It's the Assignments—A Ubiquitous and Inexpensive Strategy to Significantly Improve Higher-Order Learning

Daniel F. Sullivan & Kate Drezek McConnell

Published online: 02 Nov 2018.

Submit your article to this journal ⬚

Article views: 603

View related articles ⬚

View Crossmark data ⬚

Citing articles: 1 View citing articles ⬚

IT'S THE ASSIGNMENTS—

A Ubiquitous and Inexpensive Strategy

to Significantly Improve Higher-Order Learning

By Daniel F. Sullivan and Kate Drezek McConnell

There is now a rich research literature documenting the positive and lasting impact of "high-impact educational practices" (HIPs) on undergraduate student learning and success, from improved retention and graduation to students' sense of belonging and engagement, in all kinds of undergraduate settings (e.g., Kuh, 2008; Kuh, O'Donnell, & Reed, 2013; Kuh, O'Donnell, & Schneider, 2017). Unfortunately, the quality of the implementation of HIPs has limited their impact, and too few students are able to take advantage of them (Kuh, O'Donnell, & Reed, 2013, pp. 7–8, pp. 3–5).

In addition, disadvantaged students who stand to benefit the most from HIPs gain access to them least often (Kuh, 2008, p. 17; Finley & McNair, 2013). In response, faculty and campus leaders have sought to find ways to embed aspects of HIPs in existing classes and labs (Kuh, O'Donnell, & Reed, 2013, 10) to both lower their cost and expand their reach (Kuh, O'Donnell, & Reed, 2013); there are increasing examples of success in that effort.

That said, simply labeling a practice "high-impact" does not make it so. As part of this work, Kuh and his colleagues delineated eight characteristics of quality high impact practices (Kuh, O'Donnell, & Reed, 2013). Several of these

In Short

- When course assignments exhibit the same characteristics as high-quality HIPs—e.g., "performance expectations set at appropriately high levels" and involving "significant investment of time and effort by students over an extended period of time"—the quality of student work is much higher.

- Demanding assignments that are intentional about both higher-order learning and disciplinary content goals are an affordable, reasonable strategy for enhancing student learning.

- Effective assignments are an educational best practice that is highly scalable, broadly inclusive, and ubiquitous. They offer real hope for addressing the gaps that disadvantaged students experience in access to powerfully effective learning opportunities.

- These results remind us that student learning needs to be iterative and intentionally scaffolded across the curriculum to provide higher-order practice and achieve competence, and that essential learning occurs in multiple courses and programs, not exclusively in a particular class.

Daniel F. Sullivan is President Emeritus, St. Lawrence University and Senior Fellow, Association of American Colleges and Universities (AAC&U). He collaborates in AAC&U's VALUE Project.

Kate Drezek McConnell is Assistant Vice President, Research and Assessment, AAC&U. Her work focuses on using course-embedded assessment to improve teaching and learning while also addressing accountability and accreditation requirements.

characteristics are not simply markers of quality HIPS; they are the essence of good teaching.

In this article, we report convincing results from assessments of actual student work as part of the Association of American Colleges and Universities (AAC&U) VALUE—Valid Assessment of Learning in Undergraduate Education—Project. These assessments suggest that when course assignments exhibit the same characteristics as high-quality HIPs—e.g., "performance expectations set at appropriately high levels" and requirement of "significant investment of time and effort by students over an extended period of time" (Kuh, O'Donnell, & Reed, 2013, 8)—the quality of that student work is much higher than when those conditions are absent. Strong support for the importance of assignments, albeit from a single collegiate setting, is found in Condon and colleagues (2016) and summarized in our recent report on the VALUE Project (Sullivan & McConnell, 2017, pp. 21–23).

Insisting that faculty give assignments that are both appropriately demanding and intentional about higher-order learning goals as well as disciplinary content learning goals may be perhaps politically and managerially difficult. However, doing so is not only effective at improving students' higher-order learning, it also represents an affordable, reasonable

> ## "Assignments may be the game changer for student learning and for inclusive excellence in undergraduate education today.

strategy for enhancing student learning. Assignments are the pedagogically purposeful, faculty-created learning interventions that comprise a traditional class and can serve as the building block for any number of HIPs, from capstone courses to undergraduate research experiences.

