

A Historical Overview of Assessment: 1980s-2000s

January 2019

Excerpted from Peter T. Ewell and Tammie Cumming Chapter 1, *Enhancing Assessment in Higher Education*, Stylus (2017)

While no one has officially dated the birth of the “assessment movement” in higher education, it is probably safe to propose the First National Conference on Assessment in Higher Education held in Columbia, SC in the fall of 1985. Co sponsored by the National Institute of Education (NIE) and the American Association for Higher Education (AAHE), the origins of this conference vividly illustrate the conflicting political and intellectual traditions that have been with the field ever since. The proximate stimulus for the conference was a report called *Involvement in Learning* (NIE, 1984). Three main recommendations formed its centerpiece, strongly informed by research in the student learning tradition. In brief, they were that higher levels of student achievement could be promoted by establishing high expectations for students, by involving them in active learning environments, and by providing them with prompt and useful feedback. But the report also observed that colleges and universities as institutions could “learn” from feedback on their own performances and that appropriate research tools were now available for them to do so.

This observation might have been overlooked were it not consistent with other voices. One set came from within the academy and focused on curriculum reform, especially in general education. Symbolized by other prominent reports in 1984-85 like *Integrity in the College Curriculum* (AAC, 1985) and *To Reclaim a Legacy* (Bennett, 1984), their central argument was the need for coherent curricular experiences which could best be shaped by ongoing monitoring of student learning and development. From the outset in these discussions, the assessment of learning was presented as a form of “scholarship.” Faculties ought to be willing to engage in assessment as an integral part of their everyday work. A concomitant enlightened, but unexamined, assumption was that the tools of social science and educational measurement, deployed appropriately, could be adapted by all disciplines to further this process of ongoing inquiry and improvement.

A second set of voices arose simultaneously outside the academy, consisting largely of state-based calls for greater accountability. In part, these calls were a byproduct of the far more visible attention being paid to K-12 education, symbolized by the U.S. Department of Education’s 1983 report *A Nation at Risk* (U.S. Department of Education, 1983). In part, it stemmed from a renewed activism by governors and legislatures, based on their growing recognition that postsecondary education was a powerful engine for economic and workforce development. Both themes were apparent in yet another national report issued by the National Governors Association in 1986, revealingly titled *Time for Results* (NGA, 1986). As it was being issued, states like Colorado and South Carolina adopted assessment mandates requiring public colleges and universities to examine learning outcomes and report what they found. By 1987 when the first stock-taking of this growing policy trend occurred about a

dozen states had similar mandates. By 1989, this number had grown to more than half the states (Ewell, Finney, and Lenth, 1990).

Phase I: The Founding Period. Given this history, the motives of those attending the first national assessment conference were understandably mixed. Clear to all were the facts that they had few available tools, only a spotty literature of practice, and virtually no common intellectual foundation on which to build. Filling these yawning gaps in the period 1985–88 was a first and urgent task. In beginning this task, practitioners faced three major challenges:

- **Definitions.** One immediate problem was that the term “assessment” meant different things to different people. Initially, at least three meanings and their associated traditions of use had therefore to be sorted out. The most established had its roots in the mastery-learning tradition, where “assessment” referred to the processes used to determine an individual’s mastery of complex abilities, generally through observed performance. Adherents of this tradition emphasized development over time and continuous feedback on individual performance—symbolized by the etymological roots of the word “assessment” in the Latin *ad + sedere*, “to sit beside” (Loacker, Cromwell, and O’Brien, 1986). A far different meaning emerged from K-12 practice, where the term described large-scale testing programs like the federally-funded National Assessment of Educational Progress (NAEP) and a growing array of state-based K-12 examination programs. The primary object of such “large-scale assessment” was not to examine individual learning but rather to benchmark school and district performance in the name of accountability. Its central tools were standardized examinations founded on well-established psychometric principles, designed to produce summary performance statistics quickly and efficiently. Yet a third tradition of use defined “assessment” as a special kind of program evaluation, whose purpose was to gather evidence to improve curricula and pedagogy. Like large-scale assessment, this tradition focused on determining aggregate not individual performance, employing a range of methods including examinations, portfolios and student work samples, surveys of student and alumni experiences, and direct observations of student and faculty behaviors. An emphasis on improvement, moreover, meant that assessment was as much about *using* the resulting information as about psychometric standards.
- **Instruments.** A second challenge faced by early assessment practitioners was to quickly identify credible and useful ways to gather evidence of student learning. Virtually all the available instruments were designed to do something else. Ranging from admissions tests like the ACT Assessment and the Graduate Record Examinations, through professional registry and licensure examinations, to examinations designed to award equivalent credit, none of the available testing alternatives were really appropriate for program evaluation. Their content only approximated the domain of any given institution’s curriculum and the results they produced usually provided insufficient detail to support improvement. But this did not prevent large numbers of institutions—especially those facing state mandates—from deploying them.

