National Institute for Learning Outcomes Assessment

November 2018

Assessment 2.0: An Organic Supplement to Standard Assessment Procedure

Eric T. Metzler and Lisa Kurz

Occasional Paper #36

www.learningoutcomesassessment.org

Contents

NILOA Mission

The National Institute for Learning Outcomes Assessment's (NILOA) primary objective is to discover and disseminate ways that academic programs and institutions can productively use assessment data internally to inform and strengthen undergraduate education, and externally to communicate with policy makers, families and other stakeholders.

Please cite as:

Metzler, E. T., & Kurz, L. (2018, November). *Assessment 2.0: An organic supplement to standard assessment procedure*. (Occasional Paper 36). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

About the Authors

Dr. Eric T. Metzler is the Instructional Support and Assessment Specialist at the Kelley School of Business, Bloomington and Indianapolis. He oversees learning outcomes assessment for each of the Kelley School's degree-granting programs and assists business faculty in achieving and maintaining a national reputation for teaching excellence. He regularly teaches an interdisciplinary seminar, "Perspectives on American Consumerism," to college seniors and has also taught graduate courses on outcomes assessment at the department or program level. In his role as instructional consultant, Eric works confidentially with business faculty who wish to improve, refine, or otherwise perfect their teaching performance. As assessment specialist, Eric helps instructors, departments, programs, centers and institutes to articulate specific and measurable learning goals and outcomes and to identify innovative and workable ways to measure them. He also works with the programs' decision-making bodies, guiding the interpretation of assessment data, always with the ultimate goal in mind of using the data to improve student learning or the student experience. Eric has consulted and delivered workshops about active learning and other post-secondary pedagogical strategies, curriculum mapping, and assessment of student learning outcomes both in the U.S. and abroad in Iraq, South Africa and Jamaica. He specializes in making the work of assessment meaningful and useful to all involved in it.

Dr. Lisa Kurz is a Principal Instructional Consultant at the Center for Innovative Teaching and Learning (CITL) at Indiana University Bloomington. She provides instructional support for faculty across the University as they work to incorporate best practices in their teaching, implement innovative teaching strategies, and document their teaching practices for promotion or tenure. She provides workshops in a variety of topics, including a multi-day Course Development Institute that has led hundreds of IUB faculty through a backward design process to create or revise a course to ensure that the course's goals and learning outcomes are aligned with assessments and learning activities. As the assessment specialist in CITL, Lisa also works with departments and programs from a wide range of schools and colleges within the University, including the School of Public and Environmental Affairs, the School of Public Health, and the Maurer School of Law, to help them articulate measurable goals and student learning outcomes for their students' learning. She has also led a series of workshops for 50 departments in the University's College of Arts and Sciences to help them write goals and learning outcomes, map their curricula, and develop assessment plans for their undergraduate majors, minors, and certificate programs. Before joining the CITL, Lisa conducted institutional research at IUB, and served as the director of the university's writing center, Writing Tutorial Services.

Abstract

The discipline of assessment has matured to the point where there is general agreement on best practices. However, the field has made little progress in developing a theoretical basis — whether for assessment in general or for more specific dimensions of assessment as they emerge. Without a generalizable theory to work from, assessment professionals remain focused on the details of practice — getting it done — instead of turning their attention to systems thinking in the service of improving, revising, growing, or otherwise developing a field that is still far from perfect. In this article, the authors bring sociological theory to bear on learning outcomes assessment in order to understand its strengths and challenges from a systems point of view. Then, using this theoretical understanding, we propose an alternative method of assessment (Assessment 2.0) designed to supplement the assessment work already being done while at the same time avoiding its most difficult challenges. Assessment 2.0 is organic because it grows naturally from the professional judgment and experience of instructors rather than from the highly structured, linear procedure commonly followed in standard assessment practice.

Assessment 2.0: An Organic Supplement to Standard Assessment Procedure

Eric T. Metzler and Lisa Kurz

Using sociological theory to understand assessment's strengths and challenges from a systems point of view, we propose an alternative mode of assessment (Assessment 2.0) designed to supplement the assessment work already being done while at the same time avoiding its most difficult challenges.

While the idea of establishing student learning outcomes and assessing students' progress toward achieving them has been a part of U.S. educational practice at least since the time of John Dewey (Ewell, 2007), the assessment of learning in higher education as a movement is generally seen as beginning in the 1980s (Ewell, 2002, 2009; Weiss, Cosbey, Habel, Hanson, & Larsen, 2002), and since then it has grown in significance. The discipline of assessment—if it may so be called—has matured to the point where there is general agreement on best practices (e.g., Maki, 2010; Suskie, 2009; Walvoord, 2004). However, the field has made little progress in developing a theoretical basis—whether for assessment in general or for more specific dimensions of assessment as they emerge. Without a generalizable theory to work from, assessment professionals remain focused on the details of practice—getting it done—instead of turning their attention to systems thinking in the service of improving, revising, growing, or otherwise developing a field that is still far from perfect. In this article, we bring sociological theory to bear on learning outcomes assessment in order to understand its strengths and challenges from a systems point of view. Then, using this theoretical understanding, we propose an alternative method of assessment (Assessment 2.0) designed to *supplement* the assessment work already being done while at the same time avoiding its most difficult challenges. As we shall discuss below, our proposal for Assessment 2.0 is *organic* because it grows naturally from the professional judgment and experience of instructors rather than from the highly structured, linear procedure commonly followed in standard assessment practice.

Standard assessment practice follows a common process: articulating goals and learning outcomes, aligning the curriculum, identifying measures and instruments, collecting and analyzing data, and using the results to improve student learning (Pratto, 1996; Roscoe, 2017; Hatfield, 2009; Ewell, 2002; Martell & Calderon, 2005; Weiss et al., 2002). While this process has many strengths, including its logical, linear design, its visual depiction of whole curricula (tabular curriculum maps), and its focus on measurement of specific student learning, it can fall short when one considers the ultimate purpose of learning outcomes assessment. As assessment experts have pointed out, most institutions begin and execute the process up through the data collection step; fewer have been successful in *using* the data to make programmatic (as opposed to course-level) decisions that substantively improve student learning or the student experience. In other words, many are performing assessment to gather data; few are using the data to a useful end (Banta & Blaich, 2010; Jankowski, Timmer, Kinzie, & Kuh, 2018; Kuh & Ikenberry, 2009; Penn, 2011).

Another challenge confronting the standard process is the well-known faculty resistance and critique that perennially accompany assessment initiatives (e.g., Penn, 2011; Worthen, 2018; Eisgruber, 2012; Linkon, 2005). At our institution, research data from numerous faculty focus groups and individual faculty interviews have revealed broad frustration and dissatisfaction with the process (Katz, 2010; Weimer, 2013). For many faculty who resist assessment, the tasks are seen as a mandatory but perfunctory exercise not connected in any meaningful way to teaching and learning. We may regard such resistance as a reasonable response to a process that fails to recognize much of what faculty value about teaching and learning (Reed, 2016; Pollock, 2016). Moreover, from a faculty member's perspective, articulating student learning outcomes can reduce the richness, complexity, and beauty of their discipline to a short list of discrete knowledge, skills, and attitudes for students to master (Ewell [2007] refers to this as "fractionation"; see also McKernan, 2008; Cain & Hutchings, 2015; Powell, 2011; Hussey & Smith, 2002).

