Hot Seat." Ewell's column in 2011 (September-October) recounted the growing criticisms of regional accreditation and reviewed the principal "fixes" being proposed by various parties. Among them were re-balancing institutional coverage across the regions to achieve a more equitable distribution of institutions among commissions, creating multiple levels of institutional recognition, and aligning learning outcomes language across the regions.

By 2007, reflecting conversations occurring among members of the Spellings Commission and the Secretary's subsequent attempt to change its role through negotiated rulemaking, accreditation was coming increasingly under fire, so Ewell's (2007, January-February) column was fittingly titled "Accreditation in the Hot Seat."

In 2011 (September-October), Ewell also announced the creation of an Accreditation Task Force by the American Council on Education (ACE), on which he eventually served. This body, for the most part, reaffirmed the original purposes of institutional accreditation and ratified its importance as the country's principal quality assurance mechanism. While Task Force members also reaffirmed accreditation's basic approach—institutional self-study, followed by a site visit by peer reviewers, resulting in an accreditation finding by the appropriate commission—they did call for far more transparency in disclosing the results of accreditation reviews. Members advocated a "risk-based" stance that would accord a "lighter touch" in review to well-established institutions having no historical problems with resources or academic quality. One of Ewell's columns the next year (2012, July-August) examined the growing resistance to accreditation by elite institutions, manifested particularly on the ACE Task Force, as these institutions asserted that the process was burdensome and not needed by institutions of "obvious" quality. Putting a happy face on this phenomenon, Ewell concluded that this meant accreditation was finally getting attention from a constituency that until then could afford to ignore it. In 2015 (March-April), he examined the growing phenomenon of U.S. accreditors and offshore national quality assurance agencies reviewing and recognizing U.S. programs offered in other countries.

Assessment Methods and Strategies

Early Developments

In confronting the need to assess student learning outcomes, the first question many faculty members ask is, "I give students grades—what more is needed?" In 2000 (January-February), Banta addressed this basic issue. When the information contained in a course grade can enable one to identify what a student earning an A can do with respect to specified student learning outcomes that a student earning a B, for instance, cannot do, then grades may be aggregated to determine the strengths and weaknesses of a class of students. However, very few grading systems are based strictly on what students know and can do; some portion of the grade may be based on attendance, effort related to a group assignment, turning in a revised paper, or extra credit for participating in an out-of-class activity. So a point total or a letter grade will not usually provide the specific evidence of achievement of learning outcomes that is useful in outcomes assessment. This is not to say, however, that course-embedded assignments cannot be valuable sources of information about the strengths and weaknesses of instruction and curricula. Any homework assignment, paper, test, or group project that has been designed to assess one or more of the student learning outcomes specified for a course may be evaluated to assign a grade and to provide detailed feedback to an individual student. Then, a second look at the individual responses aggregated across students in a course can yield the information about group strengths and weaknesses that is needed for outcomes assessment—that is, to suggest warranted improvements in pedagogy, curriculum, or student support services such as advising.

Any homework assignment, paper, test, or group project that has been designed to assess one or more of the student learning outcomes specified for a course may be evaluated to assign a grade and to provide detailed feedback to an individual student.

In a 2001 column (November-December), Banta recalled the experience of working with Catherine A. Palomba on their book *Assessing Student Competence in Accredited Disciplines* (Palomba and Banta, 2001). In eight wide-ranging fields (teacher education, social work, nursing, pharmacy, computer science, engineering, business, and fine arts), names of potential chapter authors were obtained from personnel in the national offices of each of the relevant accrediting agencies. Contributing authors were asked to prepare a chapter addressing three questions: (1) What is the current state of the art of assessment in your discipline? (2) How has your accrediting body influenced campus assessment efforts in your field? (3) What has been learned thus far from assessment, and what concerns still need to be addressed? Chapter authors emphasized that many stakeholders in professional disciplines—practicing nurses, teachers, engineers—are vitally concerned about the knowledge and skills of future practitioners in their field and thus play a powerful role in stimulating faculty involvement in assessment. Similar types of assessment methods—written work, oral presentations, rubrics—are used across disciplines; and since faculty in accredited disciplines are accustomed to experiencing external influences on their work, they are less likely than faculty in fields not subject to accreditation to perceive assessment requirements as threats to their academic freedom. In fact, faculty in accredited disciplines report that they appreciate the support for assessment they receive from their accrediting bodies, such as written materials, website resources, workshops, consultants, and space on national programs for presentations on good practice in assessment.

Competence-Based Initiatives. In 2001, Banta was a member of the National Postsecondary Education Cooperative (NPEC), a panel appointed to recommend and commission studies to assist the National Center for Education Statistics (NCES) in providing indicators of education status and trends. In this capacity she was asked to serve on a working group that would study competence-based initiatives across the country, defined as purposeful actions undertaken by colleges and universities to define, teach, and assess competence. She described that experience in a 2001 (May-June) column. Eight sites were selected for in-depth study involving interviews with campus administrators and site visits: King's College, Northwest Missouri State University, Sinclair Community College, Hagerstown Community College, Western Governors University, Community Colleges of Colorado, Oregon's Proficiency-Based Admission Standards System, and Ford Motor Company. Principles of good practice in initiating and maintaining effective competence-based initiatives emerging from this study include (1) having at least one senior administrator who fosters a culture open to change and offers incentives to participants in innovation,

(2) involving appropriate stakeholders in identifying and defining competence with sufficient clarity that it can be assessed, (3) employing multiple measures in assessing competence, (4) including the competence-based initiative in a larger institution-wide plan, (5) reporting assessment findings with clarity for stakeholders, and (6) ensuring that assessment findings are used continuously to improve teaching and learning (U.S. Department of Education, NCES, 2001).

