

Rickards, W. H., & Stitt-Bergh, M. (2016). Higher education evaluation, assessment, and faculty engagement In W. H. Rickards & M. Stitt-Bergh (Eds.), *Evaluating student learning in higher education: Beyond the public rhetoric*. *New Directions for Evaluation*, 151, 11–20.

1

Higher Education Evaluation, Assessment, and Faculty Engagement

William H. Rickards, Monica Stitt-Bergh

Abstract

Evaluative practice has a long and deep history in higher education. It has been a persistent part of instructional practice and curriculum, intricately entwined with scholarly efforts to address teaching and learning. From public policy and oversight perspectives, the questions of value and worth have often focused on inputs—faculty credentials, facilities, etc.—as well as fiscal responsibility. But in the last 30 years, attention has turned to student learning as a critical outcome and the assessment of learning as a principal endeavor. The developments in higher education assessment have involved increasingly sophisticated psychometric approaches to measurement as well as more teacherly orientations to the implementation of educational assessments within the individual contexts—and intentions—of colleges and universities. In this chapter, we introduce some of the issues in the field and argue that evaluation has a unique history that is committed to systematically bringing evidence of program outcomes and processes into the discourse of educators—administrators, faculty, and staff—as they examine and build on their own operations. We briefly review the current context and challenges and support increased evaluator–faculty collaboration. We make a case for how the analysis of evaluation practices in higher education is both a means to increasing expertise in those applications and to thinking about evaluation practices across developing and complex institutions. © 2016 Wiley Periodicals, Inc., and the American Evaluation Association.

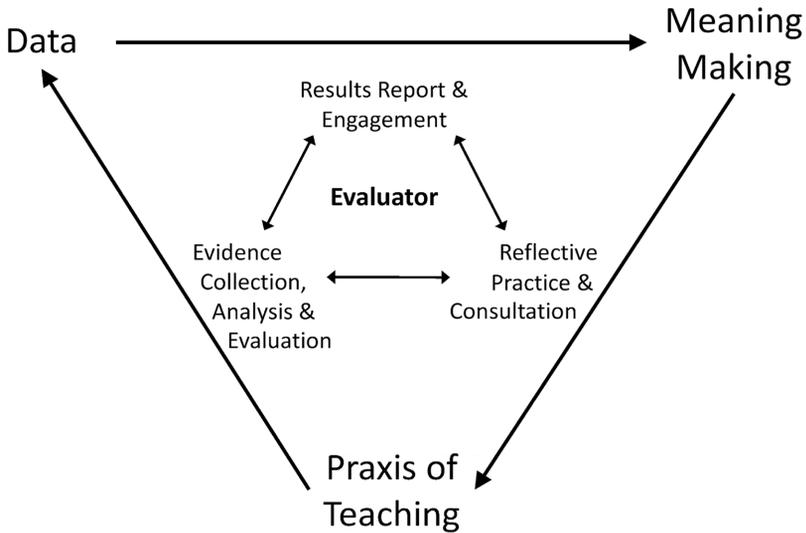
During the last 30 years in the United States, attention to accountability has led to widespread assessment of learning at the program and degree level, which is a new direction for faculty and administrators in higher education. The questions of what learning is expected, what instruments should be used to measure learning, and what evaluation practices drive changes and improvement are up for debate. Some of the public rhetoric presents an argument that academic programs are not adequately attending to student learning and that increased attention to testing would provide evidence that would result in educational transformation (Zemsky, 2007; see also Chambliss, 2007, and Secretary of Education's Commission on the Future of Higher Education, 2006). Other narratives based on extensive research locate the advance of student learning in the context of curriculum and the operation of academic units. These express equal concern about learning but emphasize the interaction of instruction, curriculum, and organizational process:

- Documentation of effective educational practice (Kuh, Kinzie, Schuh, Whitt, & Associates, 2005; Project DEEP, n.d.)
- Kezar's (2014) research on how colleges change, the role of leadership, and organizational processes
- Critical faculty engagement (Banta & Pike, 2012b)
- Use of student learning evidence to improve higher education (Kuh et al., 2015)

It can be useful to understand the interaction of three different perspectives in higher education discourse: (a) broad public concerns that are played out in national funding and policy decisions; (b) policies and practices at the institutional level (and with campus systems such as resources for evaluation and assessment staff that make complicated assessments feasible); and (c) educator perspectives that provide the context in which faculty and departments examine teaching and learning to make meaning out of available data on outcomes. It is in this context of multiple, sometimes competing, perspectives, and narratives on student learning that we describe learning outcomes assessment and the role of evaluation and evaluators.

