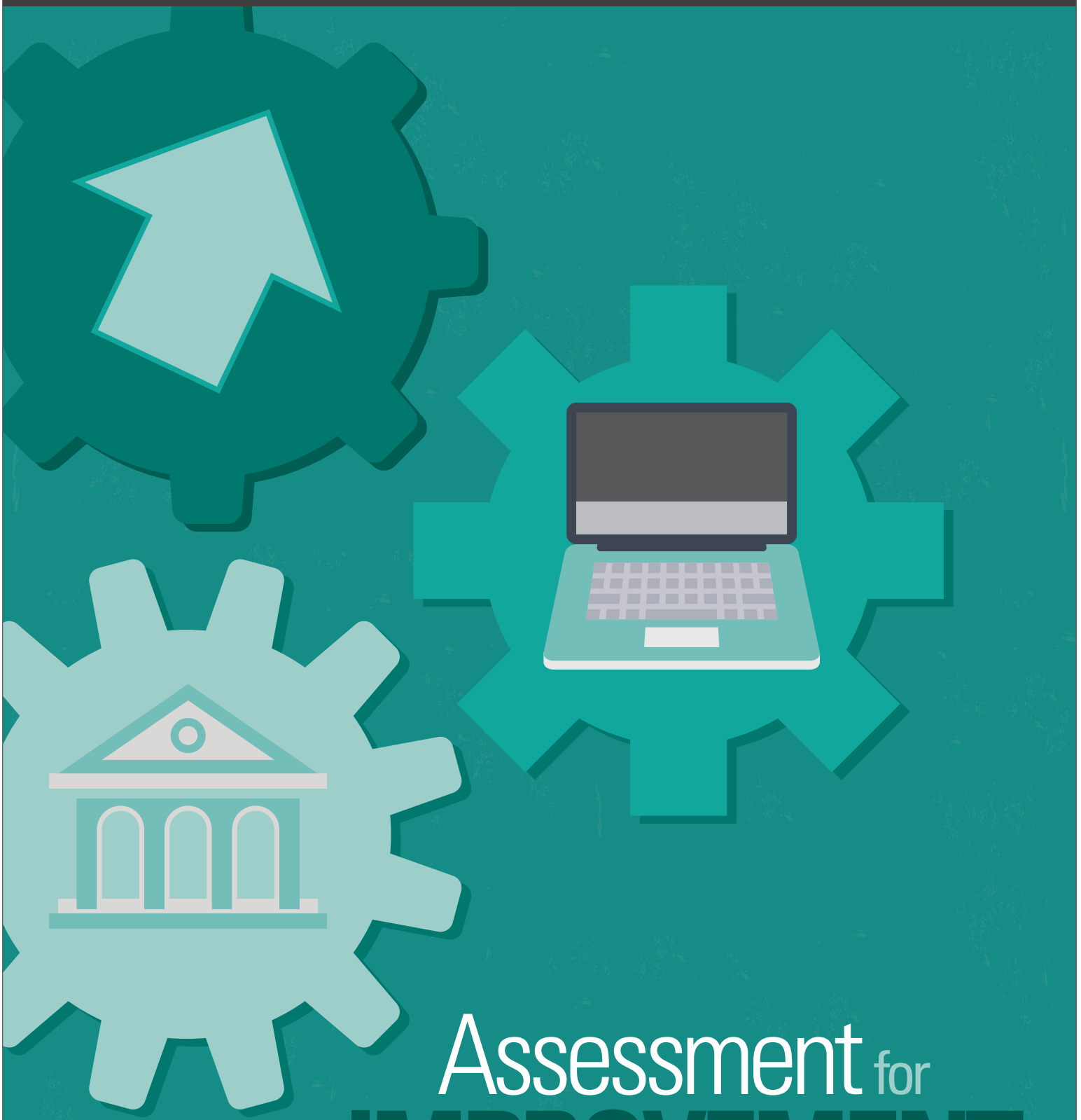

Assessment for IMPROVEMENT

ETS

Measuring the Power of Learning.®

ETS Research Report Series

Since its 1947 founding, ETS has conducted and disseminated scientific research to support its products and services, and to advance the measurement and education fields. In keeping with these goals, ETS is committed to making its research freely available to the professional community and to the general public. Published accounts of ETS research, including papers in the ETS Research Report series, undergo a formal peer-review process by ETS staff to ensure that they meet established scientific and professional standards. All such ETS-conducted peer reviews are in addition to any reviews that outside organizations may provide as part of their own publication processes. Peer review notwithstanding, the positions expressed in the ETS Research Report series and other published accounts of ETS research are those of the authors and not necessarily those of the Officers and Trustees of Educational Testing Service.

The Daniel Eignor Editorship is named in honor of Dr. Daniel R. Eignor, who from 2001 until 2011 served the Research and Development division as Editor for the ETS Research Report series. The Eignor Editorship has been created to recognize the pivotal leadership role that Dr. Eignor played in the research publication process at ETS.

