1

*This chapter gives a brief history of general education assessment, responds to common criticisms of general education assessment, and makes a case for assessing general education as a critical element of our responsibility as faculty members.*

# The Case for Assessing Complex General Education Student Learning Outcomes

*Jeremy D. Penn*

Evaluation of educational achievement has been a part of education since at least 589–613 A.D. (Pinar, Reynolds, Slattery, and Taubman, 1996). Liberal education, the foundational element for many general education programs, has existed even longer, with Aristotle, Plato, and other Greek philosophers developing its philosophical foundations in the fourth century B.C. (Mulcahy, 2008). The industrial revolution in the nineteenth century in the United States challenged the idea of a liberal education and brought rise to the factory model of schooling, in which students were "processed," separated into "age-related cohorts called classes or standards," and taught "a standard course" through "teacher-centered methods" (Hargreaves, 1994) for the purpose of preparing students for a life of work. Newman (1947), along with other defenders of the traditional view of liberal education, continued to argue for "knowledge which is its own end" (p. 98) and education as "the cultivation of the intellect" (p. 107).

Modern approaches to general education have integrated the traditional view of liberal education with preparation for work into an overall "preparation for life and personal development" (Mulcahy, 2008, p. 177) that includes learning for its own sake and workplace readiness. The National Leadership Council for Liberal Education and America's Promise issued a challenge to all postsecondary institutions to implement a curriculum with a "comprehensive set of aims and outcomes that are essential for all students" (National Leadership Council for Liberal Education and America's Promise, 2007, p. 4), including such outcomes as teamwork, critical thinking, and communication. A large number of higher education

NEW DIRECTIONS FOR INSTITUTIONAL RESEARCH, no. 149, Spring 2011 © Wiley Periodicals, Inc.
Published online in Wiley Online Library (wileyonlinelibrary.com) • DOI: 10.1002/ir.376

institutions in the United States have been responding to the changing conceptions of general education by revising their general education programs (Johnson, Ratcliff, and Gaff, 2004).

Assessment of general education also has a long history, although relative to the age of liberal education and educational evaluation it is a very recent development. One of the first recorded efforts to comprehensively assess student achievement in higher education in the United States occurred in the late 1920s and early 1930s, when many institutions had general education programs in practice but not necessarily in name. In this effort, nearly forty-five thousand high school and college students were given a multiple-choice test that assessed students' knowledge of the physical world, mathematics and science, and the social world including psychology, sociology, statistical methods, and ancient cultures (Learned and Wood, 1938). In an approach that foreshadowed the popularity of value-added methodology, the test was given to a group of students as sophomores and again two years later when they were seniors at forty-five institutions. Learning gains of 0.02 to 0.56 standard deviations were found at the school level and used to rank institutions into achievement groups. The researchers found "there were few 'surprises' in the placement of colleges as determined by their achievement-test averages" (1938, p. 15).

Further solidifying the importance of assessment of educational achievement was Ralph Tyler's book on curriculum development and evaluation. Tyler identified four questions to guide curriculum development and evaluation, focusing on identification of educational purposes, selection and organization of educational experiences, and whether or not those experiences were "actually producing the desired results" (Tyler, 1949, p. 105).

Although students' achievement of general education goals has been assessed since at least the 1930s, use of assessment of general education as an institutional improvement and accountability methodology has its roots in the mid-1980s. In 1984, the Study Group on the Conditions of Excellence in American Higher Education recommended using assessment "as a means to provide information about the teaching and learning process and as feedback to help improve the effectiveness with which students, faculty, and the institution carry out their work" (p. 53). In 1985, the First National Conference on Assessment in Higher Education was held in Columbia, South Carolina—further evidence that the higher education assessment movement was well under way (Ewell, 2002). By 1987, 55 percent of surveyed institutions had established an assessment program, and by 1993 this had increased to 98 percent (Ewell, 2002). All six regional accreditors now require assessment of general education as a condition of accreditation, firmly establishing assessment of general education as a key element of institutional accountability and improvement.