For many faculty who resist assessment, the tasks are seen as a mandatory but perfunctory exercise not connected in any meaningful way to teaching and learning.

From a systems perspective, the strengths and challenges of assessment form a correlative relationship: the strengths of standard practice bring with them a dark side that yields challenges. To be more specific, in order to conduct learning outcomes assessment systematically across schools, programs, and disciplines at university, it is necessary to devise a standard process that can be employed by faculty assessors who may have little or even no experience with assessment. A standard process enables instructors to complete assessments efficiently, with the goal of measuring student learning against program goals and student learning outcomes so that the resulting data may be used to improve a program's curriculum or the student experience. Without a standardized process, systematic assessment becomes very difficult to document or use on a large scale. And yet, the very system that enables schools to perform assessment at a high level brings with it inflexible processes that can result in compromised data, frustrated or alienated faculty, and ultimately the subversion of the very purpose of learning outcomes assessment. To better understand the complex relationship between the strengths of systematization and the challenges endemic to it, let us first turn to the systems theory of McDonaldization as developed by sociologist George Ritzer.

Assessment as Rationalized System

In 1993, Ritzer published the first of several editions of *The McDonaldization of Society*, which uses the business model of the fast food restaurant McDonald's to explore theories of rationalization as they occur in the modern world. Ritzer's project is to lay out his theory of *McDonaldization*—Ritzer's neologism describing the rationalization of all manner of structures in the late twentieth century—by explaining how it functions and giving a broad array of examples to illustrate

the theory as it plays out in modern life. The scope of the examples shows the great extent to which our world has adopted rationalized structures to do its business, and the academy is no exception. Ritzer adduces examples such as learning management systems (LMSs) and publisher-provided multiple choice assessments; student ratings of instruction and citation counting to determine a publication's influence; and Massive Open Online Courses (MOOCs) with their automated systems, standardized rubrics, and lectures as examples of rationalization in higher education (Ritzer, 2015). To this ever-growing list we may add the standard process of learning outcomes assessment, which both benefits and suffers from its highly rationalized structure.

Ritzer builds his theory around four interrelated dimensions—efficiency, calculability, predictability, and control—which together characterize rationalization as a system. Additionally, Ritzer delineates a fifth dimension—the irrationality of rationality—to explain the dark side of rationalization, when heavily rationalized processes subvert their own purpose. In other words, the strengths of rationalization exist in correlation with its weaknesses: the more a structure is rationalized, the greater the risk that its rationalization will undercut its own purpose. To better understand Ritzer's theory, let us look at each of its dimensions and the strengths they bring to the process of assessment.

Rationalization as a system is comprised of four interrelated dimensions: efficiency, calculability, predictability, and control. The strengths of rationalization exist in correlation with its weaknesses: the more a structure is rationalized, the greater the risk that its rationalization will undercut its own purpose.

Efficiency entails streamlining and optimizing processes so that the means to the end is fast, convenient, and easy. To make the assessment process efficient, we standardize the process, making it as straightforward and convenient as possible to follow. To keep the process manageable, we limit the expression of learning to a handful of goals and student learning outcomes. To make sorting easy, we use three or four buckets into which we sort student performance. To avoid extra work, we often collect data by counting the number of items students complete correctly on multiple choice exams (which are themselves highly rationalized). We create or purchase systems—often guided by computer algorithms—to manage workflow, organize data storage, and make the work of assessment as easy as possible for faculty.

Calculability is the dimension of rationalization that privileges that which can be measured, quantified, or counted with the corollary that quantity takes precedence over quality. In assessment, we may well see calculability as its premier characteristic among the four dimensions. At most schools, the focal point of outcomes assessment remains the numerical data: we look at graduation rates, employment figures, graduate salaries; we study test scores; and we translate qualitative work into numerical scores with rubrics (which are in themselves rationalized devices). We use numbers to tell our story, for we know that decision makers in our world are swayed by the allure of numbers. They are efficient. They appear clear cut. They seem objective, rational, and scientific. They help us to reduce the messy, complicated enterprise of teaching

and learning into a clear and simple snapshot that purportedly describes the extent to which students are learning the goals and outcomes we hope for. Translating complicated thinking and operations into numerical scores makes our work easier and more streamlined (hence tapping into efficiency) when making decisions about curricula and programming. Perhaps it is the only way to proceed without devolving into the chaos of multiple contingencies and complications.

The rationalized dimension of *predictability* seeks to minimize variability across time and space, emphasizing order, systematization, formalization, routine, and consistency. Given these definitions, the process of assessment is nothing if not predictable, in particular when computerized systems come into play. The digitization of assessment allows for efficient management of work flow and provides a digital receptacle not only for data, but also for artifacts of student work and other related documents. In order to fit into the system, documents must be standardized and thereby made predictable: lists of goals and learning outcomes are similar in length, format, and articulation. Student performance is sorted into a standard number of buckets, usually three or four. Forms and templates standardize the documentation. Rubrics articulate criteria so judgment about learning will be consistent and predictable. At our university, assessment is so routinized that groups of courses are assessed in *every* section taught, *every* time it is taught. Predictability enables assessment initiatives to function in large settings. Without it, no one would know how to perform assessment from course to course and from semester to semester; similarly, no one would know how to read and use the data, as it could be reported differently from assessment to assessment.

The fourth dimension of Ritzer's model of rationalization is *control* over the people involved as a means of avoiding the inefficiencies and unpredictability that inevitably accompany autonomous human action. Assessment processes depend on linear, sequential work flows that are typically organized and enforced by computer systems to ensure the desired process is followed in a timely manner. To ensure meaningful data collection, faculty are required to measure specific learning against pre-articulated goals and learning outcomes and to follow the steps of the process in a certain prescribed order. In academe, where harried faculty are burdened with heavy workloads, administrators must often mandate the completion of assessment reports, lest the number of assessments completed fail to reach a meaningful threshold. Indeed, without the dimension of control, exercised by assessment professionals, administrators, and accrediting bodies, we might reasonably expect the work of learning outcomes assessment to unravel and perhaps even grind to a halt.

When the four dimensions of rationalization work in concert, they enable large bureaucratic systems to function smoothly and serve great numbers of people

When the four dimensions of rationalization work in concert, they enable large bureaucratic systems to function smoothly and serve great numbers of people in a fair and straightforward manner.

in a fair and straightforward manner. Examples might include the Internal Revenue Service, school systems, hospitals, and insurance companies. Without rationalized processes, these and other large systems would quickly devolve into chaos. However, when the interrelated dimensions of efficiency, calculability, predictability, and control are over-emphasized or overused, when they are driven beyond their useful limits, when they take precedence over the end they are intended to serve, rationalization paradoxically begins to subvert the very goals and processes it was meant to streamline and facilitate in the first place. Ritzer identifies this paradox as the *irrationality of rationality*, the term he uses to label the fifth dimension of rationalization. And it is through this fifth dimension that Ritzer offers his main critique of rationalization. For example, drive-thru lanes at fast food restaurant chains are designed as rationalized systems aimed at streamlining the acquisition of a meal; yet when used by too many customers at once, drive-thru lanes become frustratingly inefficient, fraught with long waits, mistaken orders, and short tempers.

When pushed too far, efficiency in assessment practice rubs out the richness of a college education and promotes the collection of data without meaning. Indeed, gathering meaningful data about something as complex as the teaching and learning dyad is often expressly inefficient, time-consuming, and labor intensive.