ETS Research Report Series ISSN 2330-8516

RESEARCH REPORT

Continuing a Culture of Evidence: Assessment for Improvement

Javarro Russell & Ross Markle

Educational Testing Service, Princeton, NJ

From 2006 to 2008, Educational Testing Service (ETS) produced a series of reports titled *A Culture of Evidence*, designed to capture a changing climate in higher education assessment. A decade later, colleges and universities already face new and different challenges resulting from societal, technological, and scientific influences. This new series of reports describes current issues and trends in assessment and how they relate to the needs of higher education institutions. In this report, we discuss the shift from accountability-based assessment to improvement-based assessment. When conducted for improvement, assessment is more inclusive of institutional faculty and staff, more informative to colleges and universities, and—it is hoped—focused on enhancing instruction and learning. However, these positive outcomes do not come without challenges. Ultimately, assessment-for-improvement has implications for many of the people, processes, and tools used to measure student learning.

Keywords Assessment; student learning outcomes; higher education; improvement; accountability

doi:10.1002/ets2.12136

Over the past three decades, learning outcomes assessment has come to play a consequential role in higher education (Kuh et al., 2015). Driven at first by growing calls for accountability from key stakeholders, assessment theorists and practitioners have yielded credible measures of institutional level or academic program level performance on critical learning outcomes (Klein et al, 2009; Liu, Frankel, & Roohr, 2014; Shavelson, 2010; Sullivan & Schneider, 2015). In recent years, however, accreditation agencies and other stakeholders have widened their interests beyond the extent to which students have attained the learning outcomes specified by their institution. These stakeholders' interests have progressed to include the improvement of student learning through the use of previously obtained evidence of student learning (Provezis, 2010).

Assessment of student learning outcomes in higher education is a process for gathering information to make inferences about groups of students across the institution. These inferences can then be used to guide decisions for improving teaching and learning. Existing assessment frameworks differ in the extent to which they support institutions in obtaining information that can help improve the impact of educational experiences. Some frameworks merely provide the data required for documenting attainment of learning or "checking the box." Reconsidering assessment as an integrated, iterative, institutional process may provide higher education institutions (HEIs) with an opportunity to more directly engage faculty in improvement activities, and doing so requires a shift away from assessment as a compliance activity. It also requires new considerations for assessment development, delivery, and use.

This report outlines the changes that we believe are needed on the part of institutions, institutional faculty, researchers, the developers of assessments, and policymakers to ensure that the way assessments are used continues to evolve in ways that lead to improvements in teaching and learning and ultimately to better learning outcomes for students.

The Evolving Role of Assessment in Higher Education

It is an implicit assumption of the college experience that students will develop appropriate knowledge, skills, and behaviors (together referred to as *learning outcomes*) as a result of engagement in curricular and cocurricular activities. Along with gaining in-depth knowledge of the content associated with their chosen program of study, students are also expected to experience cognitive, psychosocial (or noncognitive), and attitudinal changes (Astin, 1993)—changes that are an important part of the broader institutional mission (Oswald, Schmitt, Kim, Ramsay, & Gillespie, 2004).

Corresponding author: R. Markle, E-mail: rmarkle@ets.org

Over the last 30 years, the need to verify this assumption through the measurement of student learning outcomes has risen, largely due to the evolving and increasing pressure from stakeholders such as state and federal governments, accreditation agencies, students, parents, and society at large (Ewell, 2009). The need to investigate institutional effectiveness has further been suggested by the results of large-scale assessment projects, as illustrated in the book *Academically Adrift* (Arum & Roksa, 2011).

Peter Ewell (2009) referred to the first calls for increased demonstration of institutional effectiveness, which took place in the early 1980s, as the *assessment movement*. Two themes have characterized this movement. First, early efforts focused on the use of more direct measures of student learning. Previous indicators of institutional quality, such as acceptance rates and graduation rates, had been criticized as invalid indicators of student learning, as they provided no evidence of the changes or learning that occurred at institutions (Ewell, 2001). Most of the emphasis on direct measures of learning, however, accompanied accountability-based assessment paradigms. Unfortunately, while encouraging institutions to participate in gathering data on student learning, efforts that focused on complying with accountability efforts unwittingly devalued the use of outcomes data for institutional or program improvement (Hutchings, 2010). As the demonstration for accountability became more synonymous with and attainable by the sharing of data on student learning outcomes, other uses of the data became an inconvenient afterthought.

The second theme of the assessment movement has been the shift to improvement-based efforts. Although the improvement of teaching and learning has always been a consideration of assessment, it has moved to the forefront in recent years thanks to several factors (Ewell, 2009). For one, whereas assessment remains, as Ewell (2009) stated, "distasteful for many faculty members, it is angrily rejected far less frequently than it was 20 years ago" (p. 6). Although assessment is not universally accepted across all institutions, faculty members have become increasingly responsive to assessment practices and the use of results, and their voice is certainly critical to the success of nearly all institutional change efforts (Kuh et al., 2015). Additionally, accreditors have since become the primary evaluator of HEIs. Although early assessment efforts existed within state-level accountability mechanisms, accreditors are now responsible for evaluating assessment practice, and some have recently placed explicit value on an institutions' ability to use data for the improvement of teaching and learning (Western Association of Schools and Colleges [WASC], 2014). Projects such as the Higher Learning Commission's (2016) Academic Quality Improvement Program, the WASC's Senior Commission, and the Southern Association of Colleges and Schools' (2011) Quality Enhancement Program demonstrate accreditors' evolving preference for the use of data to improve teaching and learning over some overall decontextualized presentation of data on student learning.