## Tackling Critiques of Assessment of General Education

The long history of general education assessment includes numerous critiques of general education assessment practices. For instance, more than seventy years ago Learned and Wood (1938) identified concerns "with the extent to which the so-called comprehensive, objective, or 'new-type' tests, less familiar ten years ago than now, would prove to justify their use for the purposes in view" (p. 13). Practices for assessing general education are often critiqued for their philosophical foundations, the quality of assessment measures, the relationship between teaching practices and assessment, and the role of academic freedom and external accountability (see Bresciani, 2007; Ewell, 2002; Hutchings, Marchese, and Wright, 1991; Kramer, 2009).

This section responds to five of the most pervasive critiques.

**General Education Learning Outcomes Cannot Be Defined.** Schwyzer (2007) summarizes the core of this critique when he writes about education as a transformative experience that can change the direction of students' lives, concluding "there's no [student learning outcome] that can measure that" (para. 13). Developing a definition of a complex learning outcome, a process usually called *construct validation* by researchers, can be challenging. Fortunately, much work in this area has already been completed. (See Chapter Two of this volume for a description of one approach grounded in psychology, and the chapters in Part Two for examples.) For example, Cronbach (1955) suggested a variety of methods to establish construct validity, or to "make clear what something is" (p. 290), including examination of differences between two groups that are expected to differ on the construct (for example, freshmen and seniors), use of correlation matrices and factor analyses across multiple measures of a construct, examination of the internal structure of the test itself, and study of changes in the scores over a number of occasions (see Campbell and Fiske, 1959, for more details). The approaches developed as part of construct validation and described in this volume can be used to define any complex general education student learning outcome.

**General Education Learning Outcomes Cannot Be Assessed with Existing Tools.** The lack of meaningful measures of student achievement on complex general education outcomes could be due to faculty disagreement on what to measure, by the expense of developing measures, or by resistance on the part of colleges (Hersh, 2005). Faculty disagreement can be addressed by implementing a process to develop a clear, research-based definition (see Chapter Two of this volume) or by implementing a systematic process for building faculty engagement (see Maki, 2004).

Some existing assessment measures are expensive, but there are many that are not, or even free. One source to use in searching for existing tools is the Buros Institute of Mental Measurements (http://www.unl.edu/buros/),

which publishes the *Mental Measurements Yearbook* approximately biennially. The yearbook summarizes and critically analyzes hundreds of newly released instruments in every volume. Another source is the *Valid Assessment of Learning in Undergraduate Education* (VALUE) project, which brought together national experts in fifteen areas to develop meta-rubrics (http://www.aacu.org/value/rubrics) that are free for use or adaption. In addition, each chapter in Part Two of this volume lists some assessment measures that are being used to gather information on student achievement of each general education learning outcome, many of which are inexpensive or free.

If an existing assessment measure cannot be located, there are a number of resources available that provide guidance on how to develop your own assessment measure (see Chapter Two of this volume). Although it takes a significant investment of time, developing your own assessment measure can be less expensive than purchasing an existing measure, and it can be one approach for building faculty members' engagement in and ownership of assessment.

Most important, remember that, in the words of Voltaire, "*le mieux est l'ennemi du bien*" (1772, p. A3); "the best is the enemy of the good." The perfect instrument to measure a specific complex general education student learning outcome may never exist. We should continue to work on improving assessment measures for general education student learning outcomes, but we cannot afford to postpone assessment until the perfect measure is developed.

**General Education Learning Outcomes Cannot Be Taught.** Seymour and Hewitt (1997) found many science, mathematics, and engineering faculty members expected substantial attrition from their programs because "those presumed to lack sufficient natural ability to continue are thought to discover their limitations, and/or their true vocation for some other discipline and leave" (p. 7). This view, that academic abilities such as critical thinking or quantitative reasoning cannot be taught and are innate and that the role of higher education is to separate out those who have them and those who do not, may be widely held in many disciplines.

