How Disciplinary Differences Shape Student Learning Outcome Assessment: A Case Study

Su Swarat
Pamella H. Oliver
Lisa Tran
J. G. Childers
Binod Tiwari
Jyenny Lee Babcock

California State University, Fullerton

Assessment of student learning outcomes (SLOs) has become increasingly important in higher education. Meaningful assessment (i.e., assessment that leads to the improvement of student learning) is impossible without faculty engagement. We argue that one way to elicit genuine faculty engagement is to embrace the disciplinary differences when implementing a university-wide SLO assessment process so that the process reflects discipline-specific cultures and practices. Framed with Biglan's discipline classification framework, we adopt a case-study approach to examine the SLO assessment practices in four undergraduate academic programs: physics, history, civil engineering, and child and adolescent studies. We demonstrate that one key factor for these programs' success in developing and implementing SLO assessment under a uniform framework of university assessment is their adaptation of the university process to embrace the unique disciplinary differences.

Keywords: *assessment, disciplinary differences, faculty engagement*

Tʜᴇ landscape of learning and teaching in higher education has shifted from the instruction paradigm to the learning paradigm (Barr & Tagg, 1995). Under the new learning paradigm, the focus is no longer the number of course offerings or pages of materials the faculty delivers but the amount of knowledge the students are able to actively construct for themselves (Halpern & Hakel, 2003). Correspondingly, the criterion for success for education changed from the quality of faculty instruction to the quality of student learning.

Accompanying this paradigm shift is the change of focus for institutional accountability. That is, instead of measuring "inputs," such as curricula, services, and resources the university provides, there is an increasing demand for evidence of "outputs," or the *impact* on student learning—namely, whether and how well students develop the knowledge, skills, and attitudes expected of them (Allen, 2004; Banta & Associates, 2002). Accreditation organizations are placing more emphasis on the centrality of evidence of student learning, or in other words, the results of student learning outcome (SLO) assessment. For example, the WASC Senior College and University Commission (2013), under which California State University, Fullerton (CSUF), is accredited, included the following "criteria for review" in its 2013 handbook: "The institution demonstrates that its graduates consistently achieve its stated learning outcomes and established

standards of performance. The institution ensures that its expectations for student learning are embedded in the standards that faculty use to evaluate student work" (p. 15). Institutions are not only encouraged but now required to establish and assess SLOs as a way of evaluating whether our higher education system is fulfilling its promise to the students and the public.

Faculty Engagement in Assessment

Few argue against the ultimate purpose of assessment—the systematic collection, review, and use of information about educational programs undertaken for the purposes of improving student learning and development (Palomba & Banta, 1999). However, creating institutionwide buy-in of SLO assessment often is not an easy task (Bresciani, 2006). Many faculty, who are at the front and center of student learning, have been skeptical of the assessment movement. At many institutions, accreditation remains the primary driver of assessment, which has caused faculty to view assessment as a means to fulfill compliance requirements and not a genuine means to examine and improve student learning (Kuh, Jankowski, Ikenberry, & Kinzie, 2014). Hutchings (2010) cited the following reasons for faculty's resistance to assessment: unfamiliarity or discomfort with

the language of assessment, lack of expertise in conducting assessment, insufficient institutional reward system for assessment, and lack of sufficient evidence that assessment makes a difference. The National Institute for Learning Outcomes Assessment (NILOA) compiled a list of case studies that showcased institutions with successful practices in using assessment information to improve teaching and learning practices, but as Gilbert's (2015) commentary in the *Chronicle of Higher Education* made clear, such success obviously is not widely spread or well known enough to convince faculty of the benefit of assessment (Baker, Jankowski, Provezis, & Kinzie, 2012).

Yet meaningful assessment (i.e., to result in improved student learning) is impossible without faculty engagement, and the institutions know it. The NILOA 2013 survey of over 1,000 U.S. institutions on the state of SLO assessment showed that "provosts rated faculty ownership and involvement as top priorities to advance the assessment agenda" (Kuh et al., 2014, p. 4). We argue in this article that one way to increase faculty ownership and involvement is to respect the disciplinary differences and perspectives of faculty that frame how they perceive and approach assessment.

The Importance of the Discipline in Assessment

The importance of the *discipline* is summarized eloquently by Kuh and colleagues (2015): "Disciplinarity must be central to assessment, allowing faculty to operate where they are most comfortable and to bring their field's distinctive questions, methods, and ways of thinking to the task of improving their students' learning (Hutchings, 2011)" (p. 103). The same principle is echoed by the Degree Qualification Profile (DQP) project by the Lumina Foundation, which highlighted *specialized knowledge* (in a disciplinary specialization) as one of the key categories to organize SLOs. DQP advocates for institutions to convene "faculty within a discipline who, with input from employers, establish discipline-specific curricular reference points and learning outcomes" (Lumina Foundation, 2014, p. 6), a process that could lead to a broad consensus among faculty on the meaning and quality of the degree (Ewell, 2012).