This more recent shift toward improvement-based assessment requires changes in both the purpose and the practice of assessment processes in order to drive meaningful changes in teaching and learning. In this section, we first discuss two assessment efforts that epitomize the accountability and improvement assessment paradigms. We then discuss the key components of effective assessment systems that can drive institutional improvement. Finally, we review the implications for key stakeholders of higher education and the actions they can take to ensure that assessment continues to serve as a valuable tool for improving teaching and learning before we turn to the next step in the evolution of assessment systems—an increased focus on individual student outcomes—in the second section of this report.

The Voluntary System of Accountability and the Valid Assessment of Learning in Undergraduate Education: Two Systems, Two Paradigms

Two of the most extensive efforts to engage HEIs and their stakeholders in the creation and use of assessment typify the differences between accountability-based assessment and the improvement-based paradigm. The Voluntary System of Accountability (VSA) is perhaps the most well-known effort to engage institutions and their stakeholders in the dissemination of student learning outcomes assessment data (Jankowski et al., 2012; McPherson & Shulenburger, 2006). It also epitomizes the use of assessments for accountability purposes and illustrates the limitations of this model in improving teaching and learning.

The VSA emerged as a result of calls for more direct measures of student learning. Although the Spellings Commission's 2006 report, *A Test of Leadership: Charting the Future of U.S. Higher Education* (U.S. Department of Education, 2006), is the most cited example of such calls, the VSA and the commission report were actually in development concurrently. The VSA provided a structure for institutions to identify measures and report results in critical thinking, analytic reasoning,

and written communication. In developing this reporting structure and promoting an easily accessible website for student learning outcomes data, the VSA aligned significantly with the accountability and transparency narrative that permeated the commission's report.

A key feature of the VSA was the use of common measures, which sought to ensure the interpretability and comparability of the data provided by institutions. The use of these common measures was supported by the VSA's Test Validity Study (Klein et al., 2009), which obtained validity evidence for established measures of critical thinking and written communication accepted by the VSA. This major step forward highlighted not only comparability, but also other test properties that would be relevant to users of these assessments. Providing this information helped to quell some critics' concerns over what was being measured by each of the assessments and how the assessments differed from each other. Additionally, and perhaps more central to its purpose, the VSA provided a structure that allowed institutions to share their assessment-related data with the public.

The lens through which the VSA envisioned assessment was one in which the provision of information was the foremost means to demonstrating accountability. The website the VSA constructed for presenting learning outcomes data allowed institutions to provide little to no indication of how the data were used to inform improvements to the curriculum and/or other educational activities (Jankowski et al., 2012). Instead, the VSA portal allowed for the provision of student learning outcome data without any institution-specific information that would provide some contextualized understanding of their assessment process. This lack of context rendered the project insufficient for promoting learning outcomes assessment as a process that is connected to the teaching and learning that occur at each institution.

Whereas the VSA used its measurement tools to focus on the accountability aspects of the assessment process, other efforts sought to provide tools that institutions could use or adapt locally as a means to inform the improvement of teaching and learning. In this way, institutions would own the data as part of their assessment process, and the focal comparison would be internal (i.e., for improvement) rather than external (i.e., for accountability). One notable example is the Association of American Colleges & Universities' Liberal Education and America's Promise (AAC&U, 2011). This effort not only defined key areas of learning in higher education but also offered the Valid Assessment of Learning in Undergraduate Education (VALUE) rubrics as tools to measure the domains of civic engagement, creative thinking, critical thinking, ethical reasoning, information literacy, inquiry and analysis, integrative learning, intercultural knowledge and competence, lifelong learning, oral communication, problem solving, quantitative literacy, reading, teamwork, and written communication.

Ultimately, tools such as the AAC&U VALUE rubrics provide an opportunity for institutions to obtain assessment results by embedding faculty members and students directly into the process of assessment. Doing so provides an opportunity for buy-in from faculty members and the communication of results with students. Unlike standardized selected-response type assessments, the VALUE rubrics allow faculty members to modify the rubrics to suit their individual needs. For example, Rhodes and Finley (2013) noted three areas in which intuitions are making modifications to the rubrics. Institutions may desire to modify the language used to describe the performance criteria to fit the institution's common language regarding assessment. Another modification may be to include additional criteria required for demonstrating the learning outcome being measured. Lastly, intuitions may create additional performance levels that further delineate student knowledge and skill on the outcomes. These options for implementing and modifying these rubrics allow faculty members to be involved in the assessment of student learning outcomes in ways that differ from implementation of standardized assessments.

An important focus of AAC&U's continued rubric development activities is standardizing its use among faculty members across institutions. There is also an ongoing effort to validate their use in improving teaching and learning, which is at the crux of a meaningful assessment process (Rhodes, 2008). These qualities are part of the reason the VALUE rubrics were included in 2012 as one of the few accepted assessments in the VSA (Keller, 2014).