A contrasting view is expressed by Ericsson, Krampe, and Tesch-Römer (1993), who proposed a theoretical framework that "explains expert performance in terms of acquired characteristics resulting from extended deliberate practice and that limits the role of innate (inherited) characteristics to general levels of activity and emotionality" (p. 363). Other researchers have focused on time, not innate ability, as the most essential element related to student achievement (see Finnegan and Hyle, 2009; Gettinger, 1984; Hong and Hong, 2009; and Millot, 1995). Carroll (1963) proposed that a learner "will succeed in learning a given task to the extent that he spends the amount of time that he *needs* to learn the task" (p. 725). Carroll goes on to suggest that perceived differences in

students' abilities can be explained by the amount of time required for a particular learning task; students who require a small amount of time to achieve a learning goal have "high aptitude" while those requiring a large amount of time to achieve a learning goal have "low aptitude" (p. 725).

If students, given sufficient time, can become concert violinists and grandmaster chess experts through deliberate practice, then undoubtedly they can also learn to become critical thinkers, quantitative reasoners, and writers through deliberate practice. Given enough time, nearly all of our students are capable of learning anything we would like to teach them.

**Results from Assessment of General Education Learning Outcomes Are Never Used for Anything.** Too often data from assessment of general education are ignored, abandoned, lost, or simply not used for anything other than proving that assessment was done. There are many reasons assessment data may not be used, among them lack of a shared vision for achievement of the general education learning outcomes, failure to clearly define the general education learning outcomes, failure to gather clear and meaningful assessment data on student achievement of the learning outcomes, distrust in the quality of the assessment measures, lack of faculty members' involvement in the assessment process, and perception that teaching the general education learning outcomes is someone else's responsibility. It does not have to be this way. A number of institutions have been able to overcome these obstacles and as a result transform their general education programs (see Bresciani, 2007). Just because results from assessment of general education may not have been used for anything does not suggest they *cannot* be used.

**Assessment of General Education Learning Outcomes Is a Threat to Academic Freedom.** Academic freedom is a foundational principle that is essential for achievement of the missions of our institutions. Unfortunately, assessment of general education is often viewed as an external intrusion into the curriculum that limits academic freedom by telling faculty members what and how to teach. Besides, critics contend, we already do assessment through grading.

Yet at its best, assessment of general education supports academic freedom. The *1940 Statement of Principles on Academic Freedom and Tenure* (American Association of University Professors, 1970) is an essential statement that protects the rights of faculty members to teach and perform research freely. These rights come with the responsibility to "seek above all to be effective teachers and scholars" (p. 8). Assessment of general education is an essential element of meeting this responsibility, by establishing a process through which students' achievement of the general education learning outcomes can be studied and improved.

Grading is an essential element of teaching a course and, along with the credit hour, is the currency of higher education. However, grading is at the wrong level for use in determining whether or not students are

achieving the general education outcomes and does not always yield useful and meaningful data for improving students' achievement. Even if a common grading scale is used, individual instructors may weight course elements differently. Some faculty members may include attendance or participation as a portion of the grade, while others may include only the results from multiple-choice tests. As a result, interpreting an individual grade or a grade point average as evidence of student achievement of the general education outcomes is problematic. Avoidance of grades as a tool in assessing students' achievement of the general education outcomes is not intended to be a rejection of faculty members' expertise in evaluating student achievement, of the value of course assignments or projects in assessing student achievement of the general education outcomes, or of the central role of courses in students' educational experiences. Rather, it is recognition that course grading serves a different purpose and is not the best way to gather evidence on students' achievement of the general education outcomes.

## The Case for Assessment of General Education

**Assessment of General Education Has the Potential to Transform Our Institutions.** Our institutions are facing many challenges: students are approaching their education in new and unfamiliar ways, colleges are reconceptualizing faculty roles, many of our institutions are facing financial crisis, political pressure for accountability is strong and getting stronger, and students matriculate unprepared and graduate not having learned as much as they should have.

To respond to these and other challenges, institutions must be able to react rapidly and effectively. Yet we know this is not how our institutions are designed. Current structures and processes in our institutions are the result of decades of tradition from which we must struggle mightily for liberation. To make matters worse, many institutions lack a systematic mechanism to use in responding to a changing environment.

Assessment of general education student learning outcomes can be one such mechanism. Assessment of general education becomes a mechanism for transformation by reframing Tyler's guiding questions on curriculum development (1949) into approachable, action-oriented questions, making them a systematic part of institutional self-reflection. These questions become:

1. Are students learning what they should be learning?
2. Which teaching, curricular, and co-curricular approaches are working well, and which approaches need to be modified?
3. What additional educational experiences should be furnished to students, and how should our existing experiences be reorganized?
4. Is our process for assessing general education working effectively?