In 2002 (January-February), Banta summarized three articles on the use of competence-based education (CBE) published in that issue of *AU*. At each of three universities employers worked with faculty to identify types and levels of knowledge and skills needed for success in their respective fields. Then faculty changed curriculum, methods of instruction, and student services such as career planning and placement to ensure that program graduates would develop the specified knowledge and skills. Benefits of CBE include the ability of students to advance through a program of study at their own pace, amassing a record of skills they have mastered as opposed to merely the time they have spent in class (number of credit hours). Students have a better idea of what is expected of them, and faculty have clear guidance for teaching in the statements of specific learning outcomes for students developed for each discipline. Transfer from one institution to another is easier when students bring with them examples of what they know and can do, and employers obtain better evidence of the skills potential employees possess. CBE requires faculty collaboration and strong administrative support for faculty development and policy changes, such as moving from transcripts that report credit hours earned to ones that actually provide evidence of learning.

Assessment Scholarship. By 2002, the term 'scholarship of assessment' had emerged, and Banta and Associates (2002) contributed fifteen chapters to *Building a Scholarship of Assessment*. Ewell wrote the first chapter of this volume, outlining the diverse intellectual roots of assessment and the evolution of its early history. Authors identified several well-established fields on which the scholarship of assessment can be based, including cognitive psychology, measurement, organizational development and change, and program evaluation. In her contribution to the book, Banta offered seventeen characteristics of effective assessment practice clustered in three phases: planning, implementing, and sustaining/improving the process. In her *Editor's Notes* (2002, March-April) she observed that assessment seemed stuck in a rut as individuals on campuses across the country simply repeat practices they have picked up from others at conferences or in their reading. She called for more assessment scholarship—systematic study of various approaches and their respective effects.

In 1997, Ewell chaired the design team for what became the National Survey of Student Engagement (NSSE) and shepherded that instrument through two pilots to its first national administration in 2000. He then chaired the National Advisory Board for the NSSE for a number of years and currently serves in this capacity for the related Community College Survey of Student Engagement (CCSSE). Development of the NSSE was originally funded by the Pew Charitable Trusts, which viewed it as a potential alternative to *U.S. News and World Report's* rankings of "America's Best Colleges." During the pilot period, however, it became clear that participating colleges and universities were opposed to this and wanted to keep NSSE data confidential. After the 2002 administration, however, a number of institutions agreed to disclose their NSSE results voluntarily, as Ewell reported in his column in 2002 (September-October). Since that time,

By 2002, the term 'scholarship of assessment' had emerged, and Banta and Associates (2002) contributed fifteen chapters to Building a Scholarship of Assessment. Ewell wrote the first chapter of this volume, outlining the diverse intellectual roots of assessment and the evolution of its early history.

NSSE results for public institutions have been featured in a number of state performance indicator systems and were included in the VSA.

Civic Engagement. The decade beginning in 2000 brought new emphasis to civic engagement—mutually beneficial collaboration of students, faculty, and staff with the community. In 2005 (March-April) Banta described some direct and indirect methods for assessing civic engagement at course, program, and campus levels. Data assembled at the campus level can provide a useful context for assessment at program and course levels. For instance, a central office may track participation in voluntary community service activities in terms of numbers of activities; of campus sites; and of student, faculty, and staff volunteers participating. An electronic civic engagement inventory may be helpful in this regard, particularly if data can be disaggregated to the program level. NSSE contains items on voting and other contributions to community welfare, as well as self-ratings of learning gains in various manifestations of civic responsibility experienced as a result of attending an institution. Program reviews can provide another perspective on the quality of civic engagement, especially if conducted by an external team that includes a community representative. This team member can assess the unit's visibility and level of involvement in the community and suggest additional avenues for civic engagement. In response to reviewers' recommendations, units may decide to appoint a community advisory board to guide civic participation on an on-going basis. At the course level, specific student learning outcomes associated with civic engagement can guide both faculty selection of learning opportunities and students' understanding of civic engagement. Electronic portfolios offer the benefit of recording students' activities and development over time. Classroom assessment techniques offer instructors a quick check on students' perceptions of their learning as well as suggestions for improving learning. Finally, items used in campus-wide surveys may be administered to classes in order to obtain unit results that can be compared to campus averages or, perhaps, to responses of students in other programs or divisions.

Academic Leadership. In 2005 (September-October), Banta and Lauryl Lefebvre interviewed eleven campus leaders: chief executives, chief academic officers, and one chief student affairs officer, at institutions well known for their good work in outcomes assessment. In response to a question about how assessment could benefit their institutions, these leaders asserted that broadening and deepening student learning is a primary goal of higher education, and assessment is a means to that end (Banta & Lefebvre, 2006). Through assessment, students can identify their own strengths and shortcomings and determine the best ways to accelerate their progress. Assessment findings can suggest to faculty the best ways to improve curriculum, instruction, and student services—and ultimately student development broadly conceived. Assessment of institutional effectiveness yields evidence of progress toward strategic goals that can be used both internally, to suggest improvements, and externally, to demonstrate to concerned stakeholders that students can apply their knowledge and skills to solve problems—that degrees have integrity.

Assessment of institutional effectiveness yields evidence of progress toward strategic goals that can be used both internally, to suggest improvements, and externally, to demonstrate to concerned stakeholders that students can apply their knowledge and skills to solve problems—that degrees have integrity.

Leaders' views on engaging stakeholders in assessment was the subject of a follow-up article based on the eleven interviews (Banta, 2005, November-December). Believing that finding ways to demonstrate and improve student learning can be a catalyst for positive change in an institution, these leaders set about creating a culture based on evidence, an environment in which decisions are guided by data, and a culture that values and rewards the collection and use of data to direct changes aimed at enhancing the student experience in college. The hope is that many faculty and staff immersed in such an environment will develop intrinsic motivation to engage in assessment. This motivation can be enhanced by providing a variety of direct support and developmental opportunities for good work in assessment such as release time, summer salary, on-campus workshops, travel to assessment-related conferences, and mini-grants. Offering expert assistance centrally for developing and administering questionnaires and focus groups, and analyzing the data and summarizing the findings, are additional incentives for encouraging colleagues to begin to use assessment in improving academic programs and student services.