Purposes of Assessment

In our description of the VSA and the VALUE rubrics, we present two distinct purposes of assessment. It is clear that the purposes differ in ways that reflect the engagement of institutions in the assessment process. Institutions that are focused on accountability rely on reporting information on student learning outcomes as the means to an end. However, institutions that utilize the process of assessment to engage faculty members and students in answering relevant questions

about student learning and applying the answers of those questions to improve student learning have set the stage for an improvement-based paradigm on their campus.

In an improvement-based paradigm of assessment, effective assessment practices require the use of evidence of student learning to inform decisions and to ultimately make improvements in student attainment of positive educational outcomes (Ewell, 2008, 2009; Kuh & Ikenberry, 2009). As Peter Ewell has repeatedly pointed out, the key underlying distinction is that improvement-based paradigms use all aspects of the assessment process to identify and act upon areas that require improvement, whereas accountability-based paradigms only implement a portion of the assessment process to collect and disseminate data as necessary to demonstrate success. Within the accountability paradigm, assessment efforts focus on collection and dissemination without giving much attention to the uses or, more importantly, the users of the results, such as faculty members and students. This practice leads to a disrupted or incomplete assessment process (Ewell, 2009).

The early promulgation of the accountability paradigm by stakeholders external to HEIs has led to assessment efforts that often focus on the types of assessments used, the nature of the assessment conducted (e.g., value-added or availability of benchmarks), and the quality of the data obtained (Kuh et al., 2015). By disregarding the actual use of data, the validity of inferences that should be drawn, and the implementation of evidence-based decisions, these stakeholders helped to spur an assessment environment in which the measure of student learning only possesses value in its ability to allow institutions to check an accreditation box and nothing more. Although meeting accreditation requirements is not negligible, doing so for the sole sake of compliance often leads to institutional disengagement in activities that focus on the improvement aspects of assessment. This disengagement is mostly due to the fact that compliance activities are episodic, whereas improvement activities are ongoing and can be resource intensive (Kuh et al., 2015).

Assessment for the sake of accountability is typical of institutions attempting to relieve administrative pressure to adhere to the requirements of regional or state-level agencies. When this type of assessment is conducted, rarely are the results shared across the institution or used for improvement of instructional activities, courses, or academic programs (Palomba & Banta, 1999). If they are shared, Palomba and Banta (1999) pointed out that less than ideal results are met with distrusting scrutiny, and the assessment or its implementation (e.g., sample size)—rather than the underlying learning phenomena or instruction—is questioned.

Institutions seeking to use data to inform learning improvement—either at the course, curriculum, program, or institutional level—understand that the assessment process includes more than gathering data and reporting to stakeholders; that is, assessment is a process and not a thing. This process includes an internally driven evidentiary inquiry into what students know, think, and can do given the educational programming provided by the institution (Fulcher, Good, Coleman, & Smith, 2014; Hutchings, 2011). Considering assessment as an integrated institutional process allows institutions to be accountable not only for obtaining information about student learning but also for responding to the information in ways that help improve teaching and learning. Without the context of the educational programming provided by the institution and student population, assessment data are stagnant numbers, not dynamic pieces of information that can be used to inform action (Kuh et al., 2015).

Instead of focusing solely on checking the box, assessment for improvement requires commitment to a genuine inquiry into an institution's effectiveness in improving the learning outcomes of students. Though this type of assessment may be initiated by external forces, it is usually driven and sustained by an institution's preparation and resolve (Fulcher et al., 2014).

Implications and Recommendations

Creating effective assessment systems will require significant changes not only within institutions but also among the stakeholders that helped frame the accountability-based approach to assessment and now play crucial roles in the shift toward improvement-based models.

Implications for Colleges and Universities

Within institutions, the shift toward an improvement paradigm is heavily dependent upon creating concurrent cultural shifts in how assessment is perceived and used.

Boosting Engagement

External entities have long influenced the involvement of institutions in the assessment process. However, for many institutions, the policies of these external entities lack the motivational appeal for faculty members and staff members to engage in the assessment process. This lack of extrinsic motivation has led to less than adequate engagement in the assessment process (Cain & Hutchings, 2015). There are few rewards for those who participate in the process. Faculty members rarely apply for or are promoted to lead assessment efforts. Instead they are deputized by previous leaders of the effort at their institution. Also, there are currently no educational or certifiable requirements of individuals involved in the process of assessment at the institution. These conditions amplify doubt among new participants in the assessment process and skepticism from those who have seen the assessment tide rise and fall. This mix of circumstances can weaken attempts to organize assessment activities on a campus and encourages institutions to engage in "checking the box" activities to satisfy accreditation demands, all contributing to institutional apathy toward assessment (Hutchings, 2010).