We may similarly consider the irrationality of rationality as it applies to the rationalized systems put into place to complete the work of assessment. We establish *efficient* processes to save time, avoid extra work, and complete our assessment obligations as expediently as possible. Yet when pushed too far, *efficiency* in assessment practice rubs out the richness of a college education and promotes the collection of data without meaning. To ensure program assessment is of manageable scope, we limit the number of learning goals and outcomes articulated for the program. Yet this limitation for the sake of efficiency also artificially reduces the wholeness of a student's curriculum to a few standardized statements of what she or he should know or be able to do at graduation. Similarly, to streamline the collection of data, we often choose measurement procedures that are easy and undemanding of our time. Yet, not infrequently, data gleaned from easy measurement tell us little about the student learning that matters most (Muller, 2018; Berg & Seeber, 2016; Champagne, 2011; Powell, 2011). Indeed, gathering meaningful data about something as complex as the teaching and learning dyad is often expressly *inefficient*, *time-consuming*, and *labor intensive*.

When we consider *calculability* through the lens of the fifth dimension, similar problems begin to emerge. Whether collecting indirect assessment data such as graduation rates, employment statistics, graduate salary figures, and grades or direct assessment data such as performance on exams, homework, presentations or projects, *numbers* prevail. In their quantifiable simplicity, they can help us identify trends and patterns quickly and efficiently, but they can also rub out the nuance, variation, details, and context of student learning such that the numbers we see actually tell us little about what students have learned well and where gaps in their learning persist. The allure of numerical summary can lead us to spend valuable time and resources counting and measuring *something*

so that we have data to show; yet as David Eubanks points out in his article lamenting the notorious variance in validity and reliability of assessment data, "[I]t is difficult to use assessment results [when] the methods of gathering and analyzing data are very poor" (Eubanks, 2017, p. 4). Calculability pushed too far can lead us to measure what doesn't really matter at the expense of measuring what does so that we have numerical results (Muller, 2018; Berg & Seeber, 2016; Powell, 2011). In assessment, the focus on collecting quantifiable data at the expense of meaning often yields spreadsheets of data and plentiful reports that are nevertheless seldom used to improve student learning.

Predictability becomes irrational when standardization—employed to ease the work of faculty, organize data logically, and make assessment reporting uniform and understandable to all—eliminates the role of professional experience and judgment in contextualizing and giving meaning to assessment data. When the reporting system is automated, i.e., controlled by a computer algorithm and therefore standardized, faculty are required to complete assessments in cookie-cutter fashion, which can lead to mindless, perfunctory tabulations in order to fill in the blanks with data of some sort (Ewell, 2002). The rationalized predictability of highly standardized procedures and documentation leads to a mentality of compliance and disengagement among faculty: enter the number, check the box, and let others decide what it all means. While "good assessment begins with real, genuine questions that educators have about their students" (P. Hutchings as quoted in Worthen, 2018, para. 12), overly predictable assessment systems can have the effect of squelching the "real, genuine questions" arising from the professional experience and judgement of educators in favor of satisfying system requirements to provide data, which rarely respond to the most salient questions about student learning.

Of the four dimensions of rationality, *control* seems to morph into the realm of the irrational the most readily in the assessment world. In rationalized systems, control over human autonomy helps to increase efficiency, ensure predictability, and foster calculability by limiting the opportunity for human error, misjudgment, and distraction. But when driven too far, control backfires, and this is especially the case in the academic context, where the primary actors are accustomed to a great deal of autonomy. Most faculty members choose a career in academe because they love their subject area, teaching, and the freedom to structure their time as they please. For most faculty, mandates from administrators are not well received, and for some, they are an affront, purportedly infringing on their prized academic freedom (Cain, 2014). When administrators *require* faculty to do the work of assessment, the dimension of *control* manifests as resentment and frustration, which often leads to resistance or counterproductive attitudes, which can then even result in sabotage. Some faculty flatly refuse. Some go through the motions, collecting and reporting unimportant, sketchy, or even falsified data. Some faculty remonstrate loudly against the administration, creating a

Control, when pushed too far, can yield institutional cultures that collect data that are not useful and complete work done with a compliance mentality rather than a spirit of asking Hutchings's "genuine questions about students" in hopes of improving student learning or the student experience.

culture of discontent and frustration. Just as with the other dimensions in the assessment milieu, *control*, when pushed too far can yield institutional cultures that collect data that are not useful and complete work with a compliance mentality rather than a spirit of asking Hutchings's "genuine questions about students" in hopes of improving student learning or the student experience (Worthen, 2018; Pollock, 2016; Champagne, 2011; Porter, 2012).

The Faculty Voice

A careful consideration of the theory of rationalization led us to the conclusion that in the abstract, the dimensions of rationalization might be a useful framework for understanding assessment in higher education. As we have outlined, the drive for efficiency, predictability, calculability, and control, and the accompanying consequences suggested by the irrationality of rationality, seemed capable of accounting for much of what we as assessment practitioners had experienced in our interactions with faculty. Yet we needed to do a more careful analysis of actual faculty responses to assessment to determine whether faculty resistance could, in fact, be aligned with the dimensions of rationalization. To do so, we elicited from faculty at our university their responses to the assessment efforts in which they had participated, specifically the curriculum mapping workshops offered to help departments complete program review. Although this part of our project was not designed to inform yet another article about faculty resistance (Gilbert, 2018; Cain & Hutchings, 2015; Pratto, 1996; Ewell, 2002; Linkon, 2005; Penn, 2011), we began our work by identifying several faculty members from the workshops who had expressed reservations or even frustration about assessment to participate in semi-structured focus groups and one-on-one interviews. Hence, we acknowledge that our recruitment methods may have skewed our findings toward more negative responses to assessment, a research decision we made in the service of better understanding the relationship between faculty resistance and rationalization.

In both the focus groups and the individual interviews, we asked participants to tell us about their experience with the process of assessment, their attitude toward the work, whether they had concerns or reservations about assessment and pointedly, whether they thought the process of assessment could be reshaped to address their reservations and yield actionable improvements to student learning. The conversations were recorded (with the participants' permission, and after review of the project by our university's Institutional Review Board), and portions of the conversations were transcribed. We organized the faculty comments into categories and looked for evidence of alignment with Ritzer's (2015) dimensions of efficiency, calculability, predictability, control, and the irrationality of rationality.

We did a more careful analysis of actual faculty responses to assessment (e.g., their assessment experiences, concerns/reservations, and thoughts on whether assessment yields actionable improvements to student learning) to determine whether faculty resistance could, in fact, be aligned with the dimensions of rationalization.

While our questions addressed assessment of learning outcomes in general, faculty in their responses generally focused on the most rationalized aspect of assessment at our university, the large-scale assessment of General Education (Gen Ed) courses. Our Gen Ed assessment system gathers data each semester from over 1,000 courses offered in several divisions and schools across the university.

The faculty comments made in our interviews and focus groups illustrate well the irrationality of rationality in each of Ritzer's four primary dimensions. In our analysis, we divided faculty comments into three categories: those that addressed the *efficiency* and *predictability* of the assessment system (grouped together because comments touched on both of these dimensions simultaneously); those that touched on *calculability*; and those that addressed *control*.