The eleven assessment leaders Banta and Lefebvre interviewed in 2005 provided quite a range of responses to the question, "What assessment methods have proven most useful in providing direction for improvement on your campus?" The chief student affairs officer wanted his staff to go beyond simply describing programs to include the number of participants, their satisfaction, and even what they had learned. Two of the leaders brought in community members to act as external examiners for senior students, evaluating them using checklists of essential graduate attributes. One CAO valued three principal components of assessment: self-evaluation by faculty, external peer review, and assessment by customers such as students and employers. A CEO said that he had bestowed additional funds on units that had used data to implement significant improvements. At one institution faculty had developed their own tests of writing, math, and information literacy that yielded scores helpful to individual students and to faculty alike. For one other campus, nationally standardized instruments such as NSSE and the California Test of Critical Thinking had been similarly helpful. At another university, students' electronic portfolios provide the medium for analyzing change in student behavior over time. At two institutions the scholarship of teaching, learning, and assessment has appealed to research-oriented faculty, but also has provided the evidence to suggest important internal improvements. Clearly there is no assessment measure that is best for all disciplines or institutions. The key is to find through experimentation that set of measures that faculty and staff find credible as sources of evidence of student learning and institutional effectiveness.

Clearly there is no assessment measure that is best for all disciplines or institutions. The key is to find through experimentation that set of measures that faculty and staff find credible as sources of evidence of student learning and institutional effectiveness.

Newer Approaches

In 2007 (March-April), Banta argued for the development of alternative measures of learning to ameliorate the current press to assess using a standardized test for the purpose of comparing institutions. First she suggested indicators such as the availability of advising for individual students and percentage of students graduating in six years. If institutions must be compared, might we use scores on standardized licensing and certification exams in professional fields and encourage national associations in other fields, such as mathematics and psychology, to develop their own achievement measures? Another possibility would be to ask experts in each field to construct rubrics for evaluating student products and/or materials in electronic portfolios. In this connection, Banta described Tuning (European Commission, 2006), the effort beginning in 2000 to bring academics from many institutions together for the purpose of harmonizing diverse institutional curricula in order to facilitate transfer of credit and credentials throughout Europe.

In 2008 (November-December), Banta, with Elizabeth A. Jones and Karen E. Black, developed the content for the book *Designing Effective Assessment* (2009) and provided a preview of the book in *Assessment Update*. The book illustrates principles of good practice in assessment using examples from 146 institutions submitting information in response to an email request. Fully developed profiles from 49 institutions appear in the book organized in categories such as general education, faculty development, program review, community colleges, and graduate programs. The authors were somewhat surprised to see that only 43 percent of the campus profiles mentioned regional or disciplinary accreditation as a major force behind the assessment initiatives being descried. Just 8 percent of the profiles that were focused on general education, generic skills like writing, or institutional effectiveness, contained evidence that a standardized test of generic skills had been used, and in no case was that test the only direct measure of student competence in place. In each case test scores were supplemented by locally developed measures such as application of scoring rubrics to papers, projects, or capstone assignments. When asked about the impact of using assessment findings, authors of profiles mentioned broader participation in assessment activities, creation of positions for campus assessment leaders, and more use of technology in assessment. But only 6 percent of the authors identified increased student learning as an outcome of conducting assessment.

Early in 2009 (March-April), Banta speculated about the reasons for so little evidence of improvement of student learning after almost three decades of work in assessing outcomes.

Early in 2009 (March-April), Banta speculated about the reasons for so little evidence of improvement of student learning after almost three decades of work in assessing outcomes. First, 62 percent of the institutions submitting profiles for the book had been engaged in assessment for fewer than five years. Perhaps this just isn't long enough to become comfortable with findings, determine what to change, make changes, then re-administer instruments to see if change has occurred. Another possibility is that personnel changes among chief executives and chief academic officers, as well as among assessment leaders, often bring differing priorities and new initiatives that cause prior assessment initiatives to be modified or abandoned. Finally, if we value continuous improvement, conditions are constantly changing, and this makes it difficult to attribute any positive or negative movement to a specific set of conditions.

In 2011 (May-June), Banta referred to an essay she had written with Charles Blaich, director of the Center of Inquiry in the Liberal Arts at Wabash College. Their article, which appeared in *Change*, was entitled "Closing the Assessment Loop." Drawing on the concept of double-loop learning developed by Chris Argyris (1992), these authors argued not just for closing the loop by using assessment evidence to direct improvements, but also for initiating the assessment loop again to test whether the improvements undertaken had actually improved the desired outcome. In the *Change* article Blaich had asserted, "An effective assessment program should spend more time and money on using data than on gathering it…" In her *AU* column Banta worried that completing the second loop would cost at least as much as closing the first and called for more funding for assessment scholarship to determine which assessment-related activities actually produce significant improvements.

Encouraging Developments

In Banta's 2010 (September-October) column, she described the 1981 Kellogg Foundation project that introduced her to Ewell as a way to encourage the research-oriented faculty at the University of Tennessee, Knoxville to engage in some of the testing and survey work that would be necessary for UTK to collect initially 2 percent and later 5.45 percent of its budget for instruction from the state—about $6 million in the early years—that would be made available through Tennessee's new performance funding initiative. In fact, she was able to use a very small amount of money—$10,000 (supplemented quietly with modest additional funding from the chancellor)—to tell skeptical faculty colleagues that "we have an opportunity" to participate in "a project funded by the Kellogg Foundation" and engage them in creative approaches to the performance funding mandates. In 2010, Banta used this example in her "Editor's Notes" to demonstrate the importance to faculty of a grant from an external source in giving them the perception of new freedom to do outcomes assessment in their own way. For more than a decade between 1986 and 1998, the federal Fund for the Improvement of Postsecondary Education provided grants of up to three years in duration for colleges and universities to do assessment in their own way. But then FIPSE priorities shifted, and for the next decade the only way for faculty to feel encouraged in their extra efforts was for creative provosts and college deans to re-direct funds for course and curriculum design and professional development to outcomes assessment projects such as pilot tests of commercial tests and surveys or development of local instruments. Fortunately this began to change in 2007 with investments by FIPSE in the development of AAC&U's VALUE rubrics; by the Teagle Foundation in consortial assessment projects; and by the Lumina Foundation, Teagle, and the Carnegie Corporation of New York in NILOA. Lumina's support of Tuning USA convinced even some skeptics in the liberal arts that they were being supported in taking the first step in outcomes assessment—specifying student learning outcomes in the major field—in their own way!