Furthermore, in implementation, educators are regularly involved in a three-way dynamic that complicates evaluation and operation: data on student learning seldom have meaning in themselves; in the absence of interpretation and meaning making, data collection can be a distracting and expensive activity. Utilization requires a robust, cyclical relationship between *data* (student performance and educational outcomes) and the *meaning-making* activities in which faculty examine their practice and student outcomes, evaluate their principles and practices, make program decisions, and engage in *praxis* (the conduct of services in the educational contexts). Within this constellation of *data*, *meaning-making* and *praxis*, the roles of the

Figure 1.1. Conceptualizing Evaluation and Utilization in a Community of Educators



evaluator fundamentally involve a collaborating narrative, explicating the relationships and the processes (e.g., how data are reported and used by faculty to confirm and challenge principles and practices). In particular, this emphasizes the importance of the critical perspective of faculty and their role in meaning making, rather than presuming that data or outcomes have an automatic impact on practice (Figure 1.1). Although these dynamics—*policy, institutional applications, educator concerns, and data, meaning making, praxis*—are drafted here in higher education terms, they are frequently at the heart of evaluative reasoning more generally (cf., evaluation-specific logic, Davidson, 2014). Thus, these dynamics become unifying and interpretive considerations in the sections to follow.

Assessment in Higher Education

In the higher education literature, thought leaders use the term *assessment* (not *evaluation*) to describe the “systematic collection, review, and use of information about educational programs undertaken for the purpose of improving student learning and development” (Palomba & Banta, 1999, p. 4). The emphasis is on the program, which sets it apart from other common forms of educational evaluation such as grading individual students. Furthermore, accrediting organizations require that colleges and universities evaluate actual student performances or products (e.g., written papers, examinations, and oral presentations), and they state that survey instruments that gather student perceptions and self-reports are helpful but insufficient evidence of learning. Although advocates argue that faculty and campuses

should conduct assessment because it is beneficial, the reality is they often conduct and report on assessment because they must do so.

Regional and professional accrediting agencies require that faculty are involved in regular interpretive activities in which data on student learning are used in curriculum review and program decision making. All but one regional accreditor explicitly states that faculty be involved in the assessment processes (Provezis, 2010). This orientation eventually brings the evaluative analysis into the basic faculty planning processes regardless of faculty readiness or level of engagement and raises questions about how assessment and reporting might be used to advance intradepartmental discourse. In response, colleges and universities have increased the resources devoted to improving teaching and learning through assessment activities. Most research universities that took part in the Association of American Universities (2013) survey reported having a newly created office or a newly assigned person specifically responsible for student learning assessment compared to 5 years ago. Across the United States, evaluators have new opportunities in offices of evaluation, assessment, and educational effectiveness. Institutions have changed criteria for program review and new program proposals. They are spending funds on cloud-based assessment management software (e.g., Academic Effect, LiveText, Taskstream, and Tk20) and externally developed tests to meet accreditation requirements (e.g., ACT's Collegiate Assessment of Academic Proficiency, the Council for Aid to Education's Collegiate Learning Assessment, and Educational Testing Service's HEIghten Outcomes Assessment Suite). Philanthropic funding is now available to support organizations such as the National Institute for Learning Outcomes Assessment and initiatives such as the Multi-state Collaborative to Advance Learning Outcomes Assessment and the Valid Assessment of Learning in Undergraduate Education. Assessment leaders, educators, and accreditors have been writing books, organizing conferences and conference tracks/interest groups, and facilitating assessment academies. The growth in the last 30 years has been tremendous and provides evaluators with more employment options and more venues in which to study evaluation in context.