Effective assignments are an educational best practice that is highly scalable, broadly inclusive, ubiquitous and—frankly—too often ignored. More effective assignments offer real hope for addressing the gaps that disadvantaged students experience in access to powerfully effective learning opportunities. In short, assignments may be the game changer for student learning and for inclusive excellence in undergraduate education today.

Operationalizing "Assignment Difficulty"

One of AAC&U's signature initiatives, VALUE is a campus-based assessment approach best known for its suite of 16 rubrics covering the Essential Learning Outcomes, from critical thinking and written communication to civic engagement and integrative learning that all students need for success in work, citizenship, and life. Over a decade ago, interdisciplinary teams of faculty and other educational professionals from institutions across the country—two- and four-year, private and public, research and liberal arts, large and small—developed the rubrics.

In the last four years alone, the rubrics have been downloaded more than 61,000 times by individuals from almost 5,900 organizations, including nearly 2,200 colleges and universities. Significantly, the rubrics provide a mechanism for measuring learning using students' own authentic work—work that originated in the college classroom in response to an assignment designed and implemented by individual faculty members.

While originally designed as tools for promoting authentic assessment efforts at the local, campus level, the VALUE

rubrics also served as the foundation for a multi-state and multi-campus experiment in student learning outcomes assessment. Between 2014 and 2017, approximately 100 2- and 4-year public and private institutions submitted samples of their students' work to a database for scoring by trained faculty from across the country. Beyond serving as an initial "proof of concept" for this approach at a multi-institutional level, the VALUE project generated results that allow us to interrogate important, nuanced questions about undergraduate teaching, learning, and assessment.

Specifically, two components of the overall VALUE Project—a Sherman Fairchild Foundation-funded collaboration among nine private colleges that are members of the Great Lakes Colleges Association (GLCA) and a ten-institution Spencer Foundation-funded public/private collaboration in Minnesota—afforded us the opportunity to explore the role of assignments in the VALUE approach more fully, including questions of intended level of rigor. Borrowing language from the practice of curricular mapping, we created a coding framework for assignment difficulty that also mapped onto the performance levels articulated by the VALUE rubrics. For all VALUE Rubrics 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 exhibit the Benchmark (minimal) level of performance. The four levels are meant to mark developmental stages in a student's path toward the competence (Capstone level), which the rubric implies students should reach by the time they complete college.

The eight-point scale shown below allowed faculty whose students' work was selected for inclusion in the GLCA and Minnesota VALUE pilots to say at what level of expectation their assignment was targeted along this developmental sequence. Faculty were asked to code their assignment's "difficulty" using this classification framework.

The resulting scale has scores that range from 1 to 8. The idea is that when students are at the "benchmark" level of competence on a higher-order learning goal like critical thinking, they need assignments that introduce them to what critical thinking is, while at the capstone level, students need assignments that ask them to demonstrate mastery of the outcome. In other words, whether an assignment is an effective one "lies in its ability to motivate relevant cognitive, behavioral, social and affective processing, not in the mere completion of the activity; that is, a critical thinking activity is only beneficial if the learner actually thinks critically during the execution" of it (McConnell & Doolittle, 2017, p. 62).

INTRODUCE *Assignment designed to introduce the outcome*		PRACTICE *Assignment designed to afford student practice with the outcome*		REINFORCE *Assignment designed to reinforce previously practiced outcome*		MASTERY *Assignment designed for students to demonstrate level of mastery of the outcome*	
1	2	3	4	5	6	7	8
BENCHMARK		MILESTONE (2)		MILESTONE (3)		CAPSTONE	

With this scale in hand we could then ask what expectations faculty had for their students on critical thinking and written communication.

But why create a second scale for assignment difficulty? Why not simply use the scale and language of the VALUE rubrics for characterizing assignment difficulty as well? It is important that we not confound aspects of faculty intentionality in the assignment design process with evidence of student learning. The VALUE rubrics are a suite of direct performance assessment tools in that they were specifically designed to allow faculty to identify evidence of student performance on an assignment against multiple dimensions of an essential learning outcome and, accordingly, provide the score that best described the quality of the learning demonstrated through the student's work.

