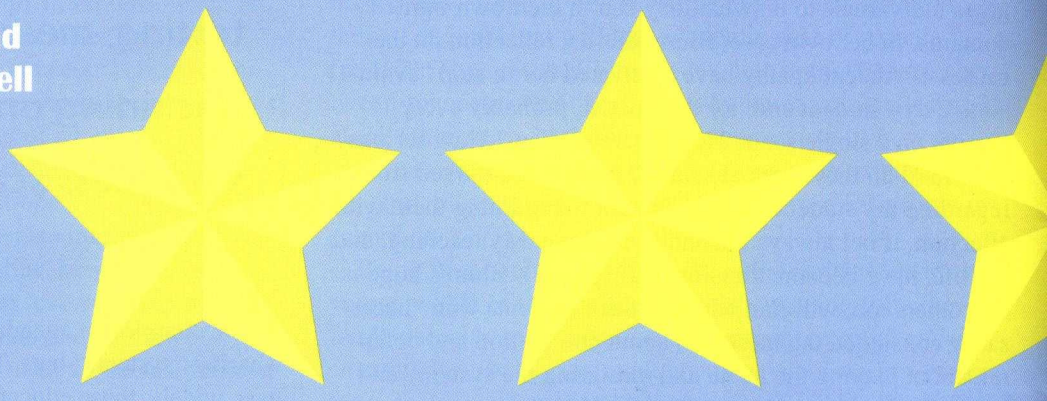
BIG PROGRESS IN Authentic Assessment,

BUT BY ITSELF NOT ENOUGH

By
Daniel F. Sullivan and
Kate Drezek McConnell



In Short

- Faculty practices must change if college student learning of both disciplinary content and higher-order skills is to meet 21st-century needs and expectations.
- Unlike results generated by many of the commercially available standardized testing options, faculty can engage with and use VALUE rubric scoring data immediately.
- The ability to deconstruct VALUE Rubric scoring results by specific dimensions provide faculty with a useful tool for asking more meaningful questions about where to allocate teaching time and where new assignments designed to improve competence should be developed.
- The connection between changes faculty make in their assignments and changes in students' learning outcomes is clear. When faculty members improve their teaching methods—or merely improve an assignment—improvement in student learning is detectable.
- The way our interpretation of academic freedom has evolved wrongly conflates the freedom faculty do and must have regarding the teaching of disciplinary content with the notion that they should be equally free—as individuals—to decide on pedagogical approaches even when strong evidence favors some kinds of pedagogy over others.

There is a growing agreement throughout American higher education and among higher education's stakeholders that:

- A high-quality, 21st-century undergraduate education—which the Association of American Colleges and Universities (AAC&U) and others refer to as a liberal education—involves both students' learning of disciplinary and/or pre-professional content and their development of higher-order thinking skills and capacities, such as AAC&U's LEAP Essential Learning Outcomes (see Figure 1).
- The Essential Learning Outcomes represent expectations of parents, policy makers, employers, and students themselves regarding student performance and college learning outcomes. (Sullivan, 2016; Association of American Colleges and Universities, 2013; Carnevale, 2009)

Daniel F. Sullivan is President Emeritus of St. Lawrence University, Canton, New York, and Senior Fellow, Association of American Colleges and Universities.

Kate Drezek McConnell is Senior Director, Research and Assessment, Association of American Colleges and Universities. Her work focuses on using course-embedded assessment to improve teaching and learning while also addressing accountability and accreditation requirements.

FIGURE 1. THE ESSENTIAL LEARNING OUTCOMES

Beginning in school, and continuing at successively higher levels across their college studies, students should prepare for twenty-first-century challenges by gaining:

▸ **KNOWLEDGE OF HUMAN CULTURES AND THE PHYSICAL AND NATURAL WORLD**

- Through study in the sciences and mathematics, social sciences, humanities, histories, languages, and the arts

Focused by engagement with big questions, both contemporary and enduring

▸ **INTELLECTUAL AND PRACTICAL SKILLS, INCLUDING**

- Inquiry and analysis
- Critical and creative thinking
- Written and oral communication
- Quantitative literacy
- Information literacy
- Teamwork and problem solving

Practiced extensively, across the curriculum, in the context of progressively more challenging problems, projects, and standards for performance

▸ **PERSONAL AND SOCIAL RESPONSIBILITY, INCLUDING**

- Civic knowledge and engagement—local and global
- Intercultural knowledge and competence
- Ethical reasoning and action
- Foundations and skills for lifelong learning

Anchored through active involvement with diverse communities and real-world challenges

▸ **INTEGRATIVE AND APPLIED LEARNING, INCLUDING**

- Synthesis and advanced accomplishment across general and specialized studies

Demonstrated through the application of knowledge, skills, and responsibilities to new settings and complex problems

NOTE: This listing was developed through a multiyear dialogue with employers and with hundreds of colleges and universities about needed goals for student learning; analysis of a long series of recommendations and reports from the business community; and analysis of the accreditation requirements for engineering, business, nursing, and teacher education. For more information, see Association of American Colleges and Universities (AAC&U), *College Learning for the New Global Century: A Report from the National Leadership Council for Liberal Education and America's Promise* (Washington, DC: AAC&U, 2007); *The LEAP Vision for Learning: Outcomes, Practices, Impact, and Employers' Views* (Washington, DC: AAC&U, 2011); or visit www.aacu.org/leap.