In the period 1986–89, the major testing organizations quickly filled the instrument gap with a range of new purpose-built group-level examinations aimed at program evaluation—all based on existing prototypes. Among the most prominent were the ACT Collegiate Assessment of Academic Proficiency (CAAP), the ETS Academic Profile, and a range of ETS Major Field Achievement Tests (MFAT). Student surveys provided another readily-available set of data-gathering tools, especially when they contained items on self-reported gain. While many institutions designed and administered their own surveys, published instruments were readily available including the CIRP Freshman and follow-up surveys, the College Student Experiences Questionnaire (CSEQ) and a range of questionnaires produced by organizations like ACT and NCHEMS.

An emphasis on improvement, moreover, meant that assessment was as much about using the resulting information as about psychometric standards.

- **Implementation.** A third challenge faced by early assessment practitioners was lack of institution-level experience about how to carry out such an initiative. One question here was cost and, as a result, some of the first “how to” publications addressed financial issues. Others considered the organizational questions involved in establishing an assessment program. But absent real exemplars, the guidance provided by such publications was at best rudimentary. Enormous early reliance was therefore placed on the lessons that could be learned from the few documented cases available. Three such “early adopters” had considerable influence. The first was Alverno College, whose “abilities-based” curriculum designed around performance assessments of every student was both inspiring and daunting. A second was Northeast Missouri (now Truman) State University, which since 1973 had employed a range of nationally-normed examinations to help establish the “integrity” of its degrees. A third was the University of Tennessee Knoxville, which under the stimulus of Tennessee’s performance funding scheme became the first major public university to develop a comprehensive multi-method system of program assessment. These three cases were very different and provided a wide range of potential models.

In the late 1980s, a “second wave” of documented cases emerged, including (among others) James Madison University, Kean College, Kings College, Ball State University, Miami-Dade Community College, and Sinclair Community College—many of which were responding to new state mandates. To a field hungry for concrete information, these examples were extremely welcome. More subtly, they helped define a “standard” approach to implementing a campus-level program, which was widely imitated.

This founding period thus generated some enduring lines for assessment’s later development. One addressed concept development and building a coherent language. The purpose here was largely to stake out the territory—though much of this early literature was frankly hortatory, intended to persuade institutions to get started. A second line of work concerned tools and techniques. A third strand comprised case studies of implementation, supplemented by a growing body of work addressing practical matters like organizational structures and faculty involvement. Finally, accountability remained a distinct topic for comment and investigation, looking primarily at state policy, but shifting later toward accreditation.

Phase II: Consolidating Assessment. By 1990, predictions that “assessment would quickly go away” seemed illusory. Most states had assessment mandates, though these varied in both substance and in the vigor with which they were enforced. Accrediting bodies, meanwhile, had grown in influence, in many cases replacing states as the primary external stimulus for institutional interest in assessment. Reflecting this shift, more and more private institutions established assessment programs. These external stimuli were largely responsible for a steady upward trend in the number of institutions reporting “involvement” with assessment. By 1993, this proportion had risen to 98%. Clearly, at least for administrators, assessment was now in the mainstream. But “entering the mainstream” meant more than just widespread reported use. It also implied consolidation of assessment’s position as a distinct and recognizable practice.

- **An Emerging Modal Type.** As institutions scrambled to “implement assessment,” it was probably inevitable that they evolved similar approaches. And despite repeated admonitions to ground assessment in each institution’s distinctive mission and student clientele, they approached the task of implementation in very similar ways. As a first step, most formed committees to plan and oversee the work. Following widespread recommendations about the importance of faculty involvement, most comprised faculty drawn from multiple disciplines. But partly because the press to implement was so great, assessment committees rarely became a permanent feature of governance or of academic administration.

The clear first task of these committees, moreover, was to develop an “assessment plan.” Often, such a product was explicitly required by an accreditor or state authority. Equally often, it was recommended by a consultant or by the burgeoning “how to” literature of practice

Assessment was now in the mainstream. But “entering the mainstream” meant more than just widespread reported use. It also implied consolidation of assessment’s position as a distinct and recognizable practice.