For more than a decade between 1986 and 1998, the federal Fund for the Improvement of Post-secondary Education provided grants of up to three years in duration for colleges and universities to do assessment in their own way.

So much of what was written about outcomes assessment prior to 2009 was in the form of stories about the use of traditional tests, papers, and surveys on individual campuses. No wonder it felt like assessment was stuck in a rut, plowing the same ground over and over (Banta, 2009, July-August). But in December 2008, George Kuh of Indiana University and Stan Ikenberry from the University of Illinois convened the first meeting of a national advisory panel for the organization soon to be named the National Institute for Learning Outcomes Assessment. Ewell and Banta were invited to serve on that panel, and Ewell has a continuing relationship with NILOA as a Senior Scholar. In March 2009, NILOA personnel undertook a web-based survey to identify and describe campus assessment initiatives and began to commission occasional papers and brief case studies of promising practices. In April 2009, Lumina introduced its Tuning USA project, bringing together faculty representatives from public and private institutions in Indiana, Minnesota, and Utah to develop common learning outcomes for graduates in a variety of disciplines. Ewell reported on this initiative in his 2010 (January-February) column, but warned that it was an incomplete answer to the accountability challenge because it focused on individual academic disciplines rather than the kinds of "generic" outcomes like communication skills and quantitative reasoning for which policymakers and employers sought to hold the academy accountable.

In 2011 (May-June), Ewell argued that the demand for common student learning outcomes was being met more effectively by Lumina's new Degree Qualifications Profile (DQP), which mirrored the kinds of qualifications frameworks already established in European countries. The need to assess generic outcomes authentically, meanwhile, was also being addressed effectively by other organizations. In May 2009, the Association of American Colleges and Universities principals brought together academics from all over the country to use the VALUE rubrics to assess a sample of student work in electronic portfolios. In the hope of providing a method of authentic assessment both to guide improvement and to demonstrate accountability, the VALUE rubrics for Written Communication, Critical Thinking, and Information Literacy had been developed over the past two years by scores of faculty from a variety of institutions. By 2013 (January-February), Ewell was able to report that use of the VALUE rubrics was seen as an effective substitute for standardized testing in Massachusetts' statewide assessment initiative, and in 2015 (July-August), he reported use of this approach in common in nine states as part of an ongoing project sponsored by the SHEEO organization.

In a 2015 (July-August) column, Banta suggested that the age-old dilemma of how to engage faculty in outcomes assessment could be addressed by discovering processes that faculty truly value and linking assessment to those processes. For example, at one time or another in their careers faculty may be particularly interested in enhancing student learning, reviewing and revising a curriculum in general education or the major field, bringing in external consultants to conduct a program review, conducting the scholarship of teaching and learning, evaluating the effectiveness of a new initiative, participating in faculty development, or being recognized and/or promoted. Ways have been found on campuses across the country to connect outcomes assessment to these valued processes, and articles in the 2015 July-August issue of *AU* illustrate some of the methods. Faculty at Gonzaga University were attracted to assessment by their curiosity about the effectiveness of linking three courses in the core curriculum. Concerns about the reliability and validity of rubrics being used to assess learning outcomes in general education courses led faculty at Mount St. Joseph University to seek professional development related to writing more effective assignments. At the University of Arizona, a requirement has been instituted that compels departments to include in their self-studies for program review sections on outcomes assessment practices, findings, and uses of findings in undergraduate and graduate programs. Assessment professionals use a rubric to evaluate the assessment sections and offer workshops and individual consultations to assist faculty in improving their assessment approaches. At Franklin Pierce University, faculty gave undergraduate students an opportunity to apply their learning about survey research by designing and administering a questionnaire to gather data from psychology alumni for a psychology program review.

In 2013, Banta used the technique of appreciative inquiry (Hammond, 1996) to explore which aspects of being a member of the campus-wide assessment committee at IUPUI were considered of most value by the membership (Banta, 2014, March-April). She was surprised to learn that members valued most their involvement in preparing for reaffirmation of accreditation by the Higher Learning Commission (HLC), which had taken place more than a year earlier. They enjoyed hearing from each other at meetings about various assessment approaches and about the assessment research funded with mini-grants from the committee. Members said they were proud when they learned of the comment in the HLC visiting team's report, "Assessment is one of the University's strengths." Focusing on what works well about their organization put members at ease and in a frame of mind to participate actively in discussions. Participants' responses to the questions shaped by appreciative inquiry provided committee leaders with guidance for developing the content of future meetings.

By 2013 (January-February), Ewell was able to report that use of the VALUE rubrics was seen as an effective substitute for standardized testing in Massachusetts' statewide assessment initiative, and in 2015 (July-August) he reported use of this approach in common in nine states as part of an ongoing project sponsored by the State Higher Education Executive Officers organization.

Summing Up

Sustained observation of developments in quality assurance in Europe and outcomes assessment in the U.S. over the last three decades invites analysis of similarities and differences. Initial reactions of academics on both sides of the Atlantic to the accountability demands of the 1980s and '90s from national education ministries in Europe or state legislatures and coordinating boards in the U.S. were predictably negative. But by 2000 programs for annual assessment conferences in these two areas of the world revealed that resentment of edicts from external actors had begun to give way to more constructive reactions—finding the best campus-specific methods for assessing and improving student learning and institutional effectiveness.

Since education is a responsibility of the states, the first accountability mandates in this country were issued by state legislatures or higher education governing and coordinating boards. But with the passage of the 1992 Higher Education Act, the federal government exerted its influence by requiring that regional and disciplinary accrediting organizations include in their standards the need for identifying, and assessing, learning related to student learning outcomes. By 2009, a survey of chief academic officers conducted by NILOA staff (Kuh & Ikenberry, 2009) (also see http://www.learningoutcomesassessment.org/MoreThanYouThink.htm) indicated that accreditation was the most important external influence on assessment practice. In addition to spurring engagement in assessment on campuses, accreditors have provided important resources for that engagement in the form of conferences, papers and articles, websites, and consultants.