In contrast, the assignment difficulty scale allows faculty to self-report their intended overall level of rigor or difficulty for the assignment they construct and have their students complete. We were concerned that utilizing the same scale for differing purposes would create confusion or lead to unhelpful leaps of logic. For example, while we argue that for students to demonstrate capstone-level learning they need to encounter an assignment at higher levels of intended difficulty (e.g., "Reinforce" or "Mastery"), there is no guarantee that a mastery-level assignment will elicit capstone-level demonstrations of learning from all or any of the resulting student work.

AAC&U's Private College VALUE Project— Exploring the Relationship Between Assignment Difficulty and VALUE Rubric Scores

Data from the 14 private institutions participating in the GLCA and Minnesota Collaboratives were utilized for this analysis. The institutions submitted actual student work across the spectrum of the undergraduate experience (e.g., first-year students, students at the mid-point of their credit hour completion, and students nearing graduation) for scoring against the VALUE rubrics for Critical Thinking and Written Communication. Participating institutions also provided key student demographic information as well as critical information about the courses and assignments that generated the student work.

As a result, we know whether the student work in our dataset comes from lower-division (e.g., 100–200 level) or upper-division (e.g., 300–400 level) coursework and if the faculty intended for their assignment to address each of the dimensions of the VALUE rubric under consideration. Assignment difficulty was introduced as a concept in Year 2 of the three-year projects. Some campuses were able to work assignment difficulty into their protocols and others were not.

In total, thirteen of the 14 private colleges provided 1,773 pieces of student work for scoring against the Critical Thinking VALUE Rubric, with 9 of the 14 also providing assignment difficulty data on approximately 50% (880 pieces) of the student work. For Written Communication, all

> *The assignment difficulty scale allows faculty to self-report their intended overall level of rigor or difficulty for the assignment they construct and have their students complete.*

14 of the private colleges provided 1,951 pieces of student work for scoring, with 7 of the 14 campuses able to provide assignment difficulty information for 33% (641 pieces) of the submitted student work. All pieces of student work were scored according to the VALUE Initiative scoring protocols (McConnell & Rhodes, 2017).

It is important to note that faculty never scored work of students from their own campus and were not privy to any information about the student or the assignment that generated the student work (e.g., faculty and/or course information, assignment directions, etc.). The actual work samples, the assignments themselves, the VALUE rubric scores, and the other data describing the course and some student characteristics were all uploaded to the databases we analyze here.

Our Results

As defined for the VALUE rubric, critical thinking: "is a habit of mind characterized by the comprehensive exploration of issues, ideas, artifacts, and events before accepting or formulating an opinion or conclusion."

The Critical Thinking VALUE Rubric is comprised of the following five dimensions (for a detailed description as well as access to the rubric itself, please see: https://www.aacu.org/value/rubrics/critical-thinking):

- **Explanation of Issues**
- **Evidence**
- **Influence of context and assumptions**
- **Student's position (perspective, thesis/hypothesis)**
- **Conclusions and related outcomes (implications and consequences)**

As defined for its VALUE rubric, written communication "is the development and expression of ideas in writing. Written communication involves learning to work in many genres and styles. It can involve working with many different writing technologies, and mixing texts, data, and images. Written communication abilities develop through iterative experiences across the curriculum."

Table 1. Mean Scores across All Five Dimensions of Critical Thinking and Written Communication

Performance Level	Critical Thinking	Written Communication
Capstone (3.2–4.0)	12% (211)	23% (443)
Milestone (2.2–3.0)	35% (615)	43% (835)
Milestone (1.2–2.0)	40% (707)	31% (606)
Benchmark (0.0–1.0)	13% (240)	3% (67)
Total	100% (1,773)	100% (1,951)

Table 2. Critical Thinking and Written Communication Overall Outcome Means by College Credits Earned

Credits Earned	Critical Thinking	Written Communication
0–30	1.92 (690)	2.37 (738)
31–60	2.29 (171)	2.57 (293)
61–90	2.29 (157)	2.62 (151)
90+	2.12 (566)	2.55 (552)
Mean Difference	+.20	+.18

The Written Communication VALUE Rubric is comprised of the following five dimensions (for a detailed description as well as access to the rubric itself, please see: https://www.aacu.org/value/rubrics/written-communication):

- **Context of and Purpose for Writing**
- **Content Development**
- **Genre and Disciplinary Conventions**
- **Sources and Evidence**
- **Control of Syntax and Mechanics**

Again, see the actual rubrics for how each level of performance is defined in the Critical Thinking and Written Communication rubrics.