Comments that related to Ritzer's dimensions of *efficiency* and *predictability* emphasized the standardization of the assessment process, particularly in the Gen Ed context. Faculty recognized the need to standardize the process because of the large quantity of data to be collected (all Gen Ed courses at our university are assessed in every semester they are offered, resulting in thousands of bits of data). But the standardized process, and the frequency of the demand to provide assessment data, led faculty to become frustrated with, or disengaged from, the process and their own data. One faculty member described her disengagement by saying, "It's easy to translate the process into something mechanical." When the process became mechanical, faculty seemed to want simply to get through it rather than engage with their data. Another professor described it this way:

> *Whether you take this seriously and try to do an authentic evaluation is voluntary. I know faculty members who've told me that what they do is they put up their essay prompt, they put up some kind of a rubric, and then they put up their grades and they use their grades to determine whether students have met their criteria or not. That's bogus.*

In addition to disengagement, the standardized process led some faculty to express frustration with the system's inflexibility: "The process of aligning grading rubrics with outcomes was problematic. It systematized things in a way that left little room for changes." Another professor said, "I got frustrated with the rubrics, so I reduced the number of assignments I require so that I wouldn't have to align so many rubrics with Gen Ed outcomes." The faculty disengagement with assessment data and results is a problem not only in itself. It also seems to lead to skepticism about the whole assessment enterprise, producing a sort of self-fulfilling spiral: faculty enter data mechanically, without reflection; that leads them to doubt not only their own data but also those of other faculty, which in turn causes them to see the entire project as useless. If the process is seen as useless, the resulting compliance mentality appears to be a reasonable response.

Faculty disengagement with assessment leads to skepticism, producing a self-fulfilling spiral: faculty enter data mechanically, without reflection; that leads them to doubt not only their own data but also those of other faculty, which in turn causes them to see the entire project as useless.

Faculty comments about the standardization associated with assessment were also connected with the idea of lowering the costs of offering courses, an aspect of efficiency. One faculty member commented that assessment "creates a level of standardization in order to find lower-cost ways to deliver courses." In this instructor's view, the drive for efficiency and predictability would eliminate "the specialness of a particular course as it's taught by a particular instructor."

Other comments related to efficiency and predictability emphasized that the assessment process failed to meet their needs. One professor noted that the efficiency of a standard process actually had the paradoxical effect of wasting time: "Some of this seemed like a waste of time because some things didn't apply to us. I wish we could have had a process that was more tailored to us." This is a clear example of the irrationality of rationality, where an efficient process subverts its own goal, leading to inefficiency. In faculty comments addressing the efficiency and predictability of the assessment system, we see faculty interpreting a system that was designed to *improve* teaching and learning instead as a system that *degrades* it.

On Ritzer's theme of *calculability*, faculty comments again illustrated the negative side of this dimension: that numbers fail to capture the nuance and complexity of teaching and learning, and that there is a propensity to count what is measurable rather than what is meaningful. Numerous faculty expressed these ideas in terms like these:

> *Students have such different backgrounds and such different reasons for taking a course, that assessment is uninformative. I'd have to write a biography of each student to tell what actually happened. So I can't take the assessment numbers seriously.*

Others expressed similar frustration with the gap between what they saw assessment as measuring, and what they saw as meaningful questions:

> *Assessment doesn't capture what students can actually do, which includes affectively laden stuff, unspoken stuff, nonverbal stuff … Assessment doesn't deal well with students who are continually struggling and taking steps forward. Does this student show curiosity, engagement, risk-taking—that's the kind of question I'm interested in. It's a distinction between measuring up to benchmarks and intellectual growth.*

Still others noted a mismatch between their perception of their students' success or failure in an assignment or course and the data they provided. One professor noted, "Having taught some of the classes that were evaluated, or worked closely with faculty in others, I knew that the results often did not mesh with what I

Some faculty noted a mismatch between their perception of their students' success or failure in an assignment or course and the data they provided. One professor noted, "Having taught some of the classes that were evaluated, or worked closely with faculty in others, I knew that the results often did not mesh with what I viewed as problems in the course."

viewed as problems in the course." Even if accurate and valid data were gathered and entered into the system (which may be unlikely, considering Eubanks' [2017] analysis of the problems with assessment methods), faculty described issues with the subsequent merging and analysis of the numbers. Referring to our university's Gen Ed assessment data entry system, one professor said:

> *The bogus part is the last part, where you have to say how many students met your criteria, exceeded your criteria, or did not meet your criteria. And when I did that, a little bar chart popped up. I suspect that what they do is they take all those little bar charts and they amalgamate them. Everybody meets the criteria—what criteria? They're all different, in all of our different classes. So what we have is "how many apples do we have? I don't know, we have a lot of fruit." It's a statistics fail.*

This quote hints at a concern expressed by other faculty in our interviews: that assessment numbers are meaningless because they're combined over so many different courses, disciplines, and teaching contexts. This concern and others in this category led to two responses from faculty that perfectly illustrate the irrationality of rationality. One is adjusting teaching to achieve better numbers, a response referred to by Muller in *The Tyranny of Metrics* (2018, p. 24) as "gaming." That response is shown in this quote:

> *If 20% of the class is with me, and 80% is lost, then I'll delete that material next year, so that I can get reasonable numbers. I'll lower the requirements in order to meet the standard.*

The other response is increasing frustration and resentment of the system and cynicism about assessment in general, as seen in these two quotes.

> *The most frustrating part is the feeling that you're out there in the classroom teaching, wanting your students to learn and doing the best you can, meaning well. When someone asks you for your numbers it's frustrating, it's annoying—it feels like they don't trust you. But there's no way to improve on that.*

> *Assessment makes us come up with high numbers of students who exceed expectations, so we move the bar. So it's not just a request to assess; it's a request to show that students are meeting or exceeding expectations.*

It is clear from these responses that the dimension of calculability is interconnected with efficiency and predictability, and that in combination, they all illustrate the dark side of assessment efforts. In striving for efficiency and predictability, the process is standardized rather than being tailored to individual departments' or faculty members' needs. By eliminating the nuances of teaching and learning

The dimension of calculability is interconnected with efficiency and predictability, and that in combination, they illustrate the dark side of assessment efforts. In striving for efficiency and predictability, the process is standardized rather than tailored to individual departments' or faculty members' needs.

in different contexts, the process seems to encourage perfunctory compliance rather than thoughtful reflection on the data. In striving for calculability, the process first asks faculty to submit (into a standardized, mechanical online system) data that can seem overly simplistic or even unrepresentative of what they think has actually happened in their courses. Then these data are amalgamated using what may be suspect methods. As a result, the process and the results are viewed with frustration and cynicism.