Between 2001 and 2008, the National Center for Public Policy in Higher Education issued 50-state *Measuring Up* report cards that attempted to compare states on collegiate preparation, participation, affordability, completion, and benefits (Callan & Finney, 2002). The measure receiving the most attention was learning, which was given the grade of Incomplete in 2001. The desire to erase that Incomplete mark gave rise to several subsequent multi-state projects, including a five-state demonstration project (Miller & Ewell, 2005) and the Multi-State Collaborative (MSC), a joint project of AAC&U and SHEEO (http://www.sheeo.org/projects/msc-multi-state-collaborative-advance-learning-outcomes-assessment), which were designed to produce better direct measures of student achievement.

U.S. Secretary of Education Margaret Spellings had a major impact on outcomes assessment between 2005 and 2012. The Commission on the Future of Higher Education that she appointed issued a report in 2006 (USDOE, 2006) calling, among other things, for the use of standardized tests of generic skills like writing and critical thinking to compare the quality of colleges and universities. Higher education leaders mobilized to develop several responsive initiatives, including the Voluntary System of Accountability (VSA). The VSA called, among other things, for institutions to test first-year and graduating students using one of three specified standardized tests of generic skills. By 2016, however, a national survey commissioned by AAC&U demonstrated that the flurry of interest in standardized tests of generic skills stimulated in part by the VSA had subsided (Fain, 2016). Chief academic officers responding to the survey reported that faculty at their institutions were much more likely to use rubrics applied to authentic examples of student work, capstone projects, student self-reports, and locally developed exams to assess learning and to provide direction for improvements in curriculum and pedagogy.

By 2000, programs for annual assessment conferences in these two areas of the world [Europe and the United States] revealed that resentment of edicts from external actors had begun to give way to more constructive reactions—finding the best campus-specific methods for assessing and improving student learning and institutional effectiveness.

During the early history of outcomes assessment in the 1980s and '90s, scholars lamented the fact that the literature of the field lacked visibility. It consisted primarily of unpublished campus-based reports. Increased use of websites in the new millennium has made even these reports more accessible. Meanwhile several journals focused on outcomes assessment have found publishers, as have a growing number of books. In 2002, the first book on the scholarship of assessment appeared (Banta & Associates, 2002) linking this new field to the more established bodies of theory and practice of program evaluation, cognitive development, psychometrics, and organizational development and change. A growing number of states have informal or formal organizations of assessment professionals who gather regularly for exchange of ideas, and there are two annual national assessment conferences. The assessment literature is still populated largely by single-institution reports of successful practice—or missteps. But principals at NILOA are leading the way toward more robust assessment scholarship.

More scholars are contributing to the knowledge base, and most disciplinary associations now have their own assessment literature, tracks at annual conferences, and even specialized workshops and conferences. More assessment tools and other resources are being developed daily, with recent advances in e-portfolios, rubrics, and digital badges.

The range of assessment strategies hasn't changed markedly since 2000. Faculty still rely on syllabus analysis; questions about writing assignments; rubrics applied to group work; portfolio reviews; and surveys for students, faculty, and alumni (Banta, 2000, May-June). However, the application of technology to each of these processes has simplified its use and made it much easier to aggregate and analyze the data it produces. Interest in competence-based education, which relies importantly on effective assessment, has ebbed and flowed over the years, but seems to be enjoying increased attention currently (see http://www.ed.gov/oii-news/competency-based-learning-or-personalized-learning).

Linda Suskie and Barbara Walvoord joined Banta as co-authors of her column (2015, January-February) following a joint presentation by the three at the 2014 Assessment Institute in Indianapolis. "The Three Tenors," as they were called in the Institute program, looked back over their professional experiences in the past year and made some educated guesses about the future of outcomes assessment. First, the three agreed that as pressures to demonstrate accountability continue to increase, there is far more acceptance of assessment now than ever before; more and more faculty and staff are recognizing that assessment is essential and are seeking ways to do it more effectively. More scholars are contributing to the knowledge base, and most disciplinary associations now have their own assessment literature, tracks at annual conferences, and even specialized workshops and conferences. More assessment tools and other resources are being developed daily, with recent advances in e-portfolios, rubrics, and digital badges. More vendors are offering software to help institutions collect, store, and analyze assessment information, and these developments will just intensify as we seek massive quantities of data for analytics to individualize learning. As we look for ways to customize high-impact practices to fit our own settings, accreditors are being encouraged by the U.S. Department of Education and Congressional critics to require more uniform standards. The need for more commonality may be addressed as cross-disciplinary and cross-institutional collaboration produces better rubrics and pushes expectations for student achievement ever higher. As all of these changes unfold, the need for continuous professional development for faculty and staff becomes critical.

Epilogue

As Banta and Ewell head into the twilight of their assessment careers and Cogswell is launched in hers, we reflect on what we have learned and what may lie ahead.

Although much of the history of outcomes assessment that we have reported has involved responding to external threats (e.g., several calls for national testing using standardized tests so that institutions can be compared) or mandates (e.g., state performance funding initiatives), there are many signs that campus faculty and administrators are ready to assume responsibility for designing their own improvement-oriented approaches using authentic measures.

Developing a culture that promotes data-based decision-making requires strong leadership from presidents, provosts, and well-informed assessment professionals. There is encouraging evidence that this kind of leadership is emerging within more and more institutions.

Evolving technologies will make the collection, aggregation, and analysis of assessment data easier. And as the science of analytics matures, confidence in using outcomes information to predict the consequences of various routes to improvement will grow. Finally, these developments will increase and strengthen assessment scholarship, which will, in turn, lead to more competent and confident assessment practice. We imagine a bright future for this field!