Table 1 shows, for Critical Thinking and Written Communication, the percentage of scores at each level of performance, averaging the scores of the five dimensions together for each student. Recall that the scores for individual dimensions range from "four," where students were assessed to have reached the "Capstone" level of performance, to "zero," where students' work does not exhibit the Benchmark (minimal) level of performance.

To be classified at the Capstone level for this analysis we decided that a student needed to have a score of "4" on at least one of the five dimensions—an average score of 3.2 or higher. Using this method of averaging the scores, 47% of students meet a Capstone or Milestone (3) standard on

> **Taking an upper division course doesn't guarantee that all assignments will be demanding in the ways we are using the term 'Mastery.'**

critical thinking and 66% meet a Capstone or Milestone (3) standard on written communication.

Even though we do not have cohort data—which means that we do not know if individual students made learning gains—we can still look to see if work produced by students later in their college careers, on average, scores higher on the Critical Thinking and Written Communication VALUE Rubrics. For a large fraction of the students whose work we have in our databases institutions gave us the number of credits students had earned by the time the work was produced.

In Table 2 we show the outcome means separately for students who had earned 0–30, 31–60, 61–90, and 90+ credits by the time they produced the work we have. In both cases what we see is a curvilinear relationship—the mean goes up and then goes down. The scores of seniors are somewhat higher on average than the scores of freshmen but lower than the scores of sophomores and juniors. Why aren't the scores of seniors much higher? Might it really be the case, a la Arum and Roksa (2011), that most students' higher-order skills do not improve in college?

We know two things about the student work in our databases that might help sort this out—we know whether the course in which the work was done was upper division (junior-senior level) or lower division (freshman-sophomore level), and we know for at least a fraction of the work products how demanding the assignment was. We assume, generally, that work expectations are higher in upper division courses than lower division courses, and we know that very little of the senior work in our databases came from lower division courses; similarly, very little freshmen work came from upper division courses. That might explain it— seniors are taking harder courses and that keeps them from scoring at the Capstone level in critical thinking or written communication.

With our corresponding 8-point scale for assignment difficulty in hand we can then ask what expectations faculty teaching upper division courses had for their students on critical thinking and written communication compared to those teaching lower division courses. Upper division critical thinking assignments were 24 percentage points more likely and upper division written communication

TABLE 3. CRITICAL THINKING AND WRITTEN COMMUNICATION OVERALL OUTCOME MEANS BY COURSE LEVEL

Course Level	Critical Thinking	Written Communication
Lower Division	2.02 (898)	2.37 (974)
Upper Division	2.12 (875)	2.60 (977)
Mean Difference	+.10	+.23

TABLE 4. CRITICAL THINKING AND WRITTEN COMMUNICATION OVERALL OUTCOME MEANS BY ASSIGNMENT DIFFICULTY

Assignment Difficulty	Critical Thinking	Written Communication
Introduce	1.70 (71)	2.05 (88)
Practice	2.03 (285)	2.43 (300)
Reinforce	2.21 (387)	2.59 (207)
Mastery	2.31 (137)	2.96 (49)
Mean Difference	+.61	+.91

assignments 19 percentage points more likely to involve Mastery level expectations on the part of faculty than lower division assignments. But, by faculty members' own admission, **only 31% of critical thinking and 21% of written communication upper division assignments were actually at the Mastery level.** Taking an upper division course doesn't guarantee that all assignments will be demanding in the ways we are using the term "Mastery."