Unlike the comments related to efficiency, predictability, and calculability, which focused on the assessment process, faculty comments related to Ritzer's (2015) dimension of *control* focus primarily on the relationship between the faculty and university administration. Faculty concerns touching on this dimension fell into two categories: resistance to assessment as a top-down mandate from administrators, and concerns with academic freedom. Regarding the first of these, one faculty member we interviewed noted that she hates the phrase, "a culture of assessment." When asked why, she said, "Again, it has this sort of top-down, authoritarian feel to it." In discussing the curriculum mapping workshops, another faculty member noted that "We *should* be mindful of what our courses address. But the top-down mandate felt like the tail wagging the dog — it was completely compliance-driven." The issue of academic freedom was raised in terms of concerns over the content of courses and curricula. Almost all faculty we interviewed expressed an attitude like this one: "The system isn't useful, and faculty feel like they're losing control over their curriculum." This idea was often tied to concerns about standardization of courses (an aspect of the dimension of efficiency):

> *There has been pressure from [the state commission for higher education] and the legislature, who want to standardize courses in order to drive down costs. But this leads to the faculty losing oversight over the course, and losing the specialness of a particular course as it's taught by a particular instructor.*

This quote articulates not only the perceived loss of control over the curriculum that results from assessment, but also the concern that assessment will rub out the individuality and unique nuances of a particular professor's teaching of a particular course. This is a theme connected to our analysis of efficiency, again showing the overlap in these dimensions. When assessment was articulated as an issue of control by administrators over faculty, the affective responses were strongly negative, as predicted by our earlier theoretical analysis.

> *We have a conception of what our jobs are, and now we're being told that our conception is totally wrong. We're being told we need to rethink our notions of what teaching is, and that's upsetting.*

When assessment was articulated as an issue of control by administrators over faculty, the affective responses from faculty were strongly negative.

I had to collect syllabi from all of our faculty. All the syllabi were carefullycrafted, and all the faculty were doing the best they could. And then they get asked by someone who knows nothing about the field to provide more information. People who don't know what they're talking about are asking us for more information. It's frustrating as well as insulting.

A more cynical view of the issue of control was expressed by another faculty member: "It's just a matter of figuring out what the gods of assessment want." Underlying concerns related to Ritzer's dimension of control is a hint of a lack of trust: faculty question the rationale for doing assessment and wonder whether administrators have hidden agendas. Several faculty expressed views like this: "What is the evidence that assessment makes anything better? How would we assess assessment?" In the absence of clear answers to these questions, some faculty suggested that assessment may be a politically-driven move by administrators to assess *faculty*, or to make decisions about departments and programs (despite repeated assurances by upper level administrators that this is not the case).

In our interviews and focus groups, we asked faculty not only about their experiences with assessment and attitudes toward it; we also asked them to speculate about how assessment could be reshaped to address their concerns. In their responses, faculty had few concrete suggestions. Some suggested that a less top-down process might be less frustrating and more useful. The faculty member quoted above, who said that she hates the phrase, "a culture of assessment," continued this way:

We asked faculty to speculate about how assessment could be reshaped to address their concerns. In their responses, some faculty suggested that a less top-down process might be less frustrating and more useful.

But the idea that [the phrase "a culture of assessment"] expresses is an important one—one where everybody in their teaching is identifying what students are learning, what students are having trouble learning, and they're pooling their data not in that quantitative way but in a qualitative way. So the individuals are gathering quantitative [data], and it has to be quantitative at a certain level because that's a way of summarizing… And then you bring that data to conversations within the department. And then we have the opportunity to move forward as a collectivity. That's what programmatic assessment is supposed to be, but to a lot of people that still feels top-down.

This quote touches on several themes expressed by other faculty. One is that assessment would be better if it focused at a local level within individual departments, rather than university-wide: "It has to be indigenous. If it comes down from above, it's the kiss of death." Another faculty member agreed: "Make it so that it's us telling us what to do, not somebody else telling us what to do." Yet another said, "The best way that administrators could do this is to ask faculty to do something. It'll be unique to your department, but do something, and document it when you're done."

Another theme captured in the quote above is the notion that assessment could be done in a more holistic, qualitative way. Some said that faculty should be allowed "to provide context and additional information in assessment reports," and to "include both qualitative and quantitative information" in their assessment reports. They suggested that "faculty should have more conversations about these issues without the terminology getting in the way." One faculty member noted that "face-to-face conversations … would make it feel less compliance-driven. It would also get faculty to think beyond their own syllabus."

In sum, our analysis of faculty responses to assessment leads us to conclude that the theory of rationalization provides a useful framework for understanding both the process of assessment and faculty resistance to it. The dimensions of efficiency, predictability, calculability, control, and the irrationality of rationality (Ritzer, 2015) help identify both the strengths and the problems of assessment as it is currently practiced. Rationalization makes large-scale assessment possible, but paradoxically, it also eventually undermines its own purpose, as well as creating strong faculty resistance to the assessment enterprise.

The problems engendered by a noble attempt to make assessment possible on a mass level must be corrected if we are ever to make assessment relevant to the teaching life of our faculty.

The problems engendered by a noble attempt to make assessment possible on a mass level must be corrected if we are ever to make assessment relevant to the teaching life of our faculty. And since rationalized procedures are so deeply embedded in current assessment practice, we believe that the key to re-envisioning assessment lies in using an awareness of the dimensions of rationalization in order to avoid the most problematic expressions of rationalization. In other words, we will apply an *inverse of McDonaldization* to guide our thought. The goal is to develop a new way of doing assessment based on avoiding the problems highlighted by rationalization so that the enterprise can more effectively move beyond data collection to action. Such an approach could make assessment more acceptable to faculty and at the same time focus faculty energy on improving student learning and the student experience rather than on the details of data collection.

Foundational Principles of Assessment 2.0

As we consider alternative, viable assessment methods based on *avoiding* the problems of heavily rationalized systems, we begin by recognizing the contributions of other assessment scholars whose ideas have informed our own thinking. For example, we can see various themes of Assessment 2.0 in Ewell's (2009; see also Huba & Freed, 2000; Roscoe, 2017) description of the *"assessment for improvement" paradigm*, which emphasizes the formative purpose of assessment, and which Ewell contrasts with the *assessment for accountability paradigm*. His assessment for improvement paradigm stresses a

focus on assessment *methods and questions that are meaningful for the faculty* who will be conducting the assessment. As Ewell notes, "Such questions are frequently best framed in terms of particular pedagogical challenges that real faculty face in real classrooms" (2009, p. 16). Many other scholars have also emphasized this point (Hutchings, 2010; Maki, 2010; Roscoe, 2017; Linkon, 2005; Steinke and Fitch, 2011; Eubanks, 2017; Reed, 2016). Another theme of Assessment 2.0 emphasized by other scholars is the key role of *discussion, reflection, and collective meaning-making* among departmental faculty as they analyze assessment data (Ewell, 2009; Cain & Hutchings, 2015; Roscoe, 2017; Linkon, 2015; Huba and Freed, 2000). Other authors have also highlighted the benefits of drawing on both the *disciplinary expertise* and the *professional judgment* of faculty in designing and conducting assessments (McConnell, 2018; Cain & Hutchings, 2015; Cain, 2014; Hutchings, 2010; Maki, 2010; Roscoe, 2017; Eubanks, 2017). Finally, the use of *flexible assessment methods* that utilize *qualitative as well as quantitative data* has also been mentioned by numerous authors (Clark and Filinson, 2011; Maki, 2010, McConnell, 2018; Cain & Hutchings, 2015; Cain, 2014; Porter, 2012; Ewell, 2009; Linkon, 2005). It is clear that Assessment 2.0 is not a set of novel ideas or approaches to assessment. Instead, we have attempted to synthesize many principles and best practices from the literature and provide a theoretical framework that helps us understand and use them to design assessment methods that faculty can own and can use to improve student learning.

Assessment 2.0 is not a set of novel ideas or approaches to assessment. Instead, we have attempted to synthesize many principles and best practices from the literature and provide a theoretical framework that helps us understand and use them to design assessment methods that faculty can own and can use to improve student learning.