Several efforts exist to capture disciplinary differences related to assessment. The recent Measuring College Learning project brought together six discipline-specific faculty panels to articulate SLOs and identify disciplinary principles for learning assessment (Arum, Roksa, & Cook, 2016). The essays by scholars from 10 disciplines on the scholarship of teaching and learning demonstrated how the "disciplinary styles empower the scholarship of teaching, not only by giving scholars a ready-made way to imagine and present their work but also by giving shape to the problems they choose and the methods of inquiry they use" (Huber & Morreale, 2002, p. 32). The Teagle Foundation's

edited volume on assessment in the literary studies (Heiland & Rosenthal, 2011) explicitly pointed out the need for discipline-based assessment, presenting conceptual and practical arguments for the peril of alienating faculty by imposing generic assessment practices that do not reflect the subject matter itself. Literary fields aside, specific examples of assessment in other disciplines have also been documented. The Association for Institutional Research has produced the Assessment in the Disciplines series, presenting best practices in business (Martell & Calderon, 2005), engineering (Kelly, 2008), mathematics (Madison, 2006), and writing (Paretti & Powell, 2009). *Assessing Student Learning in the Disciplines* also explored different assessment practices in various disciplines (Banta, 2007). Wright (2007), in the book's opening chapter, criticized the current assessment frameworks as generic and not speaking to the disciplinary interest of faculty. In addition to advocacy by scholars and specific examples, NILOA also collected self-reported data on assessment practices at the program level (Ewell, Paulson, & Kinzie, 2011). The responses from approximately 900 randomly selected department or program heads confirmed that disciplines vary greatly in their motivation for assessment, the type of methods they use, and how much they use the results of assessment.

With increasingly heavy quality assurance requirements from accreditation organizations, it is understandable that a uniform assessment framework at the institutional level is often desired. We argue that an institutional assessment framework does not necessarily devoid the disciplines of their culture, discourse, history, and habit of mind. An institutional SLO assessment framework can be successful if it allows faculty to express their unique disciplinary tradition and to adapt the framework to the specific context of their subject field. We support our argument using a case-study approach presenting data to demonstrate and compare how four undergraduate academic programs—physics, history, civil engineering, and child and adolescent studies (CAS)—approach SLO assessment under a uniform framework for university assessment.

Institutional Context

CSUF is a large regional public university, enrolling over 39,000 students (as of fall 2016) and employing approximately 1,800 full- and part-time faculty members. The university offers 110 degree programs in eight colleges. Assessment is not new to the university, with several "pockets" on campus (e.g., measurement-oriented departments, disciplinarily accredited programs) having a long history assessing student learning. However, a universitywide assessment process was not implemented until spring 2014, following the inclusion of assessment in the university's 2013-to-2018 strategic plan. This elevated status of assessment is certainly a response to the institution's accreditation

requirements but at the same time signifies CSUF's serious commitment to student learning. Specifically, the university established a six-step assessment process (Figure 1) to guide assessment efforts at the program level. The university assessment reporting process was also redesigned to align with the six-step process.

The universitywide assessment process helps the institution coordinate and report assessment activities and results; but at the program level, its implementation varies greatly. The university recommends each undergraduate program to have five to seven SLOs to be assessed over a 5-year period, which coincides with the current university strategic plan. The programs are required to assess (i.e., collect data and determine improvement actions) at least one SLO per year and report the findings annually. The university process is overseen by the Office of Assessment and Educational Effectiveness (OAEE), which also provides training and support on how to conduct assessment following the six-step process. For example, workshops have been offered to discuss topics such as how to develop meaningful SLOs, how to connect assessment to curriculum through curriculum mapping, how to choose embedded methods, and how to integrate rubrics into SLO assessment. OAEE works very closely with the group of faculty assessment liaisons (FALs), who are nominated by each college and lead the program-level assessment effort in their respective college. The FALs keep the university informed of assessment activities and needs at the program level and provide tailored support to faculty. They also review annual assessment reports to provide feedback to each program on how to improve assessment practices.