Many institutions have already experienced transformation through assessment of general education. For example, Coker College developed a new writing effectiveness program, and Alverno College implemented a major redesign of the communications courses (Bresciani, 2007). A college in the State University of New York system received a Title III grant, a community college in the system revised its algebra/precalculus/calculus curriculum, and two additional colleges (one doctoral and one comprehensive) implemented a faculty development program (Bresciani, 2007). Others, such as the University of South Florida, completely reformed the general education program (Bresciani, 2007). North Carolina State University improved teaching through inclusion of service-learning (Banta, Jones, and Black, 2009).

Northeastern Illinois University redeveloped the writing and mathematics curricula (Banta, Jones, and Black, 2009). Oklahoma State University created faculty development initiatives and implemented new writing requirements (Banta, Jones, and Black, 2009). Assessment of general education has already transformed many institutions and has the potential to serve as a mechanism for responding to our changing educational environment.

**Assessment of General Education Can Help Meet (Not Always Unreasonable) Expectations for Accountability.** We know, from watching students grow in our classes and throughout their degree programs, that earning a degree from our institutions represents a significant achievement that will produce benefits for their future lives and for our communities. This belief is supported by a substantial body of research (Bok, 2006; Cuadras-Morató and Mateors-Planas, 2006; Mortenson, 1999; Pascarella and Terenzini, 2005; U.S. Department of Labor, Bureau of Labor Statistics, 2006). When we are confronted with demands for proof that what we believe is occurring is actually occurring in our institutions as part of an accountability movement, our initial response is often incredulity and anger. Is our professional judgment not sufficient?

At the same time, we recognize many areas of concern. The U.S. Census Bureau estimated in 2007 that 54.4 percent of the U.S. population age twenty-five and older had earned some college credit, but only 27.5 percent had completed a bachelor's degree or higher (U.S. Census Bureau, 2009). The four-year completion rate at public universities is only 28 percent (Higher Education Research Institute, 2005). Learning gains from the freshman to senior year range from 0.24 to 0.90 standard deviations for such areas as mathematical and quantitative skills, subject matter knowledge, and reflective thinking, with a high of 2.0 on epistemological sophistication (Pascarella and Terenzini, 2005). These learning gains are noteworthy, but they are not as large as they could or should be. Former Department of Education Secretary Margaret Spellings summarized these concerns, saying, "It's time to examine how we can

get the most out of our national investment [in higher education]" (Spellings, 2005).

Demands for accountability are not always unreasonable. Higher education receives considerable support from taxpayers, totaling nearly $83 billion from state and local taxes in 2007 alone (State Higher Education Executive Officers, 2009). In addition, our institutions are so important to our communities and the individuals we serve that it is not surprising there is a need for evidence on whether or not we are achieving our goals. Although the dialogue about how best to achieve accountability without distorting and corrupting the very thing we are trying to measure is ongoing (see Amrein and Berliner's uncertainty principle, 2002), we cannot, and should not, resist all accountability efforts by labeling them as unreasonable.

Assessment of general education helps us meet expectations for accountability in several ways. First, it produces clear evidence on our students' achievement on learning outcomes that are most central to our institutions. Assessment of general education also facilitates a dialogue about what we expect students to learn in our institutions and identify core knowledge, skills, abilities, and dispositions that are important for all students. At the same time, assessment of general education allows us to exhibit learning and achievements that are unique to each of our institutions, highlighting one of our higher education system's greatest strengths. If used appropriately, assessment of general education can be a meaningful and valuable component of accountability.

**Assessment Is Part of Our Responsibility as Faculty Members.** There is no better argument for assessment of general education than to say it is the right thing to do. Too often we focus only on the carrot and the stick, forgetting that it is our responsibility as faculty members to ensure that our educational programs are having their desired impact.