At the same time, we must be wary of subverting our own project by rationalizing a non-rationalized system into a one-size-fits-all solution. In the final section of this essay, we will discuss the preliminary results of non-rationalized assessment as it is playing out in our context; yet we also caution that our Assessment 2.0 procedure is not intended prescriptively, but rather as an exemplar of a successful process based on foundational principles developed from the theorization we have explored above. In other words, it is one example of many possibilities for re-envisioning successful assessment procedures that are faculty friendly and lead to action.

We offer the principles below as guidelines for consideration prior to the development of Assessment 2.0 variations, which should be tailored to the unique goals and needs of each department, program, school, or university.

Data Quality and Usefulness **over Efficiency:** While rejecting an overemphasis on efficiency does not imply expressly devising an *inefficient* assessment program, it does indicate a shift in priorities, especially in the data collection phase. For Assessment 2.0, we believe the process should focus foremost on collecting data that will inform us about student learning that *matters* and form the basis for decisions about curricular and instructional changes in the students' best interest. Hence, we believe the process should

be less concerned with completing the assessment as quickly and easily as possible, and more concerned with finding the best way to discover the gaps and fissures in student learning vis-à-vis the learning we care most about. Often, for example, instructors identify items on multiple choice exams that relate to a particular learning outcome and then complete the assessment by categorizing student achievement according to how many of the selected items each student answers correctly. This method is fast and easy (efficient) but it rarely reveals data detailed enough to inspire confidence among faculty and administrators to make curricular changes. A much less efficient method, on the other hand, might involve studying student work on one multi-step problem that reveals the thought processes and competencies indicated by a learning outcome. While this latter assessment procedure will surely take more time and effort (not efficient), it will yield data that can serve as persuasive evidence of student performance on which curricular changes and hopefully improvements can be made.

Local Applicability **over Predictability:** Negotiating the inverse of predictability can be tricky in an institutional context with many players involved in the process. On the one hand, dispensing with predictability altogether may translate into confusion, where faculty idiosyncratically complete assessments, and data are reported in every which way with no common basis or structure. Faculty committees would struggle to interpret the results or recognize program-wide patterns. Unpredictable (non-standardized) assessment procedures would also emphasize individualized, course-level improvements rather than improvement at the program, department, or school level because of the individualization of the assessment process. On the other hand, requiring faculty to complete assessments in lockstep fashion with no option for contextual influences, individual findings, or out-of-the-ordinary observations can lead to a compliance mentality: check the boxes, fill in the numbers and submit. Hence, rather than standardization, we propose a process that asks instructors to focus on the production of data that emerge naturally from their experience and professional judgment as educators with respect to the student learning in their purview. With more latitude for the instructor and less pre-determined uniformity, assessment becomes much more faculty friendly and further stands to produce data that may have gone unnoticed in a more rationalized process.

Holistic Assessment **over Reductive Calculability:** In our estimation, one of the great problems with current assessment practice is the overreliance on quantifying student learning, even when numerical data may be inappropriate or lead to inaccuracies or misunderstandings. Perhaps because of their simple elegance and ability to summarize information, numerical data have an allure that is hard to resist. In standard assessment

One of the great problems with current assessment practice is the overreliance on quantifying student learning, even when numerical data may be inappropriate or lead to inaccuracies or misunderstandings. Assessment 2.0 asks faculty instead to look at student performance holistically.

practice, rubrics have become *de rigueur* in the data collection process (see Jankowski et al., 2018). While they can be extremely useful in marking student work and transforming performance into a score, rubrics can also miss key characteristics in student work and focus our attention on aspects that don't matter, even as we fail to notice those that do. Assessment 2.0 asks faculty instead to look at student performance holistically, asking for observations about student learning based on the faculty member's experience and professional judgment about the student work they have reviewed, whether for formative or summative purposes. While we in no way wish to suggest that numerical data should be abandoned—they will always occupy an important place in the assessment world — we *do* wish to suggest that non-numeric, holistic data may be an important supplement that helps us to identify learning gaps and developmental problems that we might otherwise miss when performance is reduced to a simple numerical score.

When faculty have control and agency over the assessment process and how the data will be used, their natural interest in student learning is activated and assessment becomes an interesting and energizing activity instead of a requirement of compliance.

***Subsidiarity* over Control:** Top-down administrative control over faculty in the assessment process is perhaps the greatest threat to the entire enterprise because it typically yields resentment, frustration, resistance and sometime even sabotage. Part of the problem stems from administrative interference with the fabled independent and self-regulative work habits tacitly promised to tenured faculty. A second contributing factor, however, is the opacity that shrouds how assessment data will be used. When faculty are asked to collect data to satisfy administrative bodies with no notion of how those data will be used generally or locally to benefit them, resentment and disengagement ensue. Hence, we suggest that while top-level administrative *support* for the assessment enterprise is essential, assessment itself should be driven by the principle of *subsidiarity*, where matters are handled at the smallest, lowest, or least centralized authority. When groups of faculty who teach closely-related courses determine what to assess based on their natural interest and curiosity about student learning in their own local contexts, the work of assessment takes on direct relevance for the faculty doing the work. When faculty have control and agency over the assessment process and how the data will be used, their natural interest in student learning is activated and assessment becomes an interesting and energizing activity instead for a requirement of compliance.

***Actionability* over Data Gathering For Its Own Sake:** Closely related to subsidiarity, but deserving of its own designation is *actionability*. As noted above, all too often assessment work focuses on the procedural steps up to and including data collection with little attention to *using* the data to make changes and improvements to curricula, programs, instruction, and other dimensions of the student experience. While we may adduce several reasons for this lamentable state of affairs, a major contributing factor remains the

alienation faculty feel toward the data and their meaning. In order for the assessments to matter and be taken seriously, faculty must have agency over changing or improving student learning. They must see how the data relate to their own experience as instructors and know that they can act to close the gaps and mend the fissures they detect in their own students' learning. Accordingly, Assessment 2.0 focuses first on doing assessment that leads to *action*, not simply data collection (Walvoord, 2004). Indeed, so important is this principle that we believe assessment should not be done unless there is real possibility and openness to change. Correlatively, we believe that the most effective assessments begin with faculty curiosity about aspects of students learning they suspect are not up to par. When this context obtains, assessment can reveal hidden learning problems that instructors can then address with interest and enthusiasm.

Assessment 2.0 is organic because rather than being governed by a rationalized, external system, faculty are invited to indicate the learnings they hoped to see in their students and to use their professional experience and judgment to assess the extent to which students mastered those learnings.

Assessment 2.0: An Organic Exemplar

Beginning with the foundational principles of Assessment 2.0, we developed and piloted an assessment process that eschews the disadvantages of rationalization endemic to the traditional assessment process and yet nevertheless gathers rich, actionable information about student learning that matters to individual instructors and programs. We think of our process as *organic* because rather than being governed by a rationalized, external system, faculty are invited to indicate the learnings they hoped to see in their students and to use their professional experience and judgment to assess the extent to which students mastered those learnings. Instead of tables of quantitative data, individual faculty members offer qualitative data that are then reviewed—either by faculty themselves or by an assessment expert or consultant—to identify themes and commonalities as they naturally emerge from faculty remarks.