A Theoretical Lens for Disciplinary Differences

As diverse as the disciplines, disciplinary differences can be examined in many different ways. Although much research has taken place to examine teaching and learning in different subject areas at the K–12 level (see, for example, the seminal work by Shulman, 1986, on pedagogical content knowledge), similar studies at the university level have been rather generic. Research that goes beyond teaching and learning to look at disciplinary context, value, identity, or culture is rare. Existing frameworks on this issue have largely examined disciplinary differences based on cognitive (e.g., Biglan, 1973) or social (e.g., Becher, 1989) variations, but the most predominant disciplinary classification schemes tend to be cognitively oriented. Specifically, the classification scheme developed by Biglan (1973), which empirically identified "hard/soft" and "pure/applied" as two dimensions to differentiate the disciplines, has received much empirical validation (see Alise, 2008, for a review).

According to Biglan (1973), the hard/soft dimension is concerned with "the degree to which a paradigm exists" (p. 202). Hard disciplines (e.g., physics) have a greater
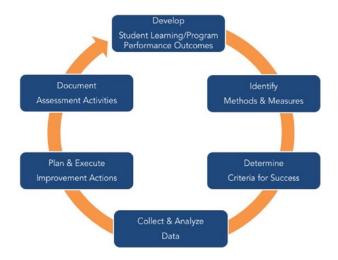

FIGURE 1. *The universitywide six-step assessment process.*

consensus about theories, principles, and methods, whereas soft disciplines (e.g., philosophy) lack such a converged view of the paradigm. The pure/applied dimension refers to "the degree of concern with application" (Biglan, 1973, p. 202), with applied disciplines more concerned about practical applications than their pure counterparts. This classification, particularly the pure/applied dimension, is reinforced in Stokes' (1997) conceptualization of Pasteur's quadrant, which suggests that research can be categorized based on two spectrums: "quest for fundamental understanding" and "considerations of use." Both Biglan's classification model and Stokes' framing of Pasteur's quadrant indicate that disciplines within the same category tend to share common values, culture, and norms of practice, whereas disciplines belonging to different categories have little in common. As such, Becher (1989) referred to the disciplines as "academic tribes" and argued that the discipline-based values, culture, and norms are what form the boundaries between and set the identities of these tribes.

We utilize Biglan's (1973) classification model as the theoretical foundation to examine disciplinary differences, as we feel that this classic theory helps us address the diversity of our campus. Neumann and colleagues (Neumann, 2001; Neumann, Parry, & Becher, 2002) further examined the interaction between Biglan's two dimensions and classified the disciplines into four categories: "hard pure," "soft pure," "hard applied," and "soft applied." They described hard pure disciplines (e.g., physics) to be "typified as having a cumulative, atomistic structure, concerned with universals, simplification and a quantitative emphasis" and soft pure disciplines (e.g., history) to be "reiterative, holistic, concerned with particulars and having a qualitative bias" (Neumann et al., 2002, p. 406). Hard applied disciplines (e.g., engineering) are characterized as "concerned with mastery of the physical environment and geared towards products and techniques," whereas soft applied disciplines

TABLE 1

Typical Assessment Approaches Used by Disciplines Based on Biglan (1973) and Neumann (2001; Neumann, Parry, & Becher, 2002)
Classification Models

	Hard	Soft
Pure	Prefer objective, focused exam questions that reveal specific knowledge acquisition	Prefer essays or papers that reveal the learners' grasp of the complexity and totality of the subject matter
Applied	Prefer exam questions that cover both factual knowledge and problem-solving skills	Prefer project-based assessment, peer assessment, and even self-reflection

(e.g., education) care about "the enhancement of professional practice and aiming to yield protocols and procedures" (Neumann et al., 2002, p. 406). Using this framework, Neumann and colleagues (Neumann, 2001; Neumann et al., 2002) have suggested differences in various aspects of teaching and learning across the disciplines. For example, hard pure disciplines tend to have a linear and hierarchical curriculum, whereas the curriculum of soft pure disciplines is more "spiral" in nature. The curricula of hard applied and soft applied disciplines have the same contrast, but they both tend to place less emphasis on "validating knowledge through examining conflicting evidence and exploring alternative explanations" (Neumann et al., 2002, p. 408). In terms of approaches to teaching, large lectures seem to be more prevalent among the hard pure disciplines, whereas small-group settings that encourage interpersonal interactions are more common among the soft pure ones. These differences also applied to the hard applied and soft applied comparison, although both focus more strongly on the immersion of practical experiences and engagement of experienced practitioners. When it comes to assessing student learning (see Table 1 for a summary), hard pure disciplines seem to prefer objective, focused exam questions that reveal specific knowledge acquisition, whereas soft pure disciplines favor essays or papers that demonstrate the learners' grasp of the complexity and totality of the subject matter. Similar to hard pure disciplines, hard applied ones prefer exam questions that cover both factual knowledge and problem-solving skills. This compares with soft applied fields that rely more on project-based assessment, peer assessment, and even self-reflection.