When we examined the relationship between credits earned and assignment difficulty we saw the same pattern. Critical thinking assignments completed by seniors were 30 percentage points more likely than assignments completed by freshmen to be explicitly designed for students to demonstrate level of mastery of the outcome, **but only 39% of seniors' assignments were at this level.** Written communication assignments completed by seniors were 22 percentage points more likely than assignments completed by freshmen to be at the Mastery level, **but only 26% of seniors' assignments were deemed sufficiently difficult or rigorous to provide students with the opportunity to demonstrate mastery.**

Our puzzle, recall, is why we see in Table 2 that seniors' outcome means for Critical Thinking and Written Communication are lower than the means for juniors and sophomores. Seniors' assignments come from upper division courses

almost exclusively, but—and this is part of the answer—students completing upper division assignments, as Table 3 shows, have a Critical Thinking outcome average that is essentially identical to the outcome average of lower division students, and a Written Communication outcome average that is only moderately higher than the outcome average of lower division students.

What's really going on, as Table 4 indicates clearly, is that **regardless of class year, students completing assignments that are more demanding achieve higher scores on the Critical Thinking and Written Communication Rubrics.** The mean Critical Thinking score when the assignment was designed for Mastery was .61 higher than when the assignment was designed to Introduce, and the mean Written Communication score when the assignment was at the Mastery level was .91 higher than when the assignment was at the introductory level. When asked to do more, students did more.

Similarly, as Table 5 shows, where we combined the Introduce and Practice levels to increase the number of cases for analysis, the mean outcome score for seniors was higher when they were asked to do Mastery level work on both Critical Thinking and Written Communication. For Critical

TABLE 5. CRITICAL THINKING AND WRITTEN COMMUNICATION OVERALL OUTCOME MEANS BY ASSIGNMENT DIFFICULTY AND CREDITS EARNED

CT Assignment Difficulty	Credits Earned				Mean Diff.
	0–30	31–60	61–90	91+	
Introduce + Practice	1.87 (177)	1.90 (151)	2.39 (42)	2.03 (31)	+.16
Reinforce	2.17 (184)	2.28 (57)	2.29 (35)	2.32 (65)	+.15
Mastery	1.61 (37)		2.65 (16)	2.75 (70)	+1.14
Mean Difference	−.18	+.38	+.26	+.72	
WC Assignment Difficulty					
Introduce + Practice	2.25 (261)	2.52 (78)	2.54 (16)	2.82 (22)	+.57
Reinforce	2.54 (104)	2.38 (33)	2.62 (25)	2.86 (45)	+.32
Mastery	2.90 (12)	2.44 (10)		3.16 (24)	+.26
Mean Difference	+.65	−.08	+.06	+.34	

> **Alas, the reverse was also true—when students were asked to do less they did less. Even so, at every level of assignment difficulty seniors did outperform students from all other classes.**

Thinking the mean was .72 higher, and for Written Communication .34 higher.

Alas, the reverse was also true—when students were asked to do less they did less. Even so, at every level of assignment difficulty seniors did outperform students from all other classes, with the exception of junior critical thinking scores at the Introduce and Practice level of assignment difficulty. This overall result **is at least consistent with** the idea that students' critical thinking and written communication skills do improve while in college.

Other numbers from the table are a bit puzzling and should be noted even as we recognize the critical role of assignment difficulty overall. For example, increasing assignment difficulty is not uniformly and linearly related to rubric scores. Freshman scores on critical thinking are lower when the assignment is "Capstone" but higher when the assignment is "Capstone" for Written Communication, though the N is small.

Concluding Thoughts

As with any preliminary investigation into a complex, nuanced topic, our work raises many more questions—philosophically, pedagogically, epistemologically, ethically—than it answers:

- *Are there developmental differences associated with individual outcomes that should be considered? For example, could it be that students are less ready to do critical thinking at a high level as freshmen than they are ready to do writing, so a "Mastery" level assignment is too challenging for them?*
- *What would the learning look like if students encountered at least one "Mastery" level assignment each year of their undergraduate careers? In other words, does—as we surely hope—students' ability to perform at the highest level demanded increase with time in college even as faculty continue to try to tune assignment difficulty generally to move students to higher levels of competence?*

- *Do students leave college with greater proficiencies on a whole series of higher-order learning goals than when they entered? For example, in the cases of Critical Thinking and Written Communication we looked at here, do individual students, on average, make an argument from evidence better or write better when they leave college than they did when they first enrolled?*
- *What are the equity implications of this work? In other words, do all students, regardless of gender, race, ethnicity, Pell eligibility, first generation, and/or transfer status not only have access to but regularly engage in courses that foster these Essential Learning Outcomes? At what level?*

For us to be able to answer these as well as many other questions, we will need to build on the exploration initiated with this paper. To get at whether students' higher-order skills improve while in college we may have to find a way to have cohorts of students in each of their four class years respond to a "Mastery-level" assignment. Ultimately, we need to follow cohorts of students from their first year through their senior year at the unit record level, and that is where we hope to take the VALUE project next.