- These same constituencies increasingly question the value of a baccalaureate degree. They are demanding that institutions show their worth, often using metrics that are absolutely essential (e.g., graduation rates, post-graduate placements, etc.) but do not explicitly address two key questions: What are students learning? What are they able to do with that new knowledge, skill, or ability?
- Ensuring that students actually receive such an undergraduate education requires that we be able to assess their level of achievement of essential learning goals and, by extension, how well colleges, universities, and other "providers" of higher education enable this kind of learning.

First, the good news: due to increased attention from accreditors, the federal and state governments, and the public, it is difficult to find a college or university that does not have at least a nascent institutional process for—if not culture of—assessing student learning. Furthermore, since the early 2000s, there has been a push for more authentic, faculty-driven, and learner-centered assessment strategies linked to

an institution's actual curriculum. This is occurring on campuses of all types and sizes, mainly in response to concerns over the potential use of one-size-fits-all approaches, such as standardized tests (e.g., Brookhart, 1999; Huba & Freed, 2000; AAC&U, 2008; McConnell & Doolittle, 2012).

There is also bad news: the existence of an assessment office or Assessment Day on a campus—even those that espouse and practice a learner-centered, course-embedded, faculty driven approach—does not necessarily translate into improved student learning. Despite the near ubiquitous assessment mantra of "closing the loop" through data-driven decision making (Fulcher, et al., 2014), there is a gap between doing sound assessment and improving learning.

Why? For one very important reason: authentic assessment alone is not enough to ensure quality in higher education—faculty practices must change if college student learning of both disciplinary content and higher-order skills is to meet 21st century needs and expectations. These changes must include assessment processes that privilege faculty judgment while focusing on student learning.

In this essay we describe the most recent work of the AAC&U VALUE (Valid Assessment of Learning in Undergraduate Education) Initiative that has been promoting these assessment processes for a decade (Rhodes, 2016). There are now empirical results from one of three AAC&U beta test projects that involve over 90 public and private institutions. Next, we present examples from recently published research of some inexpensive, highly scalable changes in faculty student work assignment practices that produced strong, positive improvements in critical thinking in a multi-year project at Washington State. Finally, we argue that the way our concept of academic freedom has evolved in American higher education inhibits the creation of accepted evidence-based best practices in teaching and learning like the evidence-based "standards of care" physicians have created to ensure the ethical and effective practice of medicine.

VALUE Proof of Concept at Scale: The Multistate Collaborative to Advance Quality Student Learning

An initiative that began a decade ago, VALUE involves the use of rubrics created by and for faculty and other higher education professionals to score actual student work in college produced in response to assignments developed by faculty (Rhodes, 2016). This approach respects the work faculty ask students to complete as part of their academic program, while at the same time providing feedback to faculty on what works to improve student learning, especially for those least advantaged under current assessment practices. And it has been embedded in the educational practices of hundreds of colleges and universities every bit as much as the testing and grading most do now to assess student learning of content (Rhodes, 2016; Sullivan, 2015).

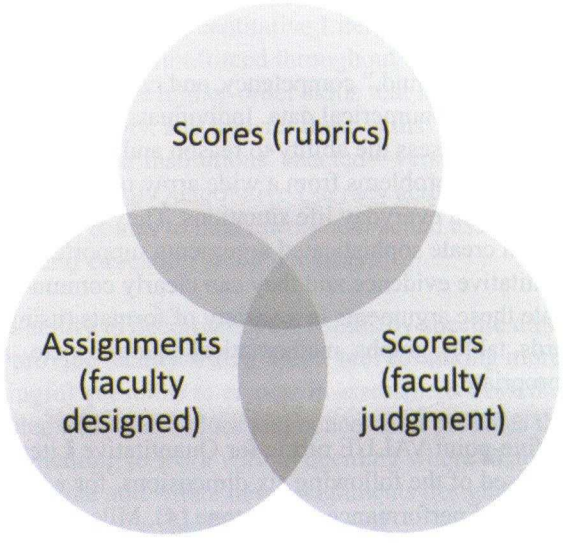
Three large beta tests—we describe them as "proofs of concept at scale"—are under way in twelve public state-systems (the Gates-funded Multi-State Collaborative), in a ten-institution Spencer-funded public/private collaboration in Minnesota, and in a Sherman Fairchild-funded collaboration among nine private colleges that are members of the Great Lakes Colleges Association.