In this article, we borrow the disciplinary classification framework by Biglan (1973) and Neumann (2001) as a lens to examine how different disciplines on our campus approach assessment, employing a case-based approach to demonstrate how the suggested disciplinary differences manifest themselves in the process of SLO assessment.

Method

The journeys of SLO assessment in four undergraduate programs are compared as case studies in this article: physics (the hard pure discipline), history (the soft pure discipline), civil engineering (the hard applied discipline), and CAS (the soft applied discipline). Yin's (2014) definition of *case study*—"an empirical inquiry that investigates a contemporary phenomenon (the "case") in depth and within its real-world context, especially when the boundaries between phenomenon and context may not be clearly evident" (p. 16)—perfectly suits the purpose of our inquiry in that we are exploring how individual disciplines carry out SLO assessment within the boundary of the discipline itself and the shared context of the university. The unit of analysis or the focus of the study is the individual program (Miles & Huberman, 1994). With these cases, the questions we attempt to address are *descriptive* in nature—what is the SLO assessment process in the discipline, and how does the process reflect the unique disciplinary identity? As stated earlier, the four cases or undergraduate programs used for this study are chosen because they represent the different categories of disciplines based on the frameworks of Biglan (1973) and Neumann (2001; Neumann et al., 2002), and they have successfully created viable SLO assessment programs within the framework of the universitywide assessment process. In other words, we believe that these programs have developed the appropriate disciplinary "twist" to the university process that allows them to implement assessment that reflects their own identity and culture.

The authors of this article are "participant-observers" of the SLO assessment process. Two authors are members of the university's OAEE, which oversees the SLO assessment efforts at the program level and provides support to faculty from diverse disciplines. Four authors serve as FALs for their respective colleges, to which the four cases of study belong. The FALs lead and/or participate in the program-level SLO assessment and, at the same time, monitor the program progress and connect the program to the university. In other words, the authors serve "dual purposes" in the program-level SLO assessment process: to engage in the assessment activities and to observe the activities. By doing so, the authors have gained both the "insider" and the "outsider" perspective (Spradley, 1980).

Given the roles of the authors, our case study engaged with multiple sources of data, with strong reliance on the direct participant observations of the assessment process and relevant artifacts (e.g., e-mails, documents, reports), which are common sources of evidence for case studies (Yin, 2014). The four FALs produced a reflection on how the SLO
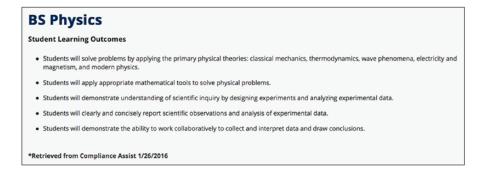
BS Physics

Student Learning Outcomes

- Students will solve problems by applying the primary physical theories: classical mechanics, thermodynamics, wave phenomena, electricity and magnetism, and modern physics.
- Students will apply appropriate mathematical tools to solve physical problems.
- Students will demonstrate understanding of scientific inquiry by designing experiments and analyzing experimental data.
- Students will clearly and concisely report scientific observations and analysis of experimental data.
- Students will demonstrate the ability to work collaboratively to collect and interpret data and draw conclusions.

*Retrieved from Compliance Assist 1/26/2016

FIGURE 2. *Physics BS student learning outcomes (as of January 2016).*

assessment process is developed and implemented in their respective disciplines. Specifically, the reflection was focused on how faculty in their respective programs perceive assessment and how the faculty approach each step of the assessment process (e.g., identifying outcomes, choosing methods, determining improvement actions). The reflections were shared among the authors, and conversations also took place to discuss the disciplinary differences. The discussions followed a mind-set of "explanation building" to explore how the different observations reflect individual disciplinary identities (Yin, 2014). The results presented below are revised reflections based on these discussions. It is our intention to tell stories from different programs of how the SLO assessment process took place and how the disciplinary culture, value, and norms shaped the process. To ensure validity, the reflection of each program was shared with respective faculty, who provided feedback and confirmed the accuracy of these case studies.

Results

We describe below how the university assessment process manifests itself in the four case disciplines: physics, history, civil engineering, and CAS. A parallel structure is followed in the descriptions so that a comparison can be easily drawn. Specifically, we discuss how faculty perceive assessment, how the SLOs are developed, how assessment measures are chosen with a specific emphasis on the use of rubrics, and how assessment results are used to improve teaching and learning practices. The use of rubrics is highlighted because rubrics have been introduced on campus as an effective way to encourage authentic assessment of student learning and are used by many programs.