We see in the assignments that we have collected that faculty are addressing important learning outcomes for their students that go beyond gaining disciplinary knowledge, preparing students for success across foundational skills and abilities regardless of area of study. Based on our conversations with participating faculty, we are moving away from the language of assignment difficulty to that of an assignment's **pedagogical purpose**.

This shift is more than mere semantics as it reflects the true nature of the VALUE approach. That approach

> **What are the equity implications of this work? In other words, do all students, regardless of gender, race, ethnicity, Pell eligibility, first generation, and/or transfer status not only have access to but regularly engage in courses that foster these Essential Learning Outcomes?**

recognizes that all levels of the four-point VALUE rubrics and the 8-point scale of assignment purpose are legitimate, appropriate, and necessary in the developmental cultivation of students' higher-order knowledge, skills, and abilities. The results we have presented reinforce for us that student learning needs to be iterative and intentionally scaffolded across the curriculum to gain higher-order practice and competence, and that essential learning occurs in multiple courses and programs, not exclusively in a particular class.

Students need to see the connections between and among not just the classes they are taking, but between and among the actual assignments they complete. This will not happen by magic, however. Faculty must not just design assignments, courses, and curricula to foster higher-order learning in their students, but also be transparent and communicate the underlying purpose and intent of each to the students themselves, early and often. The good news is that in addition to VALUE, there are complementary approaches and tools available to support faculty work in this area, including the assignment design charrettes provided by the National Institute of Student Learning Outcomes Assessment (see http://www.learningoutcomeassessment.org/assignmenttool-kit.html) and the Transparency in Learning and Teaching in Higher Education project toolkit (see https://www.unlv.edu/provost/teachingandlearning).

We at AAC&U believe we are on the cusp of something truly important—a transforming and affordable strategy for increasing the level of multiple kinds of higher-order learning and a reliable, valid way to assess students' levels of learning that feeds back into faculty practice. These emerging patterns of relationships are based on direct assessment of student work that provides faculty—the individuals who then can actually affect changes in instruction and expectations for performance—information they need to modify and improve practice. This is the unique value of the VALUE approach and a potential game-changer. C

RESOURCES

■ Arum, R. & Roksa, J. (2011). *Academically Adrift: Limited Learning on College Campuses.* Chicago, IL: University of Chicago Press.

■ Condon, W., Iverson, E.R., Manduca, C., Rutz, C., & Willett, G. (2016). *Faculty Development and Student Learning: Assessing the Connections.* Bloomington, IN: Indiana University Press.

■ Finley, A. & McNair, T. (2013). *Assessing Underserved Students' Engagement in High-Impact Practices.* Washington, DC: Association of American Colleges and Universities.

■ Hutchings, P. (2016). Aligning educational outcomes and practices. *NILOA Occasional Paper #26. Retrieved from http://learningoutcomesassessment.org/documents/Occasional%20Paper%2026.pdf*

■ Kuh, G.D. (2008). *High-Impact Educational Practices: What They Are, Who Has Access to Them, and Why They Matter.* Washington, DC: Association of American Colleges and Universities.

■ Kuh, G.D., O'Donnell, K., & Reed, S. (2013). *Ensuring Quality and Taking High-Impact Practices to Scale.* Washington, DC: Association of American Colleges and Universities.

■ Kuh, G.D., O'Donnell, K., & Schneider (2017). HIPs at Ten. *Change: The Magazine of Higher Learning, 49*(5), 8–16.

■ Sullivan, D.F. & McConnell, K.D. (2017). Big Progress in Authentic Assessment, But By Itself Not Enough. *Change: The Magazine of Higher Learning, 49*(1), 14–24.