At this juncture, assessment experts may be reacting with skepticism: What about program goals and learning outcomes? What about curricular alignment? What about looking directly at student work or performance? Just as Ewell et al. (2017) express reservation in response to Roscoe's (2017) suggestion that improvements to student learning be based on faculty conversation and research literature rather than direct (and rationalized) assessment, one might reasonably levy the same criticism on Assessment 2.0. That is, "efforts to improve that do not start by defining the objectives of instruction are akin to sailing without a compass. … [E]fforts to improve without shared instructional goals and associated measurements will not lead to systematic improvement" (Ewell et al., 2017, paragraph 11).

In response to these questions and potential criticisms, we by no means wish to suggest that the traditional structures of assessment be abandoned or elided.

Each of the well-considered steps of traditional, rationalized assessment is indispensable to the overall project of improving curricula and student learning.

Thus, we do see Assessment 2.0 as a *supplement* to standard practice, providing a fuller, more detailed picture of student learning from which to take action and plan improvements. We still begin (see our questions below) by asking faculty to tell us their core learning objective(s), but we let them emerge naturally, rather than imposing them from without. Interestingly, the independently articulated learning objectives collected from faculty aligned surprisingly well, both among faculty and with the program's previously articulated goals and student learning outcomes. Perhaps we may say that Assessment 2.0 navigates with the stars rather than a compass; yet it sails with purpose nonetheless.

We do see Assessment 2.0 as a supplement to standard practice, providing a fuller, more detailed picture of student learning from which to take action and plan improvements.

For our pilot study, we focused on a single program and selected seven faculty members who had considerable experience teaching key courses (in different content areas and at different levels) in a program's curriculum. We sent them a questionnaire and asked them to spend no more than 15 minutes responding to five simple questions:

1. What is the most important learning you hoped to see in your course this semester?
2. Based on the student work you've seen, what have been the greatest gains in student learning this semester?
3. Based on the student work you've seen, what are the biggest gaps in student learning (i.e., what do you wish students had learned better)?
4. How do you think students might have been better prepared when they entered your course—what knowledge, skills, or attitudes did they need more of?
5. Based on your observations, what do students still need to learn or improve upon as they exit your class?

We reviewed and discussed the rich qualitative data provided by all seven instructors and grouped their responses into a small number of themes. The responses to question 3, about gaps the faculty members had observed in their students' learning, particularly surprised us because of their similarity, even across the very different content areas of the courses involved: nearly all the instructors noted that their students had limited ability to integrate and apply knowledge beyond defined and familiar contexts. Students seemed to memorize rather than learn concepts. Faculty frequently offered their responses in this form: "Students struggle with X, because they're doing Y," which gave a fuller view of the learning problem than might have been possible in a traditional assessment, because it described what students *were* doing as well as what they *weren't* doing.

We presented our findings to the faculty participants and to the head of the program, and they found the results significant enough to call a meeting of key program faculty to discuss the results and brainstorm ideas for improving student learning. After that meeting, a smaller committee developed a plan to address the gaps in students' learning through faculty development events scheduled for the following semester designed to help faculty teach skills of application and deep learning.

Our pilot study of organic assessment, while preliminary and small in scale, demonstrates the feasibility of conducting assessment in a non-rationalized way, by designing the assessment in alignment with the principles of Assessment 2.0. Regarding *data quality over efficiency*, our method allowed us to obtain rich qualitative data from which consistent themes emerged—themes that touched on ways of thinking at the heart of the discipline involved. Yet while avoiding an emphasis on efficiency, our methods were paradoxically efficient: the assessment methods in each case demanded little faculty time, but yielded significant information. To enhance *local applicability over predictability*, we devised an assessment method that left room for faculty to use their professional judgment in choosing what specific issues to focus on in their responses. Interestingly, while the faculty responded individually to the questions, their responses coalesced on a small number of important concerns and issues. The method favored *holistic assessment over reductionistic quantification* by asking for qualitative information drawn from faculty members' judgment and experience rather than numbers. The emphasis on *subsidiarity over control* was perhaps the key to this pilot. There was no mandate from higher administration to conduct this assessment; participation by the faculty was voluntary. Faculty controlled the information they provided and the conclusions they drew, and the assessment focused on courses they had direct control over. Faculty were not alienated from the information they provided. The themes identified from faculty responses were significant enough to motivate them to brainstorm possible solutions to the learning gaps identified, ensuring that the results were *actionable* rather than *data gathering for its own sake*.

Our pilot study demonstrates only one of many possible ways of instantiating the principles of Assessment 2.0. The principles themselves are not intended to be restrictive; we are not proposing simply another assessment method to be rigidly applied precisely as we have described it, in place of traditional methods. Rather, we see the principles as generative, encouraging the creation of a wide variety of assessment methods to meet the needs of faculty in diverse disciplines and teaching contexts. Here are a few possibilities.

The principles of Assessment 2.0 are not intended to be restrictive; we are not proposing simply another assessment method to be rigidly applied precisely as we have described it, in place of traditional methods.

- Faculty from several related departments (e.g., different foreign language departments) might identify a teaching challenge or learning gap shared across their disciplines, and brainstorm ways to gather evidence to further analyze the challenge, or to assess potential strategies to address it.

- Faculty who teach different sections of a Gen Ed course (or related Gen Ed courses in a single discipline or department) could decide on a specific concept or skill that is a key component of the course(s), and gather evidence to determine how well students learn the concept or skill across the courses.

- Faculty might combine traditional quantitative assessment with Assessment 2.0 principles by meeting at a local level (as a department, or as a group of faculty teaching closely related courses) to discuss their quantitative data and compare them with more holistic judgments of students' learning to search for commonalities.

These ideas harness the principles of Assessment 2.0 to help faculty find ways of making assessment—even large-scale Gen Ed assessment—meaningful at a local level. But there are many other possibilities. We call on our colleagues to reflect on the methods used in their own contexts in light of our theorization, and share other examples of assessment that are consistent with the principles of Assessment 2.0.

Assessment can be designed so that faculty identify aspects of student learning that are important to them; keep decisions about assessment methods at a local level; rely on their own professional judgment as well as other sources of evidence; and arrive at actionable conclusions to address the learning gaps they have identified. When we as assessment professionals facilitate this kind of organic assessment, we may not only avoid the hostility and resistance associated with traditional assessment methods; we may also foster assessment that leads to real improvements in student learning.

When we as assessment professionals facilitate this kind of organic assessment, we may not only avoid the hostility and resistance associated with traditional assessment methods; we may also foster assessment that leads to real improvements in student learning.

References

Banta, T. W., & Blaich, C. (2010). Closing the assessment loop. *Change: The Magazine of Higher Learning, 43*(1), 22-27.

Berg, M., & Seeber, B. K. (2016). *The slow professor: Challenging the culture of speed in the academy*. Toronto: U Toronto Press.

Cain, T. R. (2014). *Assessment and academic freedom: In concert, not conflict* (Occasional Paper No. 22). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

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*. San Francisco, CA: Jossey-Bass.

Champagne, J. (2011). Teaching in the corporate university: Assessment as a labor issue. *Journal of Academic Freedom, 2*. Retrieved from: https://www.aaup.org/JAF2/teaching-corporate-university-assessment-labor-issue#.WwbhbYoh1hF

Clark, R., & Filinson, R. (2011). Kicking and screaming: How one truculent sociology department made peace with mandatory assessment. *Teaching Sociology, 39*(2), 127-137.