Institutions must find ways to cultivate the engagement of faculty members and staff members in the assessment process. In their recommendations, Kuh et al. (2015) presented active efforts to create a positive institutional attitude and culture around assessment by formally and informally recognizing its importance. Institutions can devote resources such as release time or curriculum study grants to programs that are engaging in assessment development tasks. Faculty members understand that resources, including time, funds, or accolades, are committed to activities the institution values. Providing these resources creates value for their efforts in assessment.

Institutions can also encourage campus leaders at both the academic and administrative levels to communicate openly about the importance of assessment and maintain some flexibility in how assessment is organized and executed. Empowering faculty members and administrators to discuss assessment allows for the exchange of ideas and the development of mutual understanding around assessment efforts. On a related note, Pat Hutchings (2010) noted in an occasional paper for the National Institute of Learning Outcomes Assessment (NILOA) that the increased prominence of the scholarship of teaching and learning presents a glaring opportunity to engage faculty members in assessment. Promoting this type of scholarship may lead not only to further engagement through the dissemination and sharing of ideas but also to increased ownership of the results of assessment.

Developing Assessment Competencies

Institutions can aid in the improvement of teaching and learning by providing faculty and staff members opportunities to gain the skills necessary to develop and implement quality assessment processes. An investment in assessment-related knowledge and skills is critical for institutions looking to promote improvement-focused assessment. Institutions can seek professional development workshops and other trainings in areas of assessment that allow entire academic programs to engage in best practices of assessment. By attending conferences, workshops, or training activities that focus on assessment, faculty members gain knowledge from experienced practitioners in assessment and share ideas, successes, and challenges with one another. Regional accreditors typically provide assessment workshops at their annual conferences. The same is done by accreditors of academic programs, including the Association to Advance Collegiate Schools of Business and the American Society for Engineering Education. Suskie (2009) provided a list of conferences that are focused on higher education assessment.

Another route to increasing the assessment capabilities of faculty members is to implement the theory and practice of student learning outcomes assessment into the curriculum or professional development of graduate studies programs. The Council for Graduate Schools (CGS) is encouraging graduate programs to provide graduate students the opportunity learn about the process of assessing and improving student learning. With the help of funding from the Alfred P. Sloan Foundation and the Teagle Foundation, the CGS has leveraged its Preparing Future Faculty program to educate graduate students on the use of assessment to improve student learning in higher education. Their model for developing appropriate expertise in assessment for future faculty members provides a viable option for improving assessment practice in higher education (Deneef, 2002; Tinto, 2012)

Implications for Researchers and Assessment Developers

In their contribution to a comprehensive review of assessment in higher education, Kuh et al. (2015) discussed various types of student learning measures (e.g., surveys, standardized measures, portfolios) and how those measures rate as pieces

of evidence across various constituents. They also identified three primary obstacles to effectively using data, information, and evidence to improve student learning: (a) faculty members and administrators fail to access assessment information, (b) evidence is accessed but not acted upon, and (c) actions do not lead to notable improvement in learning or instruction. Certainly, there is a host of reasons why data use might fail to occur, including faculty members' attitudes about assessment data, institutional structures to disseminate assessment results, or the willingness to implement change. However, there are steps that can be taken by those who research and develop assessments of student learning that might minimize these obstacles, which include the following.

Demonstrating Alignment

Assessment for improvement cannot occur without a detailed understanding of the outcomes institutional stakeholders hope to improve. By conducting a needs analysis, those responsible for selecting or developing assessments can identify the learning outcomes that are most important to institutions. The clarity of these learning outcomes will allow researchers and assessment developers to identify the measures to best assess their learning outcomes.

Establishing a clear alignment between the knowledge and skills measured by assessments and those valued by institutions could be helpful in promoting assessment for improvement purposes. Hutchings, Kinzie, and Kuh (2015) cited evidence from a NILOA survey that shows that faculty members often have concerns about the alignment between assessment tools and their curriculum or classroom instruction. Even though assessment development efforts often include research into commonly valued skills (e.g., Markle, Brenneman, Jackson, Burrus, & Robbins, 2013), we believe researchers and assessment developers need to explicitly communicate the link between the information produced by an assessment and the outcomes that are valued by faculty members. It is important to note that this applies to local assessment development efforts as well as large, externally developed measures (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). Without a clear alignment among assessment, learning, and instruction, faculty members and administrators will be less likely to make valid use of assessment results for improvement.

Score reporting is the primary means by which assessment data and other information are conveyed to institutional stakeholders. The assessment data and information provided at the institutional level should convey the alignment between the learning objectives and the assessment. Rather than limiting the information provided to global skills, such as critical thinking, more clear definitions or examples of those constructs might help make the reported information more useful to stakeholders charged with making improvements. If possible and appropriate to provide, subscores could give this more detailed information that faculty members can relate to specific instructional activities, thus strengthening the tie between assessment and instruction. Research has shown that developing score reports requires the consideration of separate needs of score user groups (Zapata-Rivera & Katz, 2014). By conveying clear and unambiguous information needed by the intended users of the score reports, developers may improve the utility of these score reports for the process of improving student learning (Zenisky & Hambleton, 2012).