The largest of the three is the Multi-State Collaborative to Advance Quality Student Learning (MSC), a project shared in partnership between AAC&U and the State Higher Education Executive Officers (SHEEO) association (please see http://www.aacu.org/value/msc for links to more information). From the project's first two years—the Pilot Year (2014–2015) and the Demonstration Year (2015–2016)—we have data from approximately 60 two- and four-year institutions across twelve states. We have collected over 14,000 work products from students who have all completed at least 75 percent of their two- or four-year courses of study. These work products were scored by trained faculty scorers using one of four VALUE rubrics: Civic Engagement, Critical Thinking, Quantitative Literacy, and Written Communication.

The VALUE *process* for the direct assessment of student learning used by the MSC is complex, not only methodologically, but epistemologically and pedagogically as well. The VALUE rubrics themselves represent one leg of the stool to the VALUE process, as they provide both the explicit articulation of the constituent dimensions of the essential learning outcomes as well as the progressively more sophisticated levels of possible student performance. The rubrics are only useful as an assessment tool in relation to the two other legs of the VALUE stool—student-generated work products and trained scorers, both of which leverage faculty expertise in instructional design, pedagogical strategies, and faculty's evaluative judgment of the student work in relation to the rubric (See Figure 2).

There is no existing well-tested model for what VALUE is attempting to do vis-à-vis the assessment of student learning. The variables and conditions that other approaches to assess student learning attempt to control or eliminate are—by

Figure 2. The VALUE Process



design—the variables and conditions that VALUE embraces, including:

- Using individual, faculty designed assignments taken straight off the syllabus and out of the classroom as evidence of student learning—there are no common prompts or scripts, nor are there standardized assignments relative to each rubric.
- Sampling processes that are designed to raise up, not wash out, the diversity of learners within and across institutions.
- Scorer training that is simultaneously calibration to a consensus score as well as professional development for participating faculty who represent the diversity of faculty roles on our campuses (e.g., contingent faculty, long-term instructor, tenure/tenure track, and other higher education professionals as appropriate).

Given this complexity, much of the emphasis in the first two years of the MSC has focused on establishing the methodological soundness of the approach, with significant attention dedicated to such psychometric concerns as validity, generalizability, and inter-rater reliability. While protocols for obtaining representative samples have been developed and will be implemented fully in the third, "Refinement Year" (2016–2017) of the MSC project, the data we have now are not from representative samples, and the results cannot be generalized back to a population. For purposes of our current discussion, however, the existing data allow us to see what is possible in terms of the generation, reporting, and use of assessment results, particularly with faculty.

Sample Results: Quantitative Literacy

By way of example, consider the results generated by the assessment of student learning with the Quantitative Literacy VALUE rubric during the Demonstration Year (2015–2016) for the MSC. A total of 1,363 work samples from participating institutions were submitted for scoring by 43 faculty. As defined in the VALUE rubric (for a detailed description as well as access to the rubric itself, please see: https://www.aacu.org/value/rubrics/quantitative-literacy), quantitative literacy:

> is a "habit of mind," competency, and comfort in working with numerical data. Individuals with strong QL skills possess the ability to reason and solve quantitative problems from a wide array of authentic contexts and everyday life situations. They understand and can create sophisticated arguments supported by quantitative evidence and they can clearly communicate those arguments in a variety of formats (using words, tables, graphs, mathematical equations, etc., as appropriate).

The five-point VALUE rubric for Quantitative Literacy is comprised of the following six dimensions, for which four levels of performance—Capstone (4), Milestone (3), Milestone (2), and Benchmark (1)—are defined, with "zero" scores an explicit option if students' work does not rise to the Benchmark (minimal) level of performance:

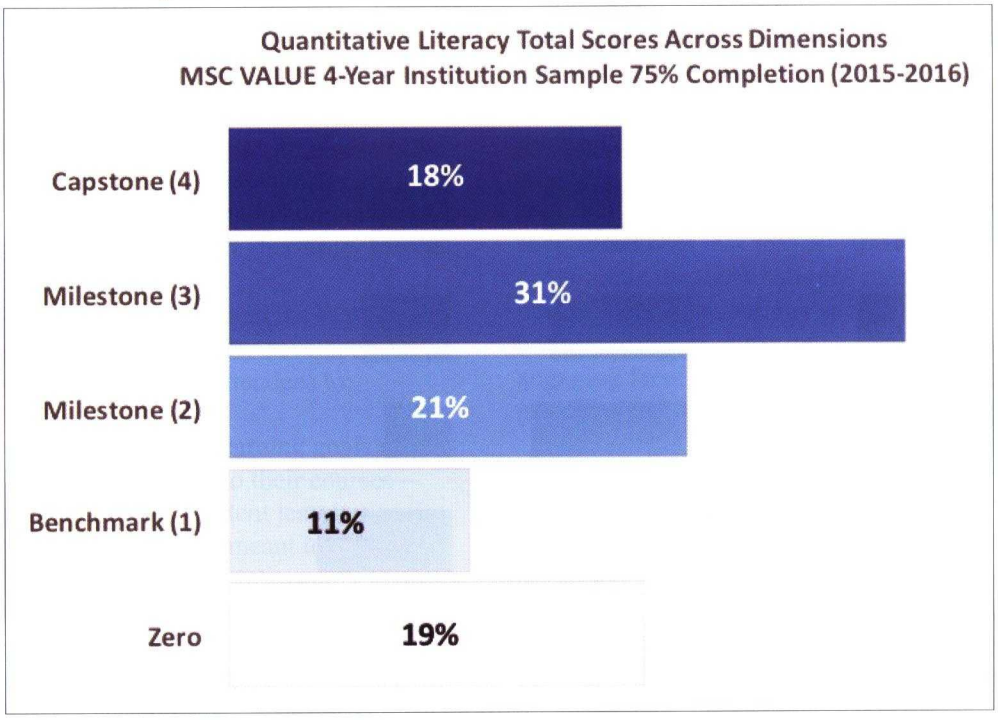
- **Interpretation:** *Ability to explain information presented in mathematical forms (e.g., equations, graphs, diagrams, tables, words)*
- **Representation:** *Ability to convert relevant information into various mathematical forms (e.g., equations, graphs, diagrams, tables, words)*
- **Calculation**
- **Application/Analysis:** *Ability to make judgments and draw appropriate conclusions based on the quantitative analysis of data, while recognizing the limits of this analysis*
- **Assumptions:** *Ability to make and evaluate important assumptions in estimation, modeling, and data analysis*
- ***Communication:*** *Expressing quantitative evidence in support of the argument or purpose of the work (in terms of what evidence is used and how it is formatted, presented and contextualized)*

Approximately 100 student work samples (with no more than 10 coming from any one course and/or faculty member) were collected on each MSC campus. The work samples were responses to an assignment that asked students to engage in and demonstrate dimensions of quantitative literacy. Faculty from all campuses were brought together and trained to score this student work. For Quantitative Literacy (QL) work samples in the Demonstration Year, most pieces of work were scored more than once. Work samples, assignments, and scorings were uploaded to a database designed for AAC&U by Task Stream, an assessment management system and ePortfolio technology provider. Scoring was done remotely from home campuses of faculty through the Task Stream platform. Faculty never scored work of students from their campus.

Figure 3 shows the percentage scores at each level of performance of QL student work products (averaging the scores of each of the six dimensions together) submitted by 4-Year Institutions. Again, although the data are not representative they do illustrate the kind of challenge we have in improving

> **The current data allow us to see what is possible in terms of the generation, reporting, and use of assessment results, particularly with faculty.**

Figure 3. Quantitative Literacy Scores, MSC Sample 2015–2016



our students' quantitative literacy. Only 49% of students overall scored at the Capstone (4) or Milestone (3) levels of competence in quantitative literacy.

But just like a standardized test score, the results in this figure really don't tell us anything that is actionable. That's where analyzing the results for each of the six dimensions of the Quantitative Literacy Rubric comes in. As shown in Figure 4 the various dimensions of the rubric allow us to consider students' scores, dimension by dimension.

