Introduction

Before diving into the history of assessment, let us first review the formal definition. The word “assessment” may be defined as the following: “an appraisal,” “an evaluation,” or “a judgment.” In terms of student-learning, the following definition by Thomas Angelo is considered: “Assessment is the ongoing process of 1) establishing clear, measurable expected outcomes of student learning, 2) ensuring that students have sufficient opportunities to achieve those outcomes, 3) systematically gathering, analyzing, and interpreting evidence to determine how well student learning matches (faculty’s) expectations, and 4) using the resulting information to understand and improve student learning.” We are concerned with the latter, and how student learning outcomes and program level outcomes are inherently associated with the assessment process. The following is a literature review on Peter Ewell’s “An Emerging Scholarship: A Brief History of Assessment.” First, the (4) areas of scholarship which gave rise to the assessment movement will be presented. Second, assessment problems which arose during this period are highlighted. Third, the relationship between assessments and mandates at both the local and national level will be discussed.

The Roots of the Assessment Movement

Ewell suggests that (4) areas of scholarship (starting from the 1930s) served as the fuel for the assessment movement which came to strong fruition in the 1970s-1980s: 1) student learning in college, 2) retention and student behavior, 3) program evaluation and scientific management, and 4) mastery and competency-based learning.
Student Learning in College. The earliest known examples of higher education assessment in the United States are from the 1930s-1940s (Learned and Wood, 1938)\textsuperscript{ii}, which focused on “individual colleges serving (the representative age group) of 18-21 year olds in traditional residential environments.” This ostensibly does not include community or junior colleges, nor an inner-city population. The assessment studies at the time focused not only on cognitive gain but also on student maturation and attitudinal development (Chickering, 1969)\textsuperscript{iii}. Ewell suggests that the research was as much an application of educational psychology as it was developmental psychology. Over the next few decades, enough work was done in this scholarly field of student learning to serve as the foundation on which assessment practitioners would eventually build upon. Eventually, the research cumulated into the following during the 1970s: longitudinal studies of net effects (Astin, 1977)\textsuperscript{iv}, public policy context (Bowen, 1977)\textsuperscript{v}, and the role of the college environment in actual student behavior (Pace, 1979)\textsuperscript{vi}. Ewell claims that the (3) aforementioned works were “especially influential” to the scholarship of assessment, providing both “conceptual and methodological contributions.” Prominent examples included early taxonomies on outcomes, which provided a common language which would facilitate the transition into a new scholarly field. Ewell also lists “early models of student growth and development” as another product of student learning scholarship. Furthermore, this period gave rise to assessment tools, such as “cognitive examinations, longitudinal and cross-sectional surveys, and quasi-experimental designs.” It is imperative to note that these early assessment tools were both quantitative and qualitative.
Retention and Student Learning. The second foundation on which the assessment movement of the 1970s was built upon is referred by Ewell as “retention and student learning.” This separate area of research complemented research on college student learning. Ewell claims that many of the ideas and methodologies in this area actually permeated into assessment scholarship. He suggests that “many practitioners (of assessment) worked to emulate” the research models utilized in retention and student learning. One model used was Tinto’s idea of “academic and social integration (1975)\textsuperscript{vii}, which proved useful in guiding research on student learning.” Academic integration is when the student, regardless of his/her talents/skills, accepts courses that are at his/her level, and can succeed in them, without losing confidence. Social integration is when the student can succeed in interacting in peers, but without being negatively influenced. A second overlap was the use of brand new assessment tools, including “longitudinal study designs, specially configured surveys, and multivariate analytical techniques, later adopted by many assessment practitioners.” Thirdly, the concept of “action research” was coined. Although the research was based off of accepted theories and common methodologies, the end goal ultimately was “informed intervention” (Lenning, Beal, and Sauer, 1980)\textsuperscript{viii}. In other words, can the data suggest ways for faculty to improve student retention?

Evaluation and Scientific Management. This is Ewell’s third suggestion which helped gave rise to the scholarship of assessment. In the 1960s and 1970s, there were several large-scale federal programs for research concerning program review. At the time, program evaluation “relied heavily on quantitative methods.” In addition, a new concept coined “scientific management” simultaneously sprang up and was widely
Literature review for Module 1: “Myth Busters of Assessment.”

adopted. This term refers to “providing an output variable for cost-benefit studies and investigations of social return on investment.” This naturally was applied to administrative projects such as strategic planning, budgeting and program review. Specifically, there was “an explicit attention to student outcomes” (Enthoven, 1970). From the tradition of program review and scientific management, out came “one of the most extensive taxonomies of collegiate outcomes ever produced and stimulated a range of surveys designed to provide campuses with information about how students used and perceived their programs (Lenning, 1977).” Moreover, qualitative tools were commonly used in program evaluation, including open-ended interviewing. These quickly appealed to those in assessment who were “skeptical of overly empirical methodologies.”