The history of the higher education "assessment movement" has been written about extensively, and we refer interested readers to texts such as *Learner-Centered Assessment on College Campuses* (Huba & Freed, 2000) and *Student Outcomes Assessment: A Historical Review and Guide to Program Development* (Sims, 1992), as well as Ewell (2002, 2008) and Wright (2002). From this history, we highlight the tension between assessment for accountability (external use of findings) and assessment for improving learning (internal use). The former is rooted in using standardized measurements to judge and compare colleges and universities. It promotes reductive accountability models that contribute to simplistic judgments and comparisons across institutions with radically different student populations and missions. (The latter comes from educational reform advocates seeking a

shift in focus from what professors teach to what students learn (e.g., Barr & Tagg, 1995; Huba & Freed, 2000; Study Group on the Conditions of Excellence in American Higher Education, 1984). It is formative and relies on internal reporting that presumes that faculty will use indicators of student learning to examine teaching and program concerns. A small number of colleges have a model in which assessment is regarded as a critical element in learning—that is, in performance, the learner makes a critical transfer to different situations that both consolidates and demonstrates the learning (cf., Abate, Stamatakis, & Haggett, 2003; Alverno College Faculty, 1994; Seaba, 2010).

These approaches have different implications for assessment and evaluators' practices. Assessment for accountability—a compliance-based model—locates the evaluation work outside of departmental discourse and inquiry. Evaluators use technologies that have been historically anchored in test and measurement traditions that can relatively quickly address concerns for substantial data and cross-institution comparisons. These tests, created by an external entity, are far from professors and classrooms, and they are fraught with issues related to validity, student motivation, and professor willingness to use findings in curricular decision making. Thus, the evaluators find themselves increasingly concerned with how judgments on student learning become part of the institutional and educational discourse on teaching and learning.

Evaluators who privilege assessment for improvement turn to a practice that integrates traditional measurement-oriented approaches with more participatory, deliberative, and developmental orientations to meet the needs of multiple stakeholders and satisfy both external accountability and internal improvement goals. Frequently cited examples in this area include the use of comprehensive exams or projects completed at program's end as performances that can be subject to an integrated analysis to examine overall program effectiveness, as well as specific elements (Segers, Dochy, & Cascallar, 2003). There are also examples of more structured action research projects focusing on a program's students for assessment and evaluation purposes. By sampling case studies of individual students or selecting specific assessments for departmental review, faculty can synthesize data in ways that serve a whole department in a critical analysis of program performance (Lava, Lehman, & Traugh, 2005). Campuses and evaluators may also use constructivist/mixed method approaches that combine performance assessment, available records, and qualitative approaches in order to address the realistic and complex needs of educators and academic units. Multiple measures along with tracking learning over time are two recommended practices in educational evaluation systems (American Evaluation Association, 2006).

Evaluation models that emphasize stakeholder perspectives and invest in structures of judgment that explicate indicators of value and worth can foster structures of meaning making that are essential to the interpretation

of evidence and true accountability. It is a development-based approach that involves dynamic and collaborative approaches in which evaluators and educators cooperate to carry out evaluation inquiry to move educational practices forward.

Evaluation models hold promise for assessment in higher education. Interestingly, the assessment literature and the evaluation literature have not intersected and would benefit from closer connections. For example, assessment experts do not often reference utilization-focused evaluation, evaluation capacity building, or theory of change and they do not acknowledge common tools such as the logic model. Evaluators and the field have much to offer higher education.

Challenges

Many challenges in evaluation cut across sectors: negative reactions, lack of leadership support, confusion about roles and purpose, inadequate resources, too much data to manage, confusion over how to analyze and use information, findings not used, etc. These challenges exist in higher education too, but we concentrate on several particular to education.