Case 1: Physics (Hard Pure Discipline)

According to Biglan (1973), physics as a field shares a strong consensus on a cognitive paradigm. That is, there is a high degree of agreement among faculty in the fundamental concepts, principles, and skills students need to develop. As expected, our physics faculty describe the curriculum as

"highly structured" and "hierarchical," with latter concepts built upon previously acquired knowledge. As they complete the curriculum-mapping exercise (i.e., mapping the courses to the program-level SLOs) required of all departments, it becomes clear that the sequential nature of courses led the course-level SLOs to largely mirror the program-level SLOs. This makes the process of SLO development fairly easy and straightforward. A small committee of faculty crafted the set of SLOs for the physics BS program and distributed it to all (full-time and senior part-time) faculty by e-mail to solicit feedback. Only a couple of minor edits were received, and at a subsequent department meeting, the faculty reached agreement and adopted these SLOs without much difficulty. Among the five SLOs (Figure 2), the first one clearly reflects the shared "cognitive paradigm" that typifies a hard pure discipline: "Students will solve problems by applying the primary physical theories: classical mechanics, thermodynamics, wave phenomena, electricity and magnetism, and modern physics."

The close alignment between the course SLOs and the program SLOs created a source of reluctance toward assessment. Most faculty believe that the course grades—which are largely based on objective, content-focused exams—are an accurate assessment of student learning; hence the separate assessment of program SLOs seems unnecessary. To help engage faculty in the process, the assessment coordinator took the approach of presenting program-level assessment as an additional opportunity to examine student learning. For example, faculty on the assessment committee were excited by the opportunity to view oral presentations in the senior laboratory course.

Assessment measures for physics faculty are primarily embedded measures, that is, existing assignments or exam questions in the appropriate courses. These measures are primarily quantitative in nature and are designed to objectively measure student learning. For example, to assess students' ability to "apply appropriate mathematical tools to solve physical problems," students are administered the Colorado Upper-Division Electrostatics Test at the start and end of a junior-level course as well as the Colorado Upper-Division Electrodynamics Test in their senior-level course. The
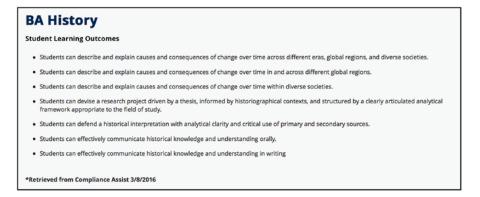
BA History

Student Learning Outcomes

- Students can describe and explain causes and consequences of change over time across different eras, global regions, and diverse societies.
- Students can describe and explain causes and consequences of change over time in and across different global regions.
- Students can describe and explain causes and consequences of change over time within diverse societies.
- Students can devise a research project driven by a thesis, informed by historiographical contexts, and structured by a clearly articulated analytical framework appropriate to the field of study.
- Students can defend a historical interpretation with analytical clarity and critical use of primary and secondary sources.
- Students can effectively communicate historical knowledge and understanding orally.
- Students can effectively communicate historical knowledge and understanding in writing

Retrieved from Compliance Assist 3/8/2016

FIGURE 3.   *History BA student learning outcomes (as of March 2016).*

grading process is typically done individually by the course instructor, without explicitly enumerating the scoring criteria used. Although the university encourages utilizing indirect measures (e.g., student reflection, self-report) as a supplement to direct measures, most of the physics faculty question the value of these results, even as a complement to direct assessment methods, and are reluctant to use them in their interpretation of the assessment results.

Rubrics, a scoring tool that naturally calls for standard-sharing among faculty, are rarely used. Even when they are used (i.e., in the scoring of oral presentations), there is no deliberate attempt to calibrate the rubrics beforehand. This is perhaps due to the belief that since the assessment measures are objective (i.e., not interpretive), the chance that two graders will score them differently is small. Physics faculty tend to prefer holistic rubrics (as opposed to analytic ones) that allow them to draw a general judgment of student mastery of the SLOs, perhaps due to the belief that "you either got it or not." This is clearly reflected in the first assessment of the program SLOs. The assessment committee created a rubric that simply listed each of the five SLOs as a separate dimension and scored oral presentations on each SLO using one of four levels of quality. In the subsequent year, with the encouragement of the assessment coordinator, the faculty readily agreed to assess only one SLO using an analytic rubric. Following the 1st year, in the instances where rubrics are used, faculty have chosen preformed rubrics modified slightly to reflect the course assignment and SLO being assessed and were able to demonstrate generally consistent agreement across the graders. Although many faculty felt that the holistic rubric was sufficient, some began to notice the benefit of the additional details that the analytic rubric provides. More faculty are now reporting that they enjoy the process of using rubrics to assess and discuss student learning, which perhaps signals a good opportunity to gradually introduce the process of calibration.