Teaching, one of our two primary responsibilities as faculty members (American Association of University Professors, 1970), is made up of three interrelated elements: instruction, curriculum, and assessment. Instructional practices have an impact on the curriculum experienced by our students. Curricular decisions about what educational experiences to offer shape our instructional practices. Assessment informs our instruction by revealing effective and ineffective practices and can also be used as an instructional tool. Assessment reveals whether or not the educational experiences we include in the curriculum are resulting in the desired student achievement. Instruction, curriculum, and assessment are inseparable. We are mistaken if we believe that we can ignore assessment of general education or pass it off to others and still hold on to instruction and curriculum. Assessment of general education may not hold the glamour of research or the gratification of teaching, but it is a critical element of our work as faculty members and can serve an essential element in transforming and improving our institutions.

## References

American Association of University Professors. "1940 Statement of Principles on Academic Freedom and Tenure (with 1970 Interpretive Comments)," 1970. Retrieved March 8, 2010, from http://www.aaup.org/NR/rdonlyres/EBB1B330-33D3-4A51-B534-CEE0C7A90DAB/0/1940StatementofPrinciplesonAcademicFreedomandTenure.pdf.

Amrein, A. L., and Berliner, D. C. "High-Stakes Testing, Uncertainty, and Student Learning." *Education Policy Analysis Archives,* 2002, *10*(18). Retrieved Sept. 11, 2006, from http://epaa.asu.edu/epaa/v10n18/.

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

Bok, D. *Our Underachieving Colleges: A Candid Look at How Much Students Learn and Why They Should Be Learning More*. Princeton: Princeton University Press, 2006.

Bresciani, M. J. "The Challenges of Assessing General Education: Questions to Consider." In M. J. Bresciani (ed.), *Assessing Student Learning in General Education: Good Practice Case Studies*. Bolton, Mass.: Anker, 2007, 1–15.

Bureau of Labor Statistics. "National Occupational Employment and Wage Estimates: May 2008," 2008. Retrieved March 9, 2010, from http://www.bls.gov/oes/2008/may/oes_nat.htm#b25-0000.

Campbell, D. T., and Fiske, D. W. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix." *Psychological Bulletin,* 1959, *56*(2), 81–105.

Carroll, J. B. "A Model of School Learning." *Teachers College Record,* 1963, *64*, 723–733.

Cronbach, L. J. "Construct Validity in Psychological Tests." *Psychological Bulletin,* 1955, *52*(4), 281–302.

Cuadras-Morató, X., and Mateors-Planas, X. "Are Changes in Education Important for the Wage Premium and Unemployment?" *International Economic Review,* 2006, *47*(1), 129–160.

Ericsson, K. A., Krampe, R. T., and Tesch-Römer, C. "The Role of Deliberate Practice in the Acquisition of Expert Performance." *Psychological Review,* 1993, *100*(3), 363–406.

Ewell, P. T. "An Emerging Scholarship: A Brief History of Assessment." In T. W. Banta and Associates (eds.), *Building a Scholarship of Assessment*. San Francisco: Jossey-Bass, 2002, 1–25.

Finnegan, D. E., and Hyle, A. E. "Assistant to 'Full': Rank and the Development of Expertise." *Teachers College Record,* 2009, *111*(2), 443–479.

Gettinger, M. "Achievement as a Function of Time Spent in Learning and Time Needed for Learning." *American Educational Research Journal,* 1984, *21*(3), 617–628.

Hargreaves, A. *Changing Teachers Changing Times: Teachers' Work and Culture in the Postmodern Age*. New York: Teachers College Press, 1994.

Hersh, R. H. "What Does College Teach?" *Atlantic Monthly*, 2005. Retrieved May 26, 2010, from http://assessment.uconn.edu/docs/resources/ARTICLES_and_REPORTS/Richard_Hersh_What_Does_College_Teach.pdf.

Higher Education Research Institute. "How 'Good' is Your Retention Rate? Using the CIRP Survey to Evaluate Undergraduate Persistence," 2005. Retrieved July 24, 2007, from http://gseis.ucla.edu/heri/PDFs/DARCU_RB.PDF.

Hong, G., and Hong, Y. "Reading Instruction Time and Homogeneous Grouping in Kindergarten: An Application of Marginal Mean Weighting Through Stratification." *Educational Evaluation and Policy Analysis,* 2009, *31*(1), 54–81.