First, the issue of ethical treatment of college students is frequently raised. The U.S. federal regulations (Protection of Human Subjects policy and the Family Educational Rights and Privacy Act) are not barriers to assessment projects that involve standard education practice and use of results for internal improvement. However, individual faculty and departments may perceive them as obstacles. Their questions about ethical treatment and the protection of student information can derail a well-designed plan because of decreased participation, loss of trust between those with the student work and those collecting it, and because an insistence on anonymity hinders use of results (confidentiality is advised). Review by an institutional review board (IRB) and student informed consent are not required. Learning outcomes assessment does not meet a key criterion that triggers IRB review: assessment is not research “designed to develop or contribute to generalizable knowledge” (Protection of Human Subjects, 1991, section 46.102(d)). Informed consent is also not a requirement and the Family Educational Rights and Privacy Act (1974) allows transfer of student information between parties with legitimate interests. Because the concern for students is justified, colleges and universities need to engage in regular methods to protect student privacy and inform students of how the institution will use their work and data.

A second challenge is negative reaction by some faculty. The professors (i.e., program staff) in academic programs have publicly put forth more negative press than staff in any other sector in which program evaluation is common. Their message is clear from their article titles, for example, “The Great Assessment Diversion” (Kelly-Woessner, 2011) and “Outcomes Assessment: No Gain, All Pain” (Fryshman, 2007). Professors point to a lack

of proof that assessment activities lead to improvement as seen in Gilbert's 2015 piece, "Does Assessment Make Colleges Better? Who Knows?" and to concerns related to validity and reliability such as Hales's 2013 article, "Who's Assessing the Assessors' Assessors?" The external mandate for assessment is one impetus for these negative articles; Powell (2011) sums up much of the critics' positions:

Outcomes assessment is an odd business. It is not to the credit of higher education that we have tolerated this external assault on our work. Its origins are suspect, its justifications abjure the science we would ordinarily require, it demands enormous efforts for very little payoff, it renounces wisdom, it requires yielding to misunderstandings, and it displaces and distracts us from more urgent tasks, like the teaching and learning it would allegedly help. (p. 21)

Vocal groups of professors have not embraced the assessment movement and some are in positions of power that stall and derail even the most optimistic advocate or well-intentioned evaluator. The negative reaction leads to a lack of engagement and reluctance to use findings, which runs counter to an assessment for improvement model that relies on collaborative participation in evaluative practice. Evaluators contemplating a move into learning outcomes assessment need to be prepared for typical evaluation challenges and also for the sometimes blatant criticism of assessment, the reluctance to engage, and the other demands on faculty time as a result of decreases in education funding.

Third, the nature of the college context presents challenges to inquiry. Most student evaluation processes in college courses are developed by professors for specific course purposes. For program- and institutional-level assessment, data from course evaluations can be aggregated and used to address broader questions of student learning—termed in the assessment literature "embedded assessment"—and this can add to the complexity of analyses. Student achievement, can be reevaluated after a change in practice (e.g., a new course, innovative technologies) to increase outcome achievement, but it will often involve between-group designs complicated by other variables (e.g., student demographics, course taking patterns, and sample size). As in other educational settings and types of educational evaluation, randomized control trials are extremely challenging, if not impossible, to implement.

Learning From and Through Evaluation Practice in Higher Education

The results of assessment efforts have shown familiar characteristics in the developing professional practice of evaluation. Continued debates about what constitutes value and merit and evidence have often been contentious (cf., Davidson, 2010). And the question of appropriate use of learning

outcome assessment results becomes particularly significant in higher education (Banta & Pike, 2012a, 2012b). Indeed, just as in K–12 education, higher education faculty can have difficulties with conversations about data in the context of their own curriculum practices.

In the midst of current educational changes and evolving priorities, evaluators are finding ways to support public conversations about the worth and outcomes of higher education while contributing to the substantive educational concerns regarding learning, curriculum, and faculty expertise. This creates multiple tensions on the evaluators' work to provide dashboard indicators and maintain the expert meaning making of faculty who are guiding the coursework and curriculum. This means that evaluators need to understand and appreciate the discourse of those who use evaluative data—in this case, faculty—and work side by side to create trustworthy and meaningful assessments that enhance the educational program. Evaluators seek to anticipate how evaluation's multiple roles are having emerging impact in continuing development in pedagogy, curriculum, and equity and enrollment. To do so, they pay attention to the relationships developing between themselves and faculty and administrators and framing evaluation studies in existing, ongoing work of educational practice. In this regard, they are often challenged by situations where initiative and leadership seem critical to keeping the evaluation discussions evolving, even when these have been less recognized as part of evaluation practice. Because of this, higher education has been a context in which evaluation has made unique contributions and from which evaluation practices can be enhanced through continued development.