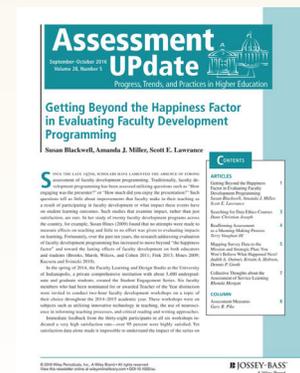
(e.g. Nichols, 1989). The resulting plans thus often had a somewhat formulaic quality. Most, for example, included a) an initial statement of principles, b) stated learning goals for general education and for each academic program, c) a charge to departments to find or develop a suitable assessment method (frequently accompanied by a list of methods to be considered) and, d) a schedule for data-collection and reporting. Implementing such plans, in turn, often involved the use of specially-funded “pilot” efforts by volunteer departments. The assessment plan of James Madison University and that of Ball State University, developed in the late 1980s and piloted in the early 1990s provide excellent examples of both. Keeping track of implementation and reporting, moreover, often demanded use of a tabular or matrix format and this too became a widespread feature of the “standard” approach (Banta, 1996). Methods, meanwhile, were healthily varied, including available standardized examinations, faculty-made tests, surveys and focus groups, and (increasingly, as the decade progressed) portfolios and work samples.

- **A Literature of Practice.** In assessment’s early days, practices and experiences were recorded in a fugitive literature of working papers, loosely-organized readings in *New Directions* sourcebooks, and conference presentations. But by the early 1990s, the foundations of a recognizable published literature could be discerned. Some of these works were by established scholars who summarized findings and provided methodological advice (e.g. Astin, 1985 and Pace, 1979). Others tried to document assessment approaches in terms that practitioner audiences could readily understand (e.g. Erwin, 1991 and Ewell, 1991). Still others continued the process of documenting institutional cases—of which there were now many—in standard or summary form (Banta and Associates, 1993).

The establishment of the movement’s own publication, *Assessment Update*, in 1989 was also an important milestone—providing relevant commentary on methods, emerging policies, institutional practices. As its editorial board envisioned, its contents were short, practical, and topical—providing the field with a single place to turn for ideas and examples. This supplemented the already-established role of *Change* magazine, which provided an early venue for assessment authors and continued to regularly print assessment-related essays. Through its Assessment Forum, moreover, AAHE issued a range of publications, building first upon conference presentations and continuing in a set of resource guides. In strong contrast to fifteen years previously, assessment practitioners in 2000 thus had a significant body of literature to guide their efforts that included systematic guides to method and implementation, well-documented examples of campus practice, and comprehensive treatises integrating assessment with the broader transformation of teaching and learning.

- **Scholarly Gatherings and Support.** Initiated on a regular annual cycle in 1987, the AAHE Assessment Forum was by 1989 *the* conference for practitioners, providing a regular gathering-place for scholarly presentation and exchange. Sessions developed for the Forum required formal documentation and often ended up as publications. The Forum also maintained professional networks, promoted idea-sharing, and provided needed moral support and encouragement. The latter was especially important in assessment’s early years because there were few practitioners and they were isolated on individual campuses. Other conferences arose at the state level including (among others) the South Carolina Higher Education Assessment (SCHEA) Network, the Washington Assessment Group (WAG), and the Virginia Assessment Group (VAG)—often directly supported by state higher education agencies. Some of these state-level groups published regular newsletters updating members on state policy initiatives and allowing campuses to showcase their programs. When the AAHE Assessment Forum ceased with the demise of its parent organization, its place was soon taken by the Assessment Institute in Indianapolis, with attendance figures topping 1500.

The establishment of *Assessment Update* in 1989 was an important milestone—providing relevant commentary on methods, emerging policies, institutional practices.



- **A “Semi-Profession.”** Although assessment remained largely a part-time activity, entering the mainstream also meant a rise in the number of permanent positions with assessment as a principal assignment. Position titles like “Assessment Coordinator” with formal job descriptions are now commonplace, usually located in Academic Affairs or merged with Institutional Research. The creation of such positions was in large measure a result of external pressure to put recognizable campus programs in place so that accreditors could notice them. Certainly such roles helped build badly-needed local capacity and infrastructure.

For assessment as a whole, one clear result of this evolution today is an established community of practice that in some ways resembles an academic discipline. Among its earmarks are an identifiable and growing body of scholarship, a well-recognized conference circuit, and a number of “sub-disciplines” each with its own literature and leading personalities. Those doing assessment, moreover, have evolved a remarkably varied and sophisticated set of tools and approaches and an effective semi-professional infrastructure to support what they do. These are significant achievements—far beyond what numerous early observers expected.