References

Adelman, C. (2008). *Learning accountability from Bologna: A higher education policy primer*. Washington, D.C.: Institute for Higher Education Policy. Retrieved from http://www.ihep.org/research/publications/learning-accountability-bologna-higher-education-policy-primer

Argyris, C. (1992). *On organizational learning*. Cambridge, MA: Blackwell.

Banta, T. W. (2000, January-February). Editor's notes: That second look. *Assessment Update, 12*(1), 3, 14, 16

Banta, T. W. (2000, May-June). Doing assessment our way. *Assessment Update, 12*(3), 3, 7

Banta, T. W. (2000, November-December). Seeing ourselves as others see us. *Assessment Update, 12*(6), 3, 13, 16.

Banta, T. W. (2001, March-April). Presenting survey data to encourage action. *Assessment Update, 13*(2), 3, 5, 11.

Banta, T. W. (2001, May-June). On competence. *Assessment Update, 13*(3), 3, 5, 10.

Banta, T. W. (2001, September-October). Confirming an empirical finding about external influences on assessment. *Assessment Update, 13*(5), 3, 13.

Banta, T. W. (2001, November-December). Is external encouragement for assessment essential? *Assessment Update, 13*(6), 3, 10, 11.

Banta, T. W. (2002, January-February). Student competence as the basis for designing curricula, instruction, and assessment. *Assessment Update, 14*(1), 3, 15.

Banta, T. W. (2002, March-April). Toward a scholarship of assessment. *Assessment Update, 14*(2), 3, 13.

Banta, T. W. (2002, November-December). Two conferences, two systems -- Similar perspectives? *Assessment Update, 14*(6), 3-4.

Banta, T. W. (2005, March-April). Assessing civic engagement. *Assessment Update, 17*(2), 3, 11.

Banta, T. W. (2005, September-October). What draws campus leaders to embrace outcomes assessment? *Assessment Update, 17*(5), 3, 14-15.

Banta, T. W. (2005, November-December). Leader's views on engaging stakeholders in assessment. *Assessment Update, 17*(6), 3, 15-16.

Banta, T. W. (2006, July-August). Reliving the history of large-scale assessment in higher education. *Assessment Update, 18*(4), 3-4, 15.

Banta, T. W. (2007, March-April). If we must compare…. *Assessment Update, 19*(2), 3-4.

Banta, T. W. (2008, March-April). Trying to clothe the emperor. *Assessment Update, 20*(2), 3-4, 15-16.

Banta, T. W. (2008, July-August). Sadly rubrics are not for everyone. *Assessment Update, 20*(4), 3.

Banta, T. W. (2008, November-December). Profiles of good practice in assessment. *Assessment Update, 20*(6), 3-4.

Banta, T. W. (2009, March-April). Demonstrating the impact of change based on assessment findings. *Assessment Update, 21*(2), 3-4.

Banta, T. W. (2009, July-August). New opportunities for pushing the wheel forward. *Assessment Update, 21*(4), 3-4,16.

Banta, T. W. (2009, November-December). Sour notes from Europe. *Assessment Update, 21*(6), 3, 7.

Banta, T. W. (2010, January-February). Are there measures for improvement and accountability? *Assessment Update, 22*(1), 3, 10.

Banta, T. W. (2010, July-August). New resource on student test-taking motivation. *Assessment Update, 22*(4), 3-4.

Banta, T. W. (2010, September-October). A little extra funding goes a long way. *Assessment Update, 22*(5), 3-4.

Banta, T. W. (2011, May-June). Double loop learning in assessment. *Assessment Update, 23*(3), 3, 15.

Banta, T. W. (2012, November-December). Addressing a question of credibility. *Assessment Update, 24*(6), 3-4, 15.

Banta, T. W. (2014, March-April). Accentuating the positive in our work. *Assessment Update, 26*(2), 3-4.

Banta, T. W. (2015, July-August). Linking outcomes assessment with valued campus processes. *Assessment Update, 27*(4), 3, 16.

Banta, T. W. & Associates. (2002). *Building a scholarship of assessment*. San Francisco, CA: Jossey-Bass.

Banta, T. W. & Blaich, C. F. (2011). Closing the assessment loop. *Change: The Magazine of Higher Learning, 43*(1), 22-27.

Banta, T. W., Jones, E. A., & Black, K. E. (2009). *Designing effective assessment: Principles and profiles of good practice*. San Francisco, CA: Jossey-Bass.

Banta, T. W. & Lefebvre, L. A. (2006). Leading change through assessment. *Effective practices for academic leaders, 1*(4), 1-16.

Banta, T. W., & Palomba, C. A. (2015). *Assessment essentials: Planning, implementing, and improving assessment in higher education. 2nd Ed.* San Francisco, CA: Jossey-Bass.

Banta, T. W., Suskie, L., & Walvoord, B. E. (2015, January-February). "Three assessment tenors look back and to the future." *Assessment Update, 27*(1), 3, 14-15.

Baum, S. (2014). Higher education earnings premium: Value, variation, and trends. *The Urban Institute*, 12 pp. Washington, D.C.: The Urban Institute.

Cahalan, M., & Perna, L. (2015). *Indicators of higher education equity in the United States: 45 year trend report*. Pell Institute for the Study of Opportunity in Higher Education. Philadelphia, PA: University of Pennsylvania.

Callan, P. M. & Finney, J. E. (2002). Assessing Education Capital: An Imperative for Policy. *Change, 34*(4), 24-31. Retrieved from http://www.jstor.org/stable/40177908

Coates, H. (2016). Assessing student learning outcomes internationally: Insights and frontiers. *Assessment & Evaluation in Higher Education*. Routledge, Taylor & Francis Group. Retrieved from http://dx.doi.org/10.1080/02602938.2016.1160273

Emrey-Aaras, M. (2014). *Higher education state funding trends and policies on affordability*. Report to the Chairman, Committee on Health, Education, Labor, and Pensions, United States Senate. GAO-15-151. Washington, D.C.: U.S. Government Accountability Office.

European Commission. (2006). *Tuning educational structures in Europe*. Retrieved from http://www.unideusto.org/tuningeu/

Ewell, P. T. (2000, January-February). From the states: "Statewide testing in Utah." *Assessment Update, 12*(1), 8-9, 16.