For example, this challenges a conceptualization of evaluation in which a singular study provides the fuel for data-driven decision making. Evaluation is an integrative activity in which value, merit, and evidence are used in complex and developmental activities to validate and inform, to think through and rethink our shared endeavors. In these activities, faculty, staff, and administrators become coinvestigators. Although evaluators may be charged with particular parts of study design, data collection, analysis, and certain levels of interpretation, it is the faculty who interpret outcomes into implementation. A sample of recent evaluation studies in higher education gives some insight into what higher education asks of evaluators and, in this regard, offers some insight to contexts in which evaluators work collaboratively with faculty and program staff to interpret program performance and to support ongoing development.

In conclusion, various forms of evaluation have long existed in higher education, but renewed attention to accountability has focused the spotlight on what students know and can do upon graduation (as well as on cost, access, and graduation rates, which are related to learning and teaching). Individual course grades, final grade point averages, and surveys of students' perceptions are now inadequate. Although new evaluation positions have been created and faculty access to evaluation expertise has increased,

pressure to conduct assessment often leads faculty to go it on their own as “accidental evaluators.” Thus, evaluators and participatory and collaborative evaluation approaches can help colleges and universities conduct better outcomes assessment. Subsequent research on evaluation in the higher education context can inform the field of evaluation, particularly other sectors with highly skilled professionals can benefit from studies of evaluators working with faculty experts who live in disciplinary subcultures.

References

- Abate, M. A., Stamatakis, M. K., & Haggett, R. R. (2003). Excellence in curriculum development and assessment. *American Journal of Pharmaceutical Education*, 67(3), 1–22.
- Alverno College Faculty. (1994). *Student assessment-as-learning at Alverno College* (3rd ed.). Milwaukee, WI: Alverno College.
- American Evaluation Association. (2006, November 1). *Public statement: Educational accountability*. Retrieved from <http://www.eval.org/p/cm/ld/fid=94>
- Association of American Universities. (2013). *Top research universities expanding efforts to assess, improve undergraduate student learning*. Washington DC: Author, Retrieved from <http://www.aau.edu/WorkArea/DownloadAsset.aspx?id=14849>
- Banta, T. W., & Pike, G. R. (2012a). Making the case against—one more time. In R. Benjamin (Ed.), *The seven red herrings about standardized assessment in higher education* (NILOA Occasional Paper No. 15, pp. 24–30). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.
- Banta, T. W., & Pike, G. R. (2012b). The bottom line: Will faculty USE assessment findings? In C. Secolsky & D. B. Dennison (Eds.), *Handbook on measurement, assessment, and evaluation in higher education* (pp. 47–56). New York, NY: Routledge.
- Barr, R., & Tagg, J. (1995). From teaching to learning—A new paradigm for undergraduate education. *Change: The Magazine of Higher Learning*, 27(6), 13–25.
- Chambliss, D. F. (2007, April 5). The flawed metaphor of the Spellings summit. *Inside Higher Education*. Retrieved from <https://www.insidehighered.com/views/2007/04/05/chambliss>
- Davidson, J. (2010, December 9). Pushing sand uphill with a pointy stick? “No value-free” in higher ed evaluation [Web log post]. Retrieved from <http://genuineevaluation.com/pushing-sand-uphill-with-a-pointy-stick-no-value-free-in-higher-ed-evaluation/>
- Davidson, J. (2014). *Evaluative Reasoning, Methodological Briefs: Impact Evaluation 4*. Florence: UNICEF.
- Ewell, P. T. (2002). An emerging scholarship: A brief history of assessment. In T. W. Banta & Associates (Eds.), *Building a scholarship of assessment* (pp. 3–25). San Francisco, CA: Jossey-Bass.
- Ewell, P. T. (2008). Assessment and accountability in America today: Background and context. In V. M. H. Borden & G. R. Pike (Eds.), *New Directions for Institutional Research: No. 51. Assessing and accounting for student learning: Beyond the Spellings Commission* (pp. 7–18). San Francisco, CA: Jossey-Bass
- Family Educational Rights and Privacy Act of 1974, 20 C.F.R. §1232g et seq.
- Fryshman, B. (2007, November 13). Outcomes assessment: No gain, all pain. *Inside Higher Education*. Retrieved from <https://www.insidehighered.com/views/2007/11/13/fryshman>
- Gilbert, E. (2015, August 15). Does assessment make colleges better? Who knows? *Chronicle of Higher Education*. Retrieved from <http://chronicle.com/article/Does-Assessment-Make-Colleges/232371/>