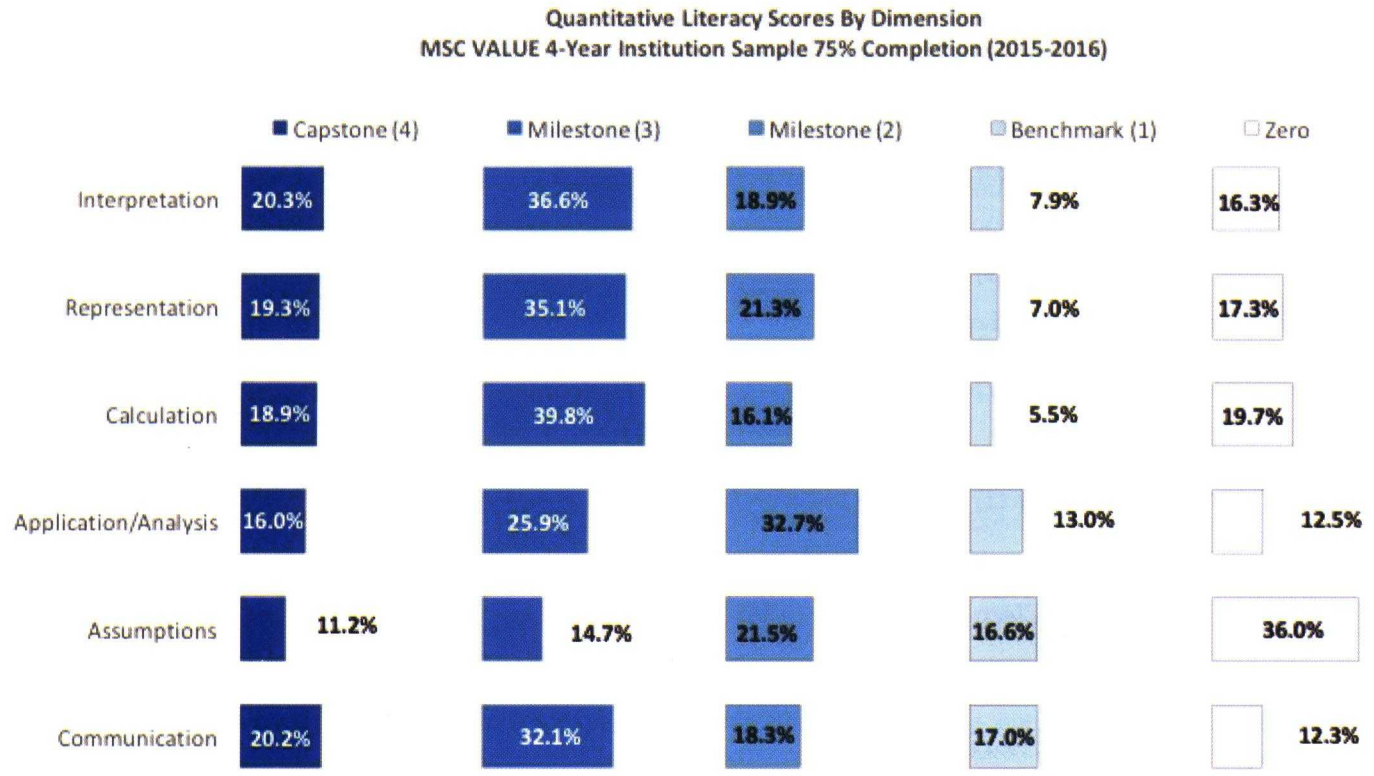
What initial observations can we make about these data? Students appear to do a reasonable job addressing the dimensions of Interpretation, Representation, and Calculation, with 57 percent of students' work scored at either the Capstone (4) or Milestone (3) levels for Interpretation, 54 percent for Representation, and 59 percent for Calculation. In the scoring distribution on the Application/Analysis dimension, however, only approximately 42 percent of students scored either at the Capstone (4) or Milestone (3) levels. For the Assumptions dimension, the percentage drops to approximately 26 percent. Interestingly, Assumptions also has the largest percentage (36%) of student work scored as a Zero. Application/Analysis and Assumptions may arguably be higher-order forms of reasoning than Interpretation, Representation, and Calculation, at least partially explaining the lower levels of success.

Even given disclaimers about generalizability, should we interpret these data to mean that students from the participating four-year institutions are not quantitatively literate? Not necessarily – and certainly not from one data point in time! These initial results raise more questions than they answer. However, unlike the overall average results for this rubric or results generated by many of the commercially available standardized testing options, faculty can in our experience engage with and use these data immediately. Faculty members almost never look at data generated by a VALUE rubric and blame the students. On the contrary data such as these raise questions like:

- Broadly, was the assignment designed to address all of the underlying dimensions captured by the VALUE rubric?
- What does the "zero" mean for Assumptions? Is it a reflection of students' skills in this area or the **absence of an opportunity** for the student to demonstrate their skills because of the assignment design?
- Our students seem to do well on the "lower order" skills associated with Quantitative Literacy. Are those skills practiced and reinforced throughout our curriculum more than the higher-order skills that help students synthesize, apply, and transfer these skills?
- What do our institutions' enacted—versus intended — definitions of Quantitative Literacy actually look like when it comes to coursework and actual assignments in general education and the majors?

The ability to deconstruct the results by specific dimensions provide faculty with a useful tool for asking more meaningful questions to engage in assessment work that resonates with their values and concerns. It provides them with something to work with in deciding where to allocate teaching time and where new assignments designed to improve competence should be developed.

Figure 4. Quantitative Literacy Scores by Dimension, MSC Sample 2015–2016



Additionally, data such as these allow for a more nuanced, fine-grained approach to "closing the loop" to make improvements in student learning. Depending upon your vantage point, the conclusions possible using only the total average scores for Quantitative Literacy from Figure 3 could be either a glass half empty or glass half full, with 49 percent scoring at Capstone (4) or Benchmark (3). Using the data by dimension, however, institutions could determine that Assumptions in particular is a key area for concern, leading to an investigation of everything from student preparation and motivation to assignment design and curricular mapping. In short, investing in a VALUE-Rubric-based assessment strategy can lead to a continuous quality improvement cycle because it grounds the entire assessment conversation in the curriculum, making it legitimately part of the work of the faculty.

Where will individual institutions and AAC&U analysts go next with these data? Individual institutions get their own data back along with the average results over the whole group of institutions in the project, so they will be able to benchmark against the overall group to see where they stand. Although we have illustrated results only from work involving quantitative literacy assignments, each MSC campus will also have results from work scored for Civic Engagement, Critical Thinking, and/or Written Communication; in addition to these outcomes, the GLCA and Minnesota Collaboratives also have work scored for Intercultural Knowledge and Competence, with many submitting for Ethical Reasoning in 2016–2017. Also under way is collection and scoring of student work completed by students at roughly 25 percent of the way through their two-year or four-year degree program. We will want to see if work collected later in students' time in college scores better than work from the one-quarter mark.