Ewell, P. T. (2000, September-October). Accountability with a vengeance: New mandates in Colorado. *Assessment Update, 12*(5), 14-16.

Ewell, P. T. (2001, January-February). Grading the states. *Assessment Update, 13*(1), 12-13, 16.

Ewell, P. T. (2001, May-June). Auditing assessment: The Teacher Education Accreditation Council. *Assessment Update, 13*(3), 8-10.

Ewell, P. T. (2002, January-February). Taking stock of accreditation reform. *Assessment Update, 14*(1), 12-14.

Ewell, P. T. (2002, September, October). Going public: NSSE encounters U.S. News. *Assessment Update, 14*(5), 11-13.

Ewell, P. T. (2003, January-February). Parallel lives: Recent developments for institutional review in England. *Assessment Update, 15*(1), 12-13, 15.

Ewell, P. T. (2003, May-June). Going for broke: The National Forum on College-Level Learning's multistate demonstration project. *Assessment Update, 15*(3), 8-9, 15.

Ewell, P. T. (2004, November-December). Tomorrow the world: Learning outcomes and the Bologna Process. *Assessment Update, 16*(6), 11-13.

Ewell, P. T. (2005, May-June). Show me: Missouri's experience with the Collegiate Learning Assessment. *Assessment Update, 17*(3), 15-16.

Ewell, P. T. (2005, September-October). Looking back on "The Year of Accountability." *Assessment Update, 17*(5), 9-11.

Ewell, P. T. (2006, May-June). Smoke signals from the hills of Texas. *Assessment Update, 18*(3), 12-13, 15.

Ewell, P. T. (2007, January-February). Accreditation in the hot seat. *Assessment Update, 19*(1), 11-13.

Ewell, P. T. (2008, July-August). Tomorrow the world: A higher education assessment initiative from the Organisation for Economic Cooperation and Development. *Assessment Update, 20*(4), 12-14.

Ewell, P. T. (2009, January-February). Rising to the occasion: The alliance for New Leadership for Student Learning and Accountability. *Assessment Update, 21*(1), 13-14.

Ewell, P. T. (2009, May-June). Well, um, we actually have a national system for tracking student progress. *Assessment Update, 21*(3), 12-13.

Ewell, P. T. (2009, September-October). "Shovel-Ready" data: The stimulus package and state longitudinal data systems. *Assessment Update, 21*(5), 11-12.

Ewell, P. T. (2010, January-February). Imitation as art: Tuning USA and higher education accountability. *Assessment Update, 22*(1), 13-14.

Ewell, P. T. (2010, September-October). The Voluntary Framework of Accountability. *Assessment Update, 22*(5), 11-12.

Ewell, P. T. (2011, May-June). Aligning academic standards: The Lumina degree profile. *Assessment Update, 23*(3), 11-12.

Ewell, P. T. (2011, September-October). Regional accreditation redux. *Assessment Update, 23*(5), 11-12, 16.

Ewell, P. T. (2012, July-August). Learning and accreditation: The elites push back. *Assessment Update, 24*(4), 10-11.

Ewell, P. T. (2013, January-February). Having it both ways: Massachusetts's approach to statewide assessment. *Assessment Update, 25*(1), 10-11.

Ewell, P. T. (2013, May-June). Assessment on steroids: Columbia's national testing program. *Assessment Update, 25*(3), 10-11.

Ewell, P. T. (2015, March-April). What happens when quality assurance crosses borders? *Assessment Update, 27*(2), 10-11.

Ewell, P. T. (2015, July-August). Multistate redux: Two initiatives expand in scope? *Assessment Update, 27*(4), 10-11.

Fain, P. (2016). Survey finds increased use of learning outcomes measures, but decline in standardized tests. *Inside Higher Ed.* Downloaded February 2016 from https://www.insidehighered.com/news/2016/02/17/survey-finds-increased-use-learning-outcomes-measures-decline-standardized-tests?width=775&height=500&iframe=true.

Gaston, P. L. (2010). *The challenge of Bologna: What United States higher education has to learn from Europe, and why it matters that we learn it.* Sterling, VA: Stylus.

Hammond, S. A. (1996). *The thin book of appreciative inquiry*. Bend, OR: Thin Book.

Harvey, L. (2002). The end of quality? *Quality in Higher Education, 8*(1), 5-22.

Jankowski, N. A., Ikenberry, S. O., Kinzie, J., Kuh, G. D., Shenoy, G. F., & Baker, G. R. (2012). *Transparency and accountability: An evaluation of the VSA College Portrait pilot.* Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

Klein, S., Liu, O. L., & Sconing, J. (2009). Test validity report. Washington, D.C.: U.S. Department of Education, Fund for the Improvement of Postsecondary Education.

Kuh, G., & Ikenberry, S. (2009). *More than you think, less than we need: Learning outcomes assessment in American higher education*. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

Medland, E. (2016). Assessment in higher education: Drivers, barriers and directions of change in the UK. *Assessment & Evaluation in Higher Education*. Routledge, Taylor & Francis Group. Retrieved from http://dx.doi.org/10.1080/02602938.2014.982072.

Miller, M. A. & Ewell, P. T. (2005). *Measuring up on college-level learning*. San Jose, CA: The National Center for Public Policy in Higher Education.

National Center for Public Policy and Higher Education. (2000). *Measuring up 2000: The state-by-state report card for higher education*. San Jose, CA: The National Center for Public Policy in Higher Education.

Palomba, C. A. & Banta, T. W. (2001). *Assessing student competence in accredited disciplines: Pioneering approaches to assessment in higher education*. Sterling, VA: Stylus Publishing, LLC.

Paris, D. (2011). *Catalyst for change: The CIC/CLA Consortium.* Washington, D.C.: Council of Independent Colleges.

Patton, M. Q. (2002). *Qualitative research and evaluation methods*. (3rd Ed.). Thousand Oaks, CA: Sage Publications.