**Mastery Learning.** This model, which Ewell considers the fourth contributor to assessment scholarship, actually sprang from elementary and secondary levels. Mastery and competency-based designs for learning are “based entirely on agreed-upon outcomes” and focus on “assessing and certifying individual student achievement.” One can easily see SLOs and PLOs embedded in the prior sentence. This concept was quickly adopted by post-secondary institutions in the mid-1960s and served as the alternative to the traditional “test and measure” approach.

**Common Problems Early On In the Assessment Movement**

Problems immediately appeared during the early parts of the assessment movement. This follows from Ewell’s premise that (4) separate areas of scholarship actually gave rise to assessment. He suggests that assessment practitioners experienced the following mishaps: 1) the word “assessment” lacked a common definition, 2) no one
Literature review for Module 1: “Myth Busters of Assessment.”

was sure of the most useful and credible ways to generate evidence of student learning, and 3) there was no prescribed way to implement the assessment process.

**The Definition of “Assessment.”** Ewell claims there to have been at least (3) different meanings for “assessment,” as the definition is subjective. Firstly, in the mastery and competency-based learning model, the practitioner’s definition would “refer to the processes used to determine an individual’s mastery of complex abilities, generally through observed performance (Alverno College Faculty, 1979)xi. A student’s development over time was also valued and emphasized. Feedback on individual performance was also the norm.

Secondly, a practitioner in the K-12 level would have a more general, less individualistic definition. “Assessment” referred to the “large-scale testing programs,” funded both at the federal and state levels. These examinations were “not to examine individual learning but rather to benchmark school and district performance in the name of accountability.” The large-scale exams were standardized and “founded on well-established psychometric principles.” A quick generic summary report with performance statistics could then be easily generated.

Thirdly, an alternative definition of “assessment” referred to a unique program evaluation, one that was designed to “gather evidence to improve curricula and pedagogy.” This approach focused more on the aggregate, rather than on the individual and used assessment tools such as “examinations, portfolios and student work examples,” and student surveys of their experiences. Since the emphasis was on improvement, “assessment was as much about *using* (data) as it was about psychometric standards.”
There would be ongoing discussion and debate in the early days of the assessment movement (and still today): 1) accountability versus improvement 2) aggregate versus individual results and 3) quantitative versus qualitative methods? Ultimately a consensus was reached “on the use of multiple methods for program improvement (American Association for Higher Education, 1992).”

**Assessment Tools.** The majority of assessment instruments at the time were not designed to accumulate evidence that supported student improvement. Standardized exams, which were the norm, were inappropriate for program evaluation, such as the Graduate Record and ACT Examinations. The content within such exams only “approximated the domain of any given institution.” However, this did not prevent colleges from moving on with program evaluation.

Ewell claims a specific example as the “harbinger.” The ACT Colleges Outcome Measures Project (COMP) examination (Forrest and Steele, 1978) evaluated GE outcomes and “supported group-level inferences about student learning.” In addition, the ACT COMP “generated GE outcomes in novel ways, emphasizing (knowledge) in real-world situations and requiring authentic demonstrations of performance.”

During 1986-1989, the major testing organizations at the time quickly followed the ACT COMP’s lead and generated purpose-built group-level examinations specifically for program evaluation. In addition to those offered by ACT and by the Educational Testing Service (ETS), many institutions started to design and administer their own exams and surveys, *heavily involving faculty* instead of educational consultants or corporations.
Implementation. Since the assessment movement was just starting to pick up steam, institutions really didn’t have concrete examples or guides on how to actually assess student learning and evaluate their programs. Another major problem was the lack of documentation which was either maintained or shared. In other words, institutional experience was missing. Ewell considers the faculty of Alverno College to be one of the first, “whose abilities-based curriculum, designed around performance assessments of every student, was both inspiring and daunting (Alverno College Faculty, 1979)". Also mentioned as a pioneer is Truman State University, where since 1973 nationally normed exams are administered to evaluate the “integrity of its degrees.” Finally, Ewell lists the University of Tennessee, which “under the stimulus of Tennessee’s performance funding scheme, became the first major public university to develop a comprehensive multimethod system of program assessment (Banta, 1985)." In the late 1980s another wave of programs followed, especially in response to a flux of new state mandates. In 1987, approximately 55% of institutions claimed they had an assessment program. By 1993, the figure rose to 98%. Assessment and its scholarship became mainstream.

Literature review for Module 1: “Myth Busters of Assessment.”


Additional Readings

General Documents on the History of Assessment in Higher Education


Pfeifer, J. www.csus.edu/indiv/p/pfeiferj/ShortHistoryAssessment.ppt

Northern Illinois University http://www.niu.edu/assessment/Manual/history.shtml


http://books.google.com/books?hl=en&lr=&id=tBPjb-Q3mDAC&oi=fnd&pg=PA67&dq=a+scholar-practitioner+model+for+assessment&ots=0Rlad1Mu-e&sig=yriREq7EiS479np8n_e_TYtwBs

Assessment History at Various Institutions

http://assessment.truman.edu/history.asp

https://oira.syr.edu/oira/Assessment/AssessSL/HistoryAssessSL.htm

http://www.missouristate.edu/fctl/103655.htm