The process of sharing assessment results is similar to that of developing the SLOs. In the 2014–2015 assessment cycle, the results were shared with the faculty, and consensus was relatively easily achieved that most students met the program expectations. Although expectations were met, the faculty noted that use of technical language was consistently weaker than other criteria. In particular, the assessment committee noted that on several occasions, students used scientific terms imprecisely. As a result, efforts were made in upper-division courses to improve this aspect of the students' communication.

Case 2: History (Soft Pure Discipline)

History is an interdisciplinary field that reflects diverse content and employs a variety of methodological and epistemological approaches. As a result, history courses reflect a wide range of topics and strategies, and even sections taught by faculty with various specializations can be quite different. As such, developing a set of SLOs that governs the entire degree program is challenging, making faculty involvement in every stage of the process critical. Initially, a committee drafted the SLOs, which were discussed and revised at faculty retreats and department meetings.

Given the difficulty of reaching a consensus on *content*, history focuses on the *skills* that are central to all history courses. Faculty agreed that history majors should be able to do the following: develop historical thinking skills that emphasize causation, comparison, and contextualization; design research projects based on primary sources and informed by scholarship; and write well-argued papers substantiated by the use of relevant historical evidence. As a result, these skills are reflected in the seven SLOs developed by the history faculty (Figure 3).

The level of faculty engagement and reflection is high, as those who teach the courses on the curriculum map consider if, when, and how they teach a particular SLO in their course. Once a consensus is reached on the mapping of the SLOs to the curriculum, faculty are asked to identify a particular assignment that addresses the assigned SLO; this information is used by the committee to plan the assessment of each SLO.

In contrast to their physics counterparts, history faculty prefer the use of written assignments, which can range from short responses to research projects. The nature and content of these assignments vary greatly, reflecting the instructors' areas of specialty and pedagogical preferences. Consistency in the type of assignment used for assessment was easiest to achieve when assessing SLOs in the senior capstone seminar, in which students produce a final research paper. However, in the assessment of other courses that do not share a common structure, particularly those with multiple sections, instructors chose the assignment they deemed appropriate as long as it addressed the SLO. A range of assignments was used, including essay and short-answer questions on an exam, graded in-class assignments, graded take-home assignments, or papers completed for extra credit.

The diverse nature of the assessment measures and different instructor approaches make it critical that a common rubric is used to ensure that student work is measured against the same criteria. Rubric calibration is an important and intentional part of assessment in history. The rubric calibration typically involves scoring samples of student work using a draft rubric, and scoring consistency is achieved through discussing disparities and revising the rubric as needed. Rubric calibration is conducted at two stages in the assessment process: during the pilot and during the full assessment. During the pilot, a committee calibrates the rubric using sample student work. During the full assessment, new faculty who will be scoring student work participate in a rubric calibration session to ensure interrater reliability, which is determined by the consistency with which faculty reach the same (or nearly same) score on the same sample. Variations in scoring engender a discussion among the faculty, who each explain the rationale for their score. Generally, discrepancies arise from different interpretations of specific words in the rubric or different standards for meeting the criteria for each score. If a consensus still cannot be reached after faculty have discussed their rationales, the rubric is revised based on faculty feedback; the revised rubric is used in subsequent rounds during the calibration session. Interrater reliability was easier to achieve with SLOs that focused on research skills (such as crafting an argumentative thesis statement) than those that required content knowledge (such as explaining causes and consequences). In the latter, discrepancies in scores stemmed from variations in faculty familiarity with the specific topic of the student paper. This underscores the importance of engaging faculty in assessment; faculty expertise is crucial to ensuring that assessment results accurately reflect student learning. In summary, the interrater reliability is accomplished through consensus, and no quantitative measures are used.

History faculty are quite reflective in their interpretation of the assessment results. Past improvement actions have generally fallen into three categories: providing assistance to students (e.g., peer tutoring, writing workshop), making curricular adjustments (e.g., adding prerequisites, creating new courses), and developing pedagogical tools for instructors (e.g., effective pedagogy workshops). History faculty also reflect upon the assessment process itself. For instance, they have noticed that sometimes low student performance is a reflection of the nature of the assignment and not necessarily student learning. To this end, efforts will be made to improve the consistency of the types of assignments used for assessment.