Eisgruber, C. L. (2012). *Promoting a 'culture of engagement,' not a 'culture of assessment.'* Remarks to presidents of the American Association of Universities. Retrieved from www.princeton.edu/president/eisgruber/speeches-writings/archive/?id=10613.

Eubanks, D. (2017). A guide for the perplexed. *Intersection*, fall 2017.

Ewell, P. (2002). An emerging scholarship: A brief history of assessment. In Banta, T.W. (Ed.), *Building a scholarship of assessment*. San Francisco, CA: Jossey-Bass.

Ewell, P. (2007). *Applying learning outcomes concepts to higher education: An overview*. Prepared for the Hong Kong University Grants Committee. Retrieved from: https://www.southernct.edu/assessment-and-planning/Applying%20Learning%20Outcomes%20Concepts%20to%20Higher%20Education%20An%20Overview.pdf

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

Ewell, P., Hutchings, P., Kinzie, J., Kuh, G., & Lingenfelter, P. (2017). *Taking stock of the assessment movement. Liberal Education, winter, 2017.* NILOA Viewpoint. Retrieved from https://www.learningoutcomesassessment.org/wp-content/uploads/2019/08/Viewpoint-Ewelletal.pdf

Gilbert, E. (2018). *An insider's take on assessment: It may be worse than you thought*. Chronicle of Higher Education. Retrieved from https://www.chronicle.com/article/An-Insider-s-Take-on/242235.

Hatfield, S. (2009). *Assessing your program-level assessment plan* (IDEA Paper #45). Manhattan, KS: The IDEA Center.

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.

Hussey, T., & Smith, P. (2002). The trouble with learning outcomes. *Active Learning in Higher Education, 3*(3), 220-233.

Hutchings, P. (2010). *Opening doors to faculty involvement in assessment* (Occasional Paper #4). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Jankowski, N. A., Timmer, J. D., Kinzie, J., & Kuh, G. D. (2018, January). *Assessment that matters: Trending toward practices that document authentic student learning*. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

Katz, S.N. (2010). Beyond crude measurement and consumerism. *Academe, 96*(5). Retrieved from https://www.aaup.org/article/beyond-crude-measurement-and-consumerism.

Kuh, G. D., & Ikenberry, S. O. (2009). *More than you think, less than we need: Learning outcomes assessment in American higher education*. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Linkon, S. L. (2005). How can assessment work for us? *Academe 91*(4). Retrieved from https://www.aaup.org/publications/Academe/2005/05ja/05jalink.htm.

Maki, P. (2010). *Assessing for learning: building a sustainable commitment across the institution* (2nd ed.). Sterling, Va.: Stylus Publishing.

Martell, K., & Calderon, T. (2005). Assessment in business schools: What it is, where we are, and where we need to go now. *Assessment of Student Learning in Business Schools, 1*(1), 1-26.

McConnell, K. D. (2018). *Assessment isn't about bureaucracy but about teaching and learning* (opinion). Inside Higher Ed. Retrieved from https://www.insidehighered.com/views/2018/03/01/assessment-isnt-about-bureaucracy-about-teaching-and-learning-opinion.

McKernan, J. (2008). *Curriculum and imagination*. London: Routledge.

Muller, J. Z. (2018). *The tyranny of metrics*. Princeton: Princeton UP.

Penn, J. D. (2011). The case for assessing complex general education student learning outcomes. *New Directions for Institutional Research, 2011*(149), 5-14.

Pollock, D. (2016). *Learning outcomes: The debate continues*. SoundIDEA Podcast. Retrieved from www.ideaedu.org.

Porter, D. (2012). Assessment as a subversive activity. *Journal of Academic Freedom, 3*.

Powell, J. W. (2011). Outcomes assessment: Conceptual and other problems. *Journal of Academic Freedom, 2*. Retrieved from: https://www.aaup.org/JAF2/outcomes-assessment-conceptual-and-other-problems#.WwbjF4oh1hF.

Pratto, D. J. (1996). Assessment abuse by design. *Teaching Sociology, 24*(1), 119-122.

Reed, M. (2016). *The whole and the sum of its parts. BLOG: Confessions of a Community College Dean*. Inside Higher Ed. Retrieved from https://www.insidehighered.com/blogs/confessions-community-college-dean/whole-and-sum-its-parts.

Ritzer, G. (2015). *The McDonaldization of society* (8th ed.). Thousand Oaks, Calif.: Sage.

Roscoe, D. D. (2017). Toward an improvement paradigm for academic quality. *Liberal Education, 103*(1). Retrieved from www.aacu.org.

Steinke, P. & Fitch, P. (2011). Outcomes assessment from the perspective of psychological science: The TAIM approach. *New Directions for Institutional Research, 2011*(149), 15-26.

Suskie, L.A. (2009). *Assessing student learning: a common sense guide*. San Francisco: Jossey-Bass.

Walvoord, B. E. (2004). *Assessment clear and simple: A practical guide for institutions, departments, and general education*. San Francisco, CA: Jossey-Bass.

Weimer, M. (2013). The assessment movement: Revisiting faculty resistance. *Faculty Focus*. Retrieved from https://www.facultyfocus.com/articles/educational-assessment/the-assessment-mov.

Weiss, G. L., Cosbey, J. R., Habel, S. K., Hanson, C. M., & Larsen, C. (2002). Improving the assessment of student learning: Advancing a research agenda in sociology. *Teaching Sociology, 30*(1), 63-79.

Worthen, M. (2018). *The misguided drive to measure 'learning outcomes.'* The New York Times Sunday Review. Retrieved from https://www.nytimes.com/2018/02/23/opinion/sunday/colleges-measure-learning-outcomes.html.

About NILOA

- The National Institute for Learning Outcomes Assessment (NILOA) was established in December 2008.
- NILOA is co-located at the University of Illinois and Indiana University.
- The NILOA website contains free assessment resources and can be found at http://www.learningoutcomesassessment.org.
- The NILOA research team has scanned institutional websites, surveyed chief academic officers, and commissioned a series of occasional papers.
- NILOA's Founding Director, George Kuh, founded the National Survey for Student Engagement (NSSE).
- The other co-principal investigator for NILOA, Stanley Ikenberry, was president of the University of Illinois from 1979 to 1995 and of the American Council of Education from 1996 to 2001.

NILOA Staff

Natasha Jankowski, *Director*

Gianina Baker, *Assistant Director*

Katie Schultz, *Project Manager*

Erick Montenegro, *Communications Coordinator and Research Analyst*

Verna F. Orr, *Research Analyst*

NILOA Senior Scholars

Peter Ewell, *Senior Scholar*

Pat Hutchings, *Senior Scholar*

Stanley Ikenberry, *Co-Principal Investigator*

Jillian Kinzie, *Senior Scholar*

George Kuh, *Founding Director, Senior Scholar, and Co-Principal Investigator*

Paul Lingenfelter, *Senior Scholar*

David Marshall, *Senior Scholar*

NILOA Sponsors

Lumina Foundation for Education

University of Illinois, College of Education

National Institute for Learning Outcomes Assessment

For more information, please contact:

National Institute for Learning Outcomes Assessment (NILOA)
University of Illinois at Urbana-Champaign
51 Gerty Drive
Suite 196, CRC, MC-672
Champaign, IL 61820

learningoutcomesassessment.org
niloa@education.illinois.edu
Phone: 217.244.2155