But the ultimate goal for participating institutions is to follow cohorts of students over time—how much do individual students, on average, gain on key dimensions of student learning while in college? AAC&U is seeking funding now to allow institutions ready to work with us on this to take their assessment work to the next level.

It was clear to us from the start that good assignments are crucial to improvement in student learning. Therefore we collected not just the student work itself but the assignments to which the work was a response. For example, how much of student performance on Assumptions assessed with the QL Rubric was a function of not being asked to think about assumptions as they approached the assignment, or not being given ahead of time a sense of what high quality quantitative analysis is? Therefore we are also seeking support to assess the quality of the assignments associated with the student work that has been scored. What do high quality QL, or other higher-order learning goal, assignments look like? When student work was produced in response to higher quality assignments was the work better? Research from a different context gives an emphatic "yes" to this last question.

Assessment by Itself is Not Enough—Faculty Practices Must Change

An important recent book (Condon et al., 2016) makes clear that if faculty practices change in the right ways, higher-order learning by students improves dramatically. The authors describe a Spencer Foundation-funded collaboration between faculty at Washington State University and Carleton College in Minnesota to understand the relationship between faculty development and higher-order learning among undergraduate students.

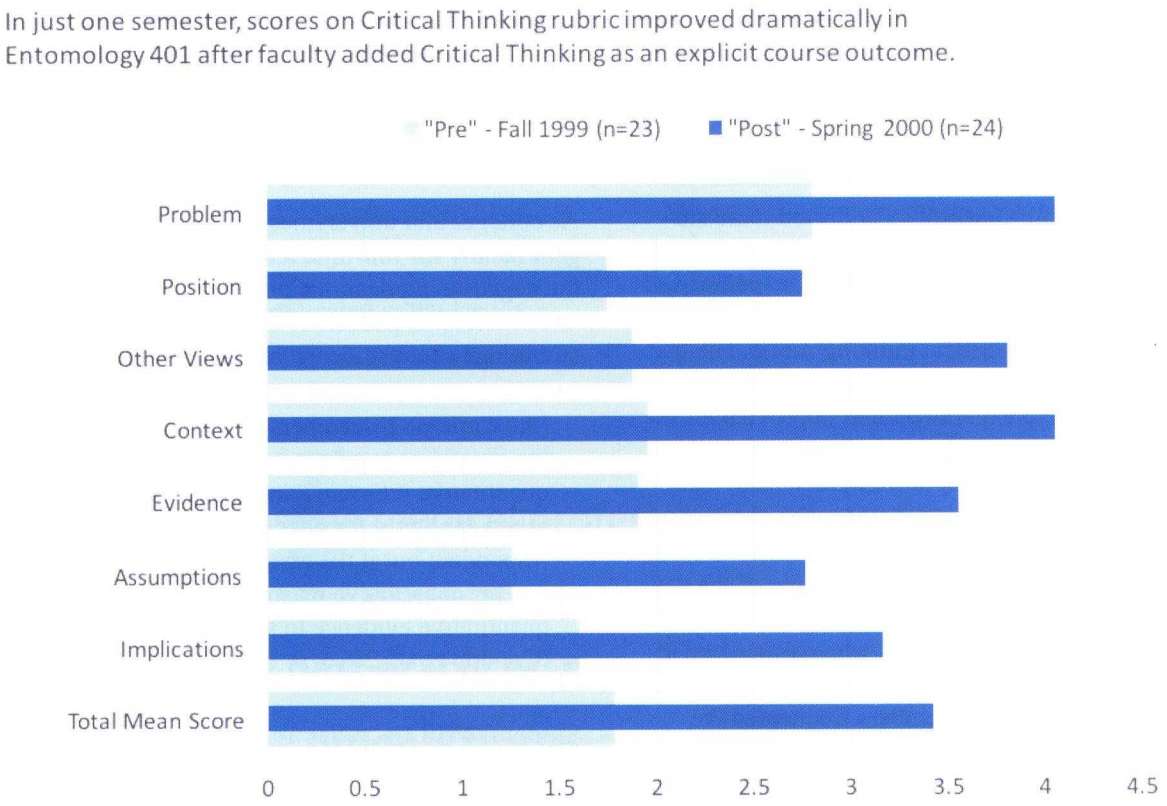
From 1999 to 2003 William Condon and his Washington State colleagues began conducting faculty development workshops focused on getting faculty attendees to:

- Explicitly introduce higher-order learning goals involving writing and critical thinking into their courses—above and beyond disciplinary content learning goals.
- Design and introduce assignments meant to help students improve their skills in those higher-order learning areas and building on the disciplinary learning they also wanted students to gain—i.e., to start with the learning goals and create an assignment pathway to improve student skills in all of the sub-dimensions of the learning goal.
- Create a common rubric so that students' actual work could be assessed for achievement of the overall learning goal and the sub-dimensions that made it up.