Peterson, M. W., Einarson, M. K., Augustine, C. H., & Vaughn, D. S. (1999). *Institutional support for student assessment: Methodology and results of a national survey*. Stanford, CA: Stanford University, National Center for Postsecondary Improvement.

Pereira, D., Flores, M. A., & Niklasson, L. (2015). Assessment revisited: A review of research in assessment and evaluation in higher education. *Assessment and Evaluation in Higher Education.* Retrieved from http://dx.doi.org/10.1080/02602938.2015.1055233.

Rhodes, T. (2009). The VALUE Project overview. *Peer Review, 11*(1), 4-7.

State Higher Education Executive Officers Association. (2012). *MSC: A multi-state collaborative to advance learning outcomes assessment*. State Higher Education Executive Officers Association. Retrieved from http://www.sheeo.org/projects/msc-multi-state-collaborative-advance-learning-outcomes-assessment

Sundre, D. L. & Colleagues. (2009). *Journal of General Education, 58*(3).

U. S. Department of Education. (2006). *A test of leadership: Charting the future of U.S. higher education*. Washington, D.C.: U. S. Department of Education.

U. S. Department of Education, National Center for Education Statistics. (2001). *Defining and assessing learning: Exploring competence-based initiatives*. NCES 2001-11, prepared by Jones, E. A., Voorhees, R. A., & Paulson, K. Washington, D.C.: U.S. Government Printing Office.

NILOA National Advisory Panel

Ex-Officio Members

NILOA Mission

NILOA's primary objective is to discover and disseminate ways that academic programs and institutions can productively use assessment data internally to inform and strengthen undergraduate education, and externally to communicate with policy makers, families and other stakeholders.

NILOA Occasional Paper Series

NILOA Occasional Papers are commissioned to examine contemporary issues that will inform the academic community of the current state-of-the art of assessing learning outcomes in American higher education. The authors are asked to write for a general audience in order to provide comprehensive, accurate information about how institutions and other organizations can become more proficient at assessing and reporting student learning outcomes for the purposes of improving student learning and responsibly fulfilling expectations for transparency and accountability to policy makers and other external audiences.

Comments and questions about this paper should be sent to
niloa@education.illinois.edu.

About NILOA

- The National Institute for Learning Outcomes Assessment (NILOA) was established in December 2008.
- NILOA is co-located at the University of Illinois and Indiana University.
- The NILOA website contains free assessment resources and can be found at http://www.learningoutcomesassessment.org/.
- The NILOA research team has scanned institutional websites, surveyed chief academic officers, and commissioned a series of occasional papers.
- One of the co-principal NILOA investigators, George Kuh, founded the National Survey for Student Engagement (NSSE).
- The other co-principal investigator for NILOA, Stanley Ikenberry, was president of the University of Illinois from 1979 to 1995 and of the American Council of Education from 1996 to 2001.

NILOA Staff

NATIONAL INSTITUTE FOR LEARNING OUTCOMES ASSESSMENT

Stanley Ikenberry, *Co-Principal Investigator*

George Kuh, *Founding Director, Senior Scholar, and Co-Principal Investigator*

Natasha Jankowski, *Director*

Gianina Baker, *Assistant Director*

Katie Schultz, *Project Manager*

Peter Ewell, *Senior Scholar*

Pat Hutchings, *Senior Scholar*

Jillian Kinzie, *Senior Scholar*

Paul Lingenfelter, *Senior Scholar*

David Marshall, *Senior Scholar*

Carrie Allen, *Research Analyst*

Erick Montenegro, *Research Analyst*

P.S. Myers, *Research Analyst*

Verna F. Orr, *Research Analyst*

Anthony B. Sullers, Jr., *Research Analyst*

Emily Teitelbaum, *Research Analyst*

Terry Vaughan III, *Research Analyst*

NILOA Sponsors

Lumina Foundation for Education

University of Illinois, College of Education

Produced by Creative Services | Public Affairs at the University of Illinois for NILOA. 10.032

ucate innovation success ingenuity intellect curiosity challenge create achievement knowledge accountability connection self-reflection educate ac
derstand communicate curiosity challenge create achievement connection self-reflection understand communicate listen learn access quality actio
tion understand communicate listen learn action understand communicate listen learn access quality innovation success ingenuity intellect curiosi
allenge knowledge accountability connection access quality self-reflection curiosity challenge create achievement learn access quality innovation s
genuity self-reflection educate action understand intellect knowledge accountability connection self-reflection educate action understand knowledg
countability connection self-reflection educate action understand communicate listen learn access quality innovation success ingenuity intellect cu
allenge connection knowledge accountability connection self-reflection educate action understand communicate listen learn access quality innovat
ccess ingenuity challenge create achievement connection self-reflection educate action understand connection self-reflection understand communi
ten learn access quality action create achievement connection self-reflection educate action understand communicate listen learn access quality in
ccess educate action communicate listen learn access quality action educate action understand communicate educate innovation success self-refle
owledge accountability communicate listen learn achievement connection self-reflection educate action understand communicate listen learn acce
novation success ingenuity intellect access quality innovation success self-reflection curiosity challenge create achievement connection self-reflect
derstand educate action understand communicate listen learn action understand communicate listen learn access quality innovation success inger
riosity challenge create achievement connection self-reflection understand communicate listen learn access quality action create achievement con
lf-reflection educate action understand communicate listen learn access quality innovation success educate action communicate listen learn acces
ality action educate action understand create achievement connection self-reflection understand communicate listen learn access quality action cr
hievement connection self-reflection educate action understand communicate listen communicate educate innovation success self-reflection knowl
countability connection self-reflection educate action understand communicate listen learn access quality innovation ingenuity intellect connection
flection understand communicate listen learn access quality action create achievement connection self-reflection educate action understand comm
ten learn access quality innovation success educate action communicate listen learn access quality action educate action understand communicate

National Institute for Learning Outcomes Assessment

For more information, please contact:

National Institute for Learning Outcomes Assessment (NILOA)
University of Illinois at Urbana-Champaign
360 Education Building
Champaign, IL 61820

learningoutcomesassessment.org
niloa@education.illinois.edu
Phone: 217.244.2155