Providing More Meaningful Scores

Often, assessments results highlight norm-referenced scores that indicate a student's or an institution's relative performance to other students' or institutions' performance. These results are commonly illustrated through the use of percentile ranks or histograms. For example, externally developed standardized assessments will provide scaled scores that allow institutions to make relative comparisons amongst their own students as well as comparisons with students from other institutions. Within an institution, there may be a desire to compare the performance of student cohorts that experience different curricular or cocurricular experiences. Investigating differences among cohorts of students over time is one way institutions can use this normative data. Similarly, institutions may seek to identify similarities between students within or across cohorts that typically perform at relatively lower levels than their peers. These students may have academic or demographic similarities that could be targeted in the institution's search for options to interventions to improve student learning.

Although normative comparisons have some utility in making improvements at the institutional or program level, faculty members are often focused on how students relate to desired learning goals rather than to other students. Criterion referenced scores can provide further details on the extent to which students demonstrate the knowledge or skills

being assessed. By providing performance level descriptors, faculty members can focus on the learning demonstrated by their students instead of scaled scores. Several researchers have noted the value of criterion-referenced scoring in formative assessment and improvement-focused assessment models (e.g., O'Donovan, Price, & Rust, 2001; Rust, Price, & O'Donovan, 2003; Wilson & Scalise, 2006). When using criterion-reference descriptors of performance, faculty can focus on the number of students who have obtained proficiency or mastery, rather than the performance of their students relative to others. This information can still be used to compare performance over time and in the aforementioned ways noted under norm-referenced comparisons.

Ultimately, the meaningfulness and perhaps the usefulness of scores, whether norm referenced or criterion referenced, will hinge upon the users' understanding of how scores relate to the curricular or cocurricular activities in which students engage. If students are expected to meet institutional learning outcomes as a result of the educational activities, then these activities are likely to be modified if assessment results are suggestive of needed improvement. As HEIs increasingly seek to enact changes that can improve teaching and learning, having access to assessments that are valid for the purposes for which institutions intend to use the scores is of upmost importance. Having clear and accurate score reports is equally important, given that they are usually the vehicle through which HEIs interact with assessment data.

Implications for Public Policy

Agents of organizations such as state systems, accreditors, legislators, and other policy makers — who largely set the context in which educational systems exist — have an important role in continuing to promote and foster improvement-focused assessment. As mentioned, this has already been done to a large extent (e.g., WASC, 2014). As institutional accountability has moved mainly to regional accreditors, there has been less interest on their part in evaluating institutional achievement (e.g., ranking assessment scores) and instead an increased focus on institutional improvement over time. That is, rather than comparing the assessment scores of one institution to another, evaluations are made based on the ability to use scores to make meaningful changes.

In order to maintain and further the progress that has been made by HEIs that seek to use assessment results for improvement, future models of accountability should explicitly include the use of results and continuous improvement as criteria for evaluation. However, other policy initiatives with conflicting approaches persist. While the VSA is currently undergoing a shift to an improvement-focused model, projects such as the Organisation for Economic Co-operation and Development's *Assessment of Higher Education Learning Outcomes* (see Tremblay, Lalancette, & Roseveare, 2012), which seek to compare institutions on a common assessment or set of assessments, will naturally foster accountability-focused assessment. Moreover, the College Ratings and Paying for Performance Plan, recently proposed by the U.S. Department of Education (2015), initially excluded any measures of student learning, let alone the consideration of assessment-based improvement. The department's proposed model, which focused on access, affordability, and enrollment outcomes (e.g., graduation and transfer rates), or similar models proposed in the future, could limit the progress made thus far. This program was ultimately abandoned for a consumer-facing data tool (Fain, 2015; Shear, 2015), but if institutions focus on these kinds of proposed metrics, they may ignore the matter that is at the core of their missions: supporting and enhancing student learning.

Conclusion

An improvement-focused assessment process includes stating learning outcomes, selecting or developing measures of those learning outcomes, collecting data from those measures, analyzing and interpreting the results, communicating those results to stakeholders, and finally using stakeholder input to improve student learning. This process repeats itself to ensure that improvement efforts are effective. Similar to a basic research design, this process feeds on specific questions or hypotheses about the achievement of student learning outcomes.

The key components of assessment processes that focus on improvement have been discussed at length in Banta, Jones, and Black's (2010) *Designing Effective Assessment: Principles and Profiles of Good Practice*, Suskie's (2009) *Assessing Student Learning*, and most recently Kuh et al.'s (2015) *Using Evidence of Student Learning to Improve Higher Education*. Some of the key aspects of assessment processes focused on the improvement of student learning include the following.

- *Student learning outcomes are made explicit and reflect the true and distinct interests of the institution or program.* Much has been written about best practices in explicitly articulating learning outcomes (Adelman, 2015). These

learning outcomes are, to a large degree, developed based on the expectations and interests of faculty members and administrators who have intimate knowledge of the educational activities they provide and the students they serve.