Hutchings, P., Marchese, T., and Wright, B. *Using Assessment to Strengthen General Education*. Washington, D.C.: American Association for Higher Education, 1991.

Johnson, D. K., Ratcliff, J. L., and Gaff, J. G. "A Decade of Change in General Education." In *Changing General Education Curriculum*. (Special Issue.) New Directions for Higher Education, no. 125, 2004.

Kramer, P. I. "The Art of Making Assessment Anti-venom: Injecting Assessment in Small Doses to Create a Faculty Culture of Assessment." *Assessment Update,* 2009, *21*(6), 8–10.

Learned, W. S., and Wood, B. D. *The Student and His Knowledge*. New York: Carnegie Foundation for the Advancement of Teaching, 1938.

Maki, P. L. *Assessing for Learning: Building a Sustainable Commitment Across the Institution*. Sterling, Va.: Stylus, 2004.

Millot, B. "Economics of Educational Time and Learning." In M. Carnoy (ed.), *International Encyclopedia of Economics Education*. Oxford, UK: Pergamon-Elsevier, 1995.

Mortenson, T. G. "Why College? Private Correlates of Educational Attainment." *Postsecondary Education Opportunity,* 1999, *81*, 1–24.

Mulcahy, D. G. *The Educated Person: Toward a New Paradigm for Liberal Education*. Lanham, Md.: Rowman and Littlefield, 2008.

National Leadership Council for Liberal Education and America's Promise. *College Learning for the New Global Century*. Washington, DC: Association of American Colleges and Universities, 2007.

Newman, J. H. C. *The Idea of a University* (ed. C. F. Harrold). New York: Longmans, Green, 1947.

Pascarella, E. T., and Terenzini, P. T. *How College Affects Students: A Third Decade of Research* (vol. 2). San Francisco: Jossey-Bass, 2005.

Pinar, W. F., Reynolds, W. M., Slattery, P., and Taubman, P. M. *Understanding Curriculum: An Introduction to the Study of Historical and Contemporary Curriculum Discourses*. New York: Peter Lang, 1996.

Schwyzer, H. (2007). "The Educrats' Attack on Teaching." InsideHigherEd.com. Retrieved Oct. 8, 2007, from http://www.insidehighered.com/views/2007/10/08/schwyzer.

Seymour, E., and Hewitt, N. M. *Talking About Leaving: Why Undergraduates Leave the Sciences*. Boulder, Colo.: Westview, 1997.

Spellings, M. "Prepared Remarks for Secretary Spellings at the Meeting of the Commission on the Future of Higher Education," Sept. 19, 2005. Charlotte, North Carolina. Retrieved April 1, 2007, from http://www2.ed.gov/news/speeches/2005/09/09192005.html.

State Higher Education Executive Officers. "State Higher Education Finance: FY 2009," 2009. Retrieved March 9, 2010, from http://www.sheeo.org/finance/shef/SHEF_FY_2009.pdf.

Study Group on the Conditions of Excellence in American Higher Education. *Involvement in Learning: Realizing the Potential of American Higher Education*. Washington, D.C.: National Institute of Education, 1984.

Tyler, R. W. *Basic Principles of Curriculum and Instruction*. Chicago: University of Chicago Press, 1949.

U.S. Census Bureau. "Educational Attainment in the United States: 2007," 2009. Retrieved March 1,, 2011, from http://www.census.gov/prod/2009pubs/p20-560.pdf.

U.S. Department of Labor, Bureau of Labor Statistics. "Employment Status of the Civilian Noninstitutional Population 25 Years and over by Educational Attainment, Sex, Race, and Hispanic or Latino Ethnicity, 2006. Retrieved July 24, 2007, from http://www.stats.bls.gov/cps/cpsaat7.pdf.

Voltaire (Arouet, F.-M.). *La Bégueule: Conte moral*, 1772. Retrieved July 12, 2010, from http://www.archive.org/details/labgueulecontem00voltgoog.

*JEREMY D. PENN is the director of assessment and testing and an adjunct faculty member in the School of Educational Studies at Oklahoma State University.*