Case 3: Civil Engineering (Hard Applied Discipline)

As in many other institutions, the civil engineering BS program at CSUF is accredited by the Engineering Accreditation Commission of ABET, which is the Accreditation Board for Engineering and Technology. Instead of developing its own SLOs, the civil engineering program adopted the 11 ABET SLOs (Figure 4) and identified their alignment with the curriculum, the program educational objectives (also required by ABET), and the university learning goals. In determining the SLOs, the program not only involved its faculty and students but also sought extensive feedback from its alumni and the Industrial Advisory Board. This practice reflects the discipline's emphasis on practitioners' experiences and perspectives. While establishing the SLOs, there was a series of discussions on whether fewer SLOs than the 11 ABET SLOs should be established or if the ABET SLOs should be adopted. After extensive discussions in multiple faculty meetings, the program came to consensus to adopt the ABET SLOs.

Civil engineering uses both direct and indirect measures to assess the SLOs. For direct measures, faculty prefer embedded questions (e.g., final exam questions). These questions are jointly developed by the faculty members who are in charge of the course, and a common rubric is used for grading. Similar to physics colleagues, civil engineering faculty tend to view the questions as objective and have found the application of the grading rubric to be straightforward. As such, no rubric calibration process has been needed. For indirect measures, the program administers self-developed surveys that address specific SLOs to current students and alumni alike, again a reflection of the program's "applied" focus. Initially, there was no consensus among the faculty members on adopting indirect measures. Many faculty argued that as the indirect measures are students' self-assessment that may not accurately reflect student learning, indirect assessment may not add value to the assessment process. However, after a series of discussions among faculty, an agreement was reached for using indirect measures, as such measures include the student voice and could therefore be used to improve course content.

A standard "criterion for success" is applied to all SLOs: Students on average need to score at least 70% on the assessment measure (e.g., exam question) to be considered
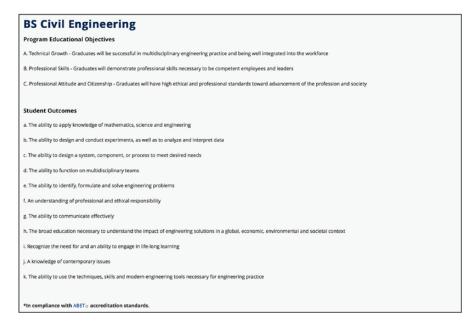
BS Civil Engineering

Program Educational Objectives

A. Technical Growth - Graduates will be successful in multidisciplinary engineering practice and being well integrated into the workforce

B. Professional Skills - Graduates will demonstrate professional skills necessary to be competent employees and leaders

C. Professional Attitude and Citizenship - Graduates will have high ethical and professional standards toward advancement of the profession and society

Student Outcomes

a. The ability to apply knowledge of mathematics, science and engineering

b. The ability to design and conduct experiments, as well as to analyze and interpret data

c. The ability to design a system, component, or process to meet desired needs

d. The ability to function on multidisciplinary teams

e. The ability to identify, formulate and solve engineering problems

f. An understanding of professional and ethical responsibility

g. The ability to communicate effectively

h. The broad education necessary to understand the impact of engineering solutions in a global, economic, environmental and societal context

i. Recognize the need for and an ability to engage in life-long learning

j. A knowledge of contemporary issues

k. The ability to use the techniques, skills and modern engineering tools necessary for engineering practice

*In compliance with ABET® accreditation standards.

FIGURE 4. *Civil engineering BS student learning outcomes (as of March 2016).*

satisfactory. Since the curriculum is hierarchical in nature, civil engineering also utilizes a "value-added" approach to tracking student progress on the SLOs at three different points of the curriculum: initiate (freshmen-/sophomore-level courses), enhance (junior-level courses), and reinforce (senior-level courses). This additional layer of assessment provides information that helps explain observed student performance near or at graduation.

To manage the 11 SLOs, civil engineering has established an assessment plan that assesses each SLO every 2 years. In the event that assessment findings suggest a particular SLO is not met, a clear improvement plan is established and implemented by the faculty, and the SLO is reassessed in the following year (instead of in 2 years). Initially, there was no consensus on how many SLOs to assess per year. A few faculty members proposed to assess all SLOs within the 1st year and reassess every 2 years thereafter. However, after an extensive discussion, the program faculty came to consensus that assessing only a few SLOs a semester provides opportunities to continuously improve the assessment process and also reduces the likelihood of overwhelming the faculty and staff during the assessment year.

The assessment of civil engineering is led by an assessment committee consisting of three to five faculty members. The committee reviews the assessment measure of choice (e.g., exam question) and the grading rubric chosen by the faculty who teach the courses and provides feedback as necessary. The committee also compiles the assessment results for each SLO and reports the findings on an annual basis to the faculty and the program's Industrial Advisory Board. This way, all program constituencies, that is, students, alumni, faculty, and industrial partners, are fully involved in the SLO assessment process.