They then followed up to see which faculty did these things, collected samples of student work from faculty who attended the workshops more and less frequently and who implemented the recommendations of the workshops more and less frequently, and then had independent scorers rate the work samples using the rubrics that were developed. As the authors say, "The unifying theme for both campuses focuses on changing *faculty* practices by engaging faculty in a learning process about best teaching practices" (p. 94).

Figure 5 (p. 97) shows just one set of results: scores for the sub-dimensions of the Critical Thinking rubric developed at Washington State and the overall Critical Thinking average score for student work obtained from a fall 1999 upper level entomology course and the spring 2000 version of the same course taught by the same faculty member. The fall 1999 version of the course did not have critical thinking as an explicit learning goal and assignments designed to improve students' critical thinking skills while the spring 2000 version did. The results are dramatic—the critical thinking

Figure 5. Pre-Post Scores on Critical Thinking Rubric After Making Critical Thinkink Rubric Outcome Explicit



Comparing results from courses taught by faculty who did and did not attend the faculty development workshops and, if they did attend, did or did not implement the recommended best practices. In all cases results were much better when best practices from faculty development efforts were implemented.

average score was over 1.5 scale points higher in the spring 2000 version of the course.

Note too that the highest student scores in the fall 1999 version were achieved on the "Problem" sub-dimension—just stating the problem—while the "Evidence", "Assumptions," and "Implications" scores—all of which are higher-order dimensions of this learning goal—were much lower. In the spring 2000 results scores on those same dimensions were much higher.

There are multiple examples like this in the book, comparing results from courses taught by faculty who did and did not attend the faculty development workshops and, if they did attend, did or did not implement the recommended best practices. In all cases results were much better when best practices from faculty development efforts were implemented.

Similarly, in 2009–2010 the project collected more student work from courses taught by the faculty who did and did not attend workshops in the 1999–2003 time period and did and did not implement best practices. They assessed that student work using the critical thinking rubric. Faculty who implemented best practices in the earlier time period continued on their own to experiment and improve their assignments, and their students' critical thinking scores were even better relative to students of non-implementers than they were in 1999–2003. There is not just a one-time improvement from this work that plateaus at the same level but a continuously increasing impact on the learning of the students of faculty adopters of these best practices (p. 99).

Another interesting finding involves who participated in faculty development efforts and implemented best practices. Adjunct and temporary faculty members frequented the faculty development workshops the most and non-tenured tenure track faculty participated the least. But it was tenured faculty who most frequently used the "best practice" advice of the workshops to modify their courses.

> Most adjunct and temporary faculty practiced what the researchers came to identify as "defensive" faculty development: they felt the need to participate in a very large number of development opportunities (some as many as 20 per year) so that their appointments would be renewed. While that need stimulated participation, it was paired with an avowedly cautious approach to experimentation in the classroom, lest student course evaluations go down. . . . the tenured faculty were far more willing to experiment with new assignments or techniques, since they did not have to worry about a temporary dip in course evaluations. Thus, for that set of high-participating faculty, higher critical thinking scores result from the extra freedom to experiment (Condon et al., 2016, p. 101).

As colleges and universities increasingly substitute contingent faculty for full-time tenure-track and tenured faculty, the negative implications for student learning are clear.

The Carleton portion of the research also indicated important things about students' response to assignments:

> [the Carleton researchers] sampled Carleton's writing portfolios to assess whether students were using quantitative reasoning and found that if an assignment asked explicitly for quantitative reasoning, students responded; but if the assignment did not specifically cue for quantitative reasoning, students rarely used quantitative evidence to support and clarify assertions (Condon et al., 2016, p. 103).

Students Will Do What You Ask Them to Do, So Ask!

Similarly, it is very interesting that Carleton students themselves saw the relationship between faculty best practices and the quality of their writing:

> Rutz and Lauer-Glebov (2005) established at Carleton that the likelihood that a student would choose a paper from a given faculty member's course for the student's writing portfolio was in direct proportion to the number of faculty development events the faculty member had attended. In effect, students endorse their teachers' learning by voting with their papers. (Condon et al., 2016, p. 109)

And finally, this overall conclusion: "… the connection between changes faculty make in their assignments and changes in students' learning outcomes is clear. When faculty improve their teaching methods—or merely improve an assignment—improvement in student learning is detectable" (p. 112).

It is astonishing that we have not previously developed this kind of powerful evidence for what each of us probably suspects must be the case. It seems so simple.

- Decide on what you want students to learn and make those learning goals explicit in curricula and syllabi.
- Create assignments specifically designed to lead students to learn what you intend them to learn.
- Develop means to assess whether students have learned what you want them to learn.
- Act on the results of the assessments to improve student learning.

What Condon et al. have demonstrated is how to make significant improvements in critical thinking, an important kind of higher-order learning, at minimal cost, and similar results occurred at Carleton. These changes in faculty practices are scalable within the existing system of higher education.