- *The design of the assessment process allows the institution to gain reliable, valid, and trustworthy data to inform broad action in response to internally driven inquiries about student learning*. Each aspect of the assessment design, from selecting measures to using results, aligns with the institution's purpose for conducting assessment.
- *Resources are available to encourage the sustained engagement of faculty members and administrators in the process of assessment*. These include opportunities for assessment-related professional development, internal structures for communication of campus-wide assessment activities, and professional incentives that demonstrate the value institutions place on time dedicated to assessment (Huber, Hutchings, & Gale, 2005; Hutchings, 2010; Maki, 2002). Although increasing the resources to improve assessment practices is important, opportunities to improve teaching are also necessary. Improved teaching practices may be helpful in producing educational activities that can help students improve their learning.
- *Students are engaged in assessment, not in a compliance manner, but in a way that demonstrates the integration of assessment into the learning process at the institution*. Just as institutions seek to use learning outcomes data to improve educational activities, students may be able to use similar information to improve their own attainment of learning outcomes (Metcalf & Kornell, 2007).

Each of these aspects is neither easily implemented nor cultivated at institutions, given the complex cultures and distinct drivers of assessment at the institutional level. However, with support from policymakers, researchers, and assessment developers, institutions can see these aspects as goals that will allow them to shift from an accountability-based assessment program to a much more internally useful improvement-based assessment program.

References

Adelman, C. (2015). *To imagine a verb: The language and syntax of learning outcomes statements* (Occasional Paper No. 24). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago, IL: University of Chicago Press.

Association of American Colleges and Universities. (2011). *The LEAP vision for learning: Outcomes, practices, impact, and employers' views*. Washington, DC: Author.

Astin, A. W. (1993). *What matters in college: Four critical years revisited*. San Francisco, CA: Jossey-Bass.

Banta, T. W., Jones, E. A., & Black, K. E. (2010). *Designing effective assessment: Principles and profiles of good practice*. New York, NY: Wiley.

Cain, T. R., & Hutchings, P. (2015). Faculty and students: Assessment at the intersection of teaching and learning. In G. D. Kuh, S. O. Ikenberry, N. A. Jankowski, T. R. Cain, P. T. Ewell, P. Hutchings, & J. Kinzie (Eds.), *Using evidence of student learning to improve higher education* (pp. 95–116). San Francisco, CA: Jossey-Bass.

Deneef, A. L. (2002). *The preparing future faculty program: What difference does it make?* Washington, DC: Association of American Colleges and Universities.

Ewell, P. T. (2001). *Accreditation and student learning outcomes: A proposed point of departure*. Retrieved from http://www.chea.org/award/StudentLearningOutcomes2001.pdf

Ewell, P. T. (2008). Assessment and accountability in America today: Background and context. In V. M. H. Borden & G. Pike (Eds.), *Assessing and accounting for student learning: Beyond the Spellings Commission* (pp. 7–18). San Francisco, CA: Jossey-Bass.

Ewell, P. T. (2009). *Assessment, accountability, and improvement: Revisiting the tension* (NILOA Occasional Paper No. 1). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Fain, P. (2015, June 25). Ratings without . . . ratings. *Inside Higher Ed.* Retrieved from https://www.insidehighered.com/news/2015/06/25/education-department-says-rating-system-will-be-consumer-tool-rather-comparison

Fulcher, K. H., Good, M. R., Coleman, C. M., & Smith, K. L. (2014). *A simple model for learning improvement: Weigh pig, feed pig, weigh pig* (NILOA Occasional Paper No. 23). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Higher Learning Commission. (2016). *AQIP pathway overview*. Retrieved from https://www.hlcommission.org/Pathways/aqip-overview.html

Huber, M. T., Hutchings, P., & Gale, R. (2005). Integrative learning for liberal education. *Peer Review*, *7*(3), 4–7.

Hutchings, P. (2010, April). *Opening doors to faculty involvement in assessment* (NILOA Occasional Paper No. 4). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Hutchings, P. (2011). From department to disciplinary assessment: Deepening faculty engagement. *Change, 43*(5), 36–43.

Hutchings, P., Kinzie, J., & Kuh, G. D. (2015). Evidence of student learning: What counts and what matters for improvement. In G. D. Kuh, S. O. Ikenberry, N. A. Jankowski, T. R. Cain, P. T. Ewell, P. Hutchings, & J. Kinzie (Eds.), *Using evidence of student learning to improve higher education* (pp. 27–50). San Francisco, CA: Jossey-Bass.

Jankowski, N. A., Ikenberry, S. O., Kinzie, J., Kuh, G. D., Shenoy, G. F., & Baker, G. R. (2012). *Transparency & accountability: An evaluation of the VSA college portrait pilot.* Urbana: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment. Retrieved from http://www.learningoutcomesassessment.org/documents/VSA_001.pdf

Keller, C. M. (2014). Lessons from the Voluntary System of Accountability (VSA): The intersection of collective action & public policy. *Change*, *46*(5), 23–29.

Klein, S., Liu, O. L., Sconing, J., Bolus, R., Bridgeman, B., Kugelmass, H., & Steedle, J. (2009). *Test Validity Study (TVS) report*. Princeton, NJ: Educational Testing Service.