- Hales, S. (2013, March 11). Who's assessing the assessors' assessors? *Chronicle of Higher Education*. Retrieved from <http://chronicle.com/article/Whos-Assessing-the-Assessors/137829/>
- Huba, M. E., & Freed, J. E. (2000). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning*. Needham Heights, MA: Allyn & Bacon.
- Kelly-Woessner, A. (2011, February 13). The great assessment diversion. *Chronicle of Higher Education*. Retrieved from <http://chronicle.com/article/The-Great-Assessment-Diversion/126347/>
- Kezar, A. (2014). *How colleges change: Understanding, leading, and enacting change*. New York, NY: Routledge.
- Kuh, G. D., Ikenberry, S. O., Jankowski, N. A., Cain, T. R., Ewell, P. T., Hutchings, P., & Kinzie, J. (2015). *Using evidence of student learning to improve higher education*. San Francisco: Wiley & Sons.
- Kuh, G. D., Kinzie, J., Schuh, J. H., & Whitt, E. J., & Associates. (2005). *Student success in college: Creating conditions that matter*. Washington DC: American Association for Higher Education.
- Lava, V., Lehman, L., & Traugh, C. (2005). Re-imagining one urban teacher education program: Transformation through inquiry, difference and field experience. In P. M. Jenlink & K. E. Jenlink (Eds.), *Portraits of teacher preparation: Learning to teach in a changing America* (pp. 73–94). Lanham, MD: Rowman & Littlefield Education.
- Palomba, C. A., & Banta, T. W. (1999). *Assessment essentials*. San Francisco, CA: Jossey-Bass.
- Powell, J. W. (2011). Outcomes assessment: Conceptual and other problems. *AAUP Journal of Academic Freedom*, 2(2), 1–25.
- Project DEEP: Documenting effective educational practice. (n.d.). Retrieved from <http://nsse.indiana.edu/html/projectDEEP.cfm>
- Protection of Human Subjects. (1991). 45 C.F.R. §46 et seq.
- Provezion, S. (2010). *Regional accreditation and student learning outcomes: Mapping the territory*. Champaign, IL: National Institute for Learning Outcomes Assessment.
- Seaba, H. H. (2010). Assessment in the College of Pharmacy. Retrieved from http://pharmacy.uiowa.edu/files/pharmacy.uiowa.edu/files/wysiwyg_uploads/assessment_in_the_college_of_pharmacy_1999-2010.pdf
- Secretary of Education's Commission on the Future of Higher Education. (2006). *A test of leadership: Charting the future of U.S. higher education*. Washington, DC: U.S. Department of Education.
- Segers, M., Dochy, F., & Cascallar, E. (2003). *Optimising new modes of assessment: In search of qualities and standards*. Dordrecht, The Netherlands: Kluwer Academic.
- Sims, S. J. (1992). *Student outcomes assessment: A historical review and guide to program development*. New York, NY: Greenwood Press.
- 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. Department of Education, National Institute of Education. Retrieved from <http://files.eric.ed.gov/fulltext/ED246833.pdf>
- Wright, B. D. (2002). Accreditation and the scholarship of assessment. In *Building a scholarship of assessment* (pp. 240–258). San Francisco, CA: Jossey-Bass.
- Zemsky, R. (2007). The rise and fall of the Spellings Commission. *Chronicle of Higher Education*, 53(21), B6.

WILLIAM H. RICKARDS is a consultant at Independent Evaluation Inquiry.

MONICA STITT-BERGH is an associate specialist of assessment at the University of Hawai'i at Mānoa.