Academic Freedom and Pedagogy—Learning From the Profession of Medicine

We in higher education—as a profession—need to decide that we are going to make the effort to adopt evidence-based best practices broadly. The students of the tenured professors at Washington State who participated in the faculty development program outperformed the students of other faculty, but only a small portion of Washington State's tenured faculty were regular participants in it.

The way our interpretation of academic freedom has evolved wrongly conflates the freedom faculty do and must have regarding the teaching of disciplinary content, including content that is "controversial," with the notion that they should be equally free, as individuals, to decide on pedagogical approaches—even when strong evidence favors some kinds of pedagogy over others.

The 1940 AAUP Statement on Academic Freedom says nothing about pedagogy; it's all about content. However, some more recent statements (Nelson, 2010) argue (and it is often assumed) that an individual faculty member has the freedom to choose his or her pedagogical strrategy without reference to anything like a "standard practice" or "standard of care."

In contrast, the profession of medicine has a concept of "standard practice." The American College of Physicians oversees a process whereby physicians themselves create and vet Clinical Practice Guidelines and Guidance Statements. These are evidence-based best practices clarifying the treatment every patient has a right to expect—the best available treatment. (Qaseem et al., 2010). Physicians are not free to use techniques or therapies known to produce inferior results when there is clear evidence that standard practice is superior, and they are not free to experiment with new techniques or therapies without patients' consent or outside the accepted ethical protocols.

Of course—and this is absolutely critical—it is the medical profession that oversees the rigorous process and creates consensus on best practices, not politicians, not institutional trustees, or others unprepared by education and experience to do so. As a profession, physicians take the responsibility to do it, publish their analyses and their conclusions, and thereby help create societal trust that medicine is committed to a cycle of continuous quality improvement.

Maybe we in higher education do not have a consensus on what truly "best practices" for facilitating and supporting student learning look like because we have spent so much less time assessing the effectiveness of our practices in

"It is the medical profession that oversees the rigorous process and creates consensus on best practices, not politicians, not institutional trustees, or others unprepared by education and experience to do so.

teaching and learning. In response, other stakeholders have unfortunately stepped into the void.

But as books like Condon et al. make clear, we are entering a new age. We know more about "what works" in teaching and learning with each day. Like the medical profession, when there is clear evidence that some forms of pedagogy work better than others, we believe faculty—nationally, in some kind of organized way—need to figure out how to declare that to be so.

AAC&U is leading the way to a new world in higher education where we can assess even higher-order student learning in ways that respect the complexity of what we do and what we need students to learn. We now have ways to assess the effectiveness of our practices for facilitating student learning. We need to embrace this revolution—this "VALUE Breakthrough" (Sullivan, 2015a)—to respond to the large portion of the nation that has lost faith in higher education. We have to be clear about our learning goals; we have to build them into the design of our curricula and syllabi; we have to create assignments and practice regimens intended to facilitate student learning of the outcomes we have chosen; we have to assess the results in a way worthy of our enterprise; and we have to use assessment results to improve our practices. We can and should do this. C

Resources

- Association of American Colleges and Universities. (2008). *College Learning for the New Global Century.*
- Association of American Colleges and Universities. (2013). Hart Research Associates, "It Takes More than a Major: Employer Priorities for College Learning and Student Success."
- Brookhart, S. M. (1999). *The art and science of classroom assessment: The missing part of pedagogy.* San Francisco, CA: Jossey-Bass.
- Carnevale, A. (2009). "The Economic Value of Liberal Education," Association of American Colleges and Universities, (June).
- Condon, W., Iverson, E. R., Manduca, C., Rutz, C., & Willett, G. (2016). *Faculty Development and Student Learning: Assessing the Connections.* Indiana University Press, Bloomington, Indiana.
- 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," National Institute for Learning Outcomes Assessment, University of Indiana Urbana-Champaign, Occasional Paper #23.
- Huba, M. E., & Freed, J. E. (2000). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning*. Boston, MA: Allyn & Bacon.
- McConnell, K.D. & Doolittle, P.E. (2012). Classroom-level assessment: Aligning pedagogical practices to enhance student learning. In C. Secolsky & D.B. Denison (Eds.), *Handbook on Measurement, Assessment, and Evaluation in Higher Education*. New York, NY: Routledge.
- Nelson, C. (2010). "Defining Academic Freedom," *Inside Higher Ed*, December 21.
- Qaseem, A., Snow, V., Owens, D. K., & Shekelle, P. (2010) "The Development of Clinical Practice Guidelines and Guidance Statements of the American College of Physicians: Summary of Methods," *Annals of Internal Medicine*, 3 August, Vol. 153, No. 3.
- Rhodes, T. (2016) "The VALUE of Assessment: Transforming the Culture of Learning," *Change, 48*(5), 36–43.
- Sullivan, D. F. (2015). *The VALUE Breakthrough: Getting the assessment of student learning in college right.* Washington, DC: Association of American Colleges and Universities.
- Sullivan, D. F. (2016). "AAC&U Has It Right—It's 'Liberal Education,'" Association of American Colleges and Universities, Washington, DC. *Liberal Education* (Summer).