Kuh, G., & Ikenberry, S. (2009, October). *More than you think, less than we need: Learning outcomes assessment in American higher education*. Retrieved from http://www.learningoutcomesassessment.org/documents/niloafullreportfinal2.pdf

Kuh, G. D., Ikenberry, S. O., Jankowski, N., Cain, T. R., Hutchings, P., & Kinzie, J. (2015). *Using evidence of student learning to improve higher education*. New York, NY: Wiley.

Liu, O. L., Frankel, L., & Roohr, K. C. (2014). *Assessing critical thinking in higher education: Current state and directions for next-generation assessment* (Research Report No. RR-14-10). Princeton, NJ: Educational Testing Service. 10.1002/ets2.12009

Maki, P. L. (2002). Developing an assessment plan to learn about student learning. *The Journal of Academic Librarianship*, *28*(1), 8–13.

Markle, R., Brenneman, M., Jackson, T., Burrus, J., & Robbins, S. (2013). *Synthesizing frameworks of higher education student learning outcomes* (Research Report No. RR-13-22). Princeton, NJ: Educational Testing Service. 10.1002/j.2333-8504.2013.tb02329.x

McPherson, P., & Shulenburger, D. (2006). *Toward a voluntary system of accountability (VSA) for public colleges and universities.* Washington, DC: National Association of State Universities and Land-Grant Colleges.

Metcalf, J., & Kornell, N. (2007). Principles of cognitive science in education: The effects of generation, errors, and feedback. *Psychonomic Bulletin & Review, 14*(2), 225–229. Retrieved from http://link.springer.com/content/pdf/10.3758%2FBF03194056.pdf

O'Donovan, B., Price, M., & Rust, C. (2001). The student experience of criterion-referenced assessment (through the introduction of a common criteria assessment grid). *Innovations in Education and Teaching International*, *38*(1), 74–85.

Oswald, F. L., Schmitt, N., Kim, B. H., Ramsay, L. J., & Gillespie, M. A. (2004). Developing a biodata measure and situational judgment inventory as predictors of college student performance. *Journal of Applied Psychology*, *89*, 187–207.

Palomba, C. A., & Banta, T. W. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education.* San Francisco, CA: Jossey-Bass.

Provezis, S. J. (2010). *Regional accreditation and student learning outcomes: Mapping the territory* (NILOA Occasional Paper No. 6). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Rhodes, T. L. (2008). VALUE: Valid assessment of learning in undergraduate education. In V. Borden and G. Pike (Ed.), *Assessing and accounting for student learning: Beyond the Spellings Commission* (pp. 59–70). San Francisco, CA: Jossey-Bass.

Rhodes, T. L., & Finley, A. P. (2013). *Using the VALUE rubrics for improvement of learning and authentic assessment*. Washington, DC: Association of American Colleges and Universities.

Rust, C., Price, M., & O'Donovan, B. (2003). Improving students' learning by developing their understanding of assessment criteria and processes. *Assessment & Evaluation in Higher Education*, *28*(2), 147–164.

Shavelson, R. (2010). *Measuring college learning responsibly: Accountability in a new era*. Stanford, CA: Stanford University Press.

Shear, M. D. (2015, September 12). With website to research colleges, Obama abandons ranking system. *The New York Times.* Retrieved from http://www.nytimes.com/2015/09/13/us/with-website-to-research-colleges-obama-abandons-ranking-system.html?_r=0

Southern Association of Colleges and Schools' Commission on Colleges. (2011). *The principles of accreditation: Foundations for quality enhancement*. Retrieved from http://www.sacscoc.org/pdf/2012PrinciplesOfAcreditation.pdf

Suskie, L. (2009). *Assessing student learning: A common sense guide*. New York, NY: Wiley.

Tinto, V. (2012). *Completing college: Rethinking institutional action*. Chicago, IL: University of Chicago Press.

Tremblay, K., Lalancette, D., & Roseveare, D. (2012). *Assessment of higher education learning outcomes: Feasibility study report volume 1 – Design and implementation.* Paris, France: Organisation for Economic Co-operation and Development.

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

U.S. Department of Education. (2015). *For public feedback: A college ratings framework*. Retrieved from http://www.ed.gov/collegeratings

Western Association of Schools and Colleges. (2014). *Using evidence in the WASC accreditation process: A guide* (2nd ed.) Alameda, CA: Author.

Wilson, M., & Scalise, K. (2006). Assessment to improve learning in higher education: The BEAR Assessment System. *Higher Education*, *52*(4), 635–663.

Zapata-Rivera, D., & Katz, I. (2014). Keeping your audience in mind: Applying audience analysis to the design of score reports. *Assessment in Education: Principles, Policy & Practice, 21*(4), 442–463.

Zenisky, A. L., & Hambleton, R. K. (2012). Developing test score reports that work: The process and best practices for effective communication. *Educational Measurement: Issues and Practice, 31*(2), 21–26.

Suggested citation:

Russell, J., & Markle, R. (2017). *Continuing a culture of evidence: Assessment for improvement* (Research Report No. RR-17-08). Princeton, NJ: Educational Testing Service. http://dx.doi.org/10.1002/ets2.12136

Action Editor: Donald Powers

Reviewers: Richard Tannenbaum and Cathy Wendler