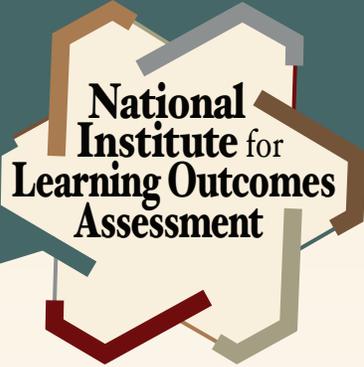
Into the Future. In assessment’s first decade, the question of “when will it go away?” was frequently posed. This was largely because the movement was diagnosed by many as a typical “management fad,” like Total Quality or Management by Objectives (MBO), that would quickly run its course (Birnbaum, 2000). Yet assessment has shown remarkable staying power and has undoubtedly attained a measure of permanence, at least in the form of a visible infrastructure. Several factors appear responsible for this phenomenon. Probably the most important is that external stakeholders will not let the matter drop. State interest is now stronger than ever, fueled by demand-driven needs to improve “learning productivity” and by burgeoning state efforts to implement standards-based education in K-12 education (Ewell, 1997). Accreditation organizations, meanwhile, have grown increasingly vigorous in their demands that institutions examine learning outcomes, though they are also allowing institutions more flexibility in how they proceed. Market forces and the media are not only more powerful, but are also far more performance-conscious and data-hungry than they were two decades ago. Assessment has thus become an unavoidable condition of doing business: institutions can no more abandon assessment than they can do without a development office.

The last twenty years have also seen a revolution in undergraduate instruction. In part, this results from technology. In part, it reflects the impact of multiple other “movements” including writing across the curriculum, learning communities, problem-based learning, and service learning. Together, these forces are fundamentally altering the shape and content of undergraduate study. These changes are sustaining assessment in at least two ways. Most immediately, new instructional approaches are forced to demonstrate their relative effectiveness precisely because they are new. Assessment activities are therefore frequently undertaken as an integral part of their implementation. More subtly, the very nature of these new approaches shifts the focus of attention from teaching to learning. In some cases, for instance, direct determination of mastery is integral to curricular design. In others, common rubrics for judging performance are required to ensure coherence in the absence of more visible curricular structure. Assessment has thus been sustained in part because it has become a necessary condition for undertaking meaningful undergraduate reform—just as the authors of *Involvement in Learning* foresaw.

For assessment as a whole, one clear result of this evolution today is an established community of practice with identifiable and growing body of scholarship, a well-recognized conference circuit, and a number of “sub-disciplines” each with its own literature and leading personalities.

References

- Association of American Colleges (1985). *Integrity in the college curriculum: A report to the academic community*. Washington, DC: AAC.
- Astin, A.W. (1985). *Achieving educational excellence*. San Francisco: Jossey-Bass.
- Banta, T.W. (1996). The power of a matrix. *Assessment Update*, 8(4), 3-13.
- Banta, T.W., and Associates (1993). *Making a difference: Outcomes of a decade of assessment in higher education*. San Francisco: Jossey-Bass.
- Bennett, W.J. (1984). *To reclaim a legacy: A report on the humanities in higher education*. Washington, DC: NEH.
- Birnbaum, R. (2000). *Management fads in higher education: Where they come from, what they do, why they fail*. San Francisco: Jossey-Bass.
- Erwin, T.D. (1991). *Assessing student learning and development*. San Francisco: Jossey-Bass.
- Ewell, P.T. (2002). An emerging scholarship: A brief history of assessment. In T.W. Banta and Associates, *Building a scholarship of assessment*. San Francisco: Jossey-Bass, 3-25.
- Ewell, P.T. (1997). Accountability and assessment in a second decade: New looks or same old Story? In AAHE, *Assessing impact, evidence and action*, 7-22. Washington, DC: AAHE.
- Ewell, P.T.; Finney, J.E.; and Lenth, C. (1990). Filling in the mosaic: The emerging pattern of state-based assessment. *AAHE Bulletin*, 42, 3-7.
- Loacker, G.; Cromwell, L.; and O'Brien, K. (1986). Assessment in higher education: To serve the learner. In OERI, *Assessment in American higher education: Issues and contexts*. Washington, DC: OERI, U.S. Department of Education, 47-62.
- National Governors' Association (1986). *Time for results*. Washington, DC: National Governors' Association (NGA).
- National Institute of Education, Study Group on the Conditions of Excellence in American Higher Education (1984). *Involvement in learning: Realizing the potential of American higher education*. Washington, DC: U.S. Government Printing Office.
- Nichols, J.O. (1989). *Institutional effectiveness and outcomes assessment implementation on campus: A practitioner's handbook*. New York: Agathon Press.
- Pace, C.R. (1979). *Measuring the outcomes of college*. San Francisco: Jossey-Bass.
- U.S. Department of Education, National Commission on Excellence in Education (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: U.S. Government Printing Office.



National Institute for Learning Outcomes Assessment

Please Cite As:

National Institute for Learning Outcomes Assessment. (2019, January). *A historical overview of assessment: 1980s-2000s* (Assessment Brief). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

Follow us on social media:



@NILOA_web



@LearningOutcomesAssessment

Sign up to receive our monthly **NILOA Newsletter** and stay up to date with our research and publications.



Assessment
Brief:

History of Assessment

www.learningoutcomesassessment.org