Case 4: CAS (Soft Applied Discipline)

CAS focuses on preparing students for professional, ethical, and reflective evidence-based practice with diverse populations of children, adolescents, and families. CAS has a long history of utilizing assessment for ongoing improvement of practice. This is because assessment is an integral part of the discipline, with measurement and assessment being parts of the curriculum. As the CAS assessment coordinator summarized, "The culture of our department embraces assessment." As such, the integration of the CAS program assessment with the university assessment plan presented no major difficulties, although of course there were many long discussions at every step before agreement was reached.

The CAS curriculum reflects multiple paradigms that make up the discipline (e.g., psychology, education, human development), which poses some challenges for reaching consensus on SLOs at the program level. At the beginning of the process, much time was spent determining the SLOs to represent what various faculty members thought were the crucial elements, dependent upon their own particular disciplines. However, the faculty were open to collaboration and worked together at faculty retreats to develop the program SLOs. The program has 10 SLOs (Figure 5), which, although quite a lot, are necessary to reflect the interdisciplinary nature of the program. The SLOs focus on both content and practice. They cover comprehension of theory, concepts, and research. They also include the application and integration of theories, research, communication skills, cultural

BS Child and Adolescent Development

Student Learning Outcomes

- Students can describe and/or explain relevant theories, concepts and related research findings.
- Students can identify and describe normative development and individual and group differences.
- Students can describe biological, psychological, cultural and environmental influences on development.
- Students can identify and describe key components of cultural competence.
- Students will be able to identify, access, analyze and synthesize relevant sources, including research studies.
- Students write effectively in APA style, taking purpose and audience into account.
- Students make effective oral presentations, taking purpose and audience into account.
- Students can effectively apply theories, concepts, and research findings to promote child well-being across diverse populations.
- Students can identify relevant ethical principles and legal issues and the impact of possible actions in real-world situations.
- Students can identify funding, services, and advocacy strategies at the local, state, federal and international levels that support children, adolescents, families and communities.

*Retrieved from Compliance Assist 3/1/2016

FIGURE 5.   *Child and Adolescent Development BS student learning outcomes (as of March 2016).*

competence, and professional and ethical evidence-based practice. In other words, the SLOs cover both factual knowledge and its application.

As practitioners in the field of child and adolescent development are required to use multiple methods to assess children, adolescents, and families, students in CAS are trained to use self-reports, observations, standardized tests, and a variety of other appropriate assessment methods. Consequently, faculty conduct assessment in the same manner, creating multiple measures, including senior papers, observations of oral presentations, multiple-choice tests, and short essays. They also actively utilize indirect measures, such as student surveys. The faculty approach assessment with the same rigor as their own research practice. For each SLO, they develop a pilot measure in the fall semester, which, upon improvement, becomes the final assessment tool in the spring. They also sample the student work through a systematic process, which results in an appropriate and representative random sample.

Given the types of assessments used, as well as the wide range of knowledge and skills assessed, CAS relies heavily on rubrics. Careful consideration was given to ensure the validity and reliability of the rubrics. CAS is the only program on campus that calculates interrater reliability and completes a double-scoring calibration process before every rubric is used and double-scores for the final assessment. This undoubtedly reflects the disciplinary culture and faculty background (e.g., psychology, statistics) of the program. Although the faculty agreed on the importance of interrater reliability, they also needed many conversations of the appropriate measure for reliability as well as the stringency that should be required for agreement. Some faculty wanted a level of agreement commensurate with research requirements, and others were amenable to a consensus approach. After many conversations, the department went with a stringent requirement for reliability, comparable to what would be acceptable in research.

The collaborative culture helps facilitate the use of assessment results as well. The results are reported to all faculty at faculty retreats. Faculty take an evidence-based approach to suggest changes needed based on the findings of the assessment in the areas of curriculum, pedagogy, or additional student or faculty support, as well as revisions of the assessment measure itself. The faculty's response to assessment results indicating that students had difficulty with synthesizing information from a variety of scholarly research articles demonstrates this "closing-the-loop" process. A faculty workshop was presented on techniques and in-class exercises that could be used to teach the skill of synthesizing. Particular consideration was made to reach part-time instructors via video recordings and dissemination of these workshops presented in response to a need for faculty professional development. CAS is fortunate the nature of the discipline draws faculty who are motivated to improve their teaching and student learning and who appreciate empirical evidence of success or the need to make changes.

Discussion

It is clear through the assessment "stories" of the four disciplines that the identity and culture of each discipline play a significant role in shaping how SLO assessment takes place. Although all faculty care about student learning, their enthusiasm toward assessment varies. Because assessment places more emphasis on practice and advocates for both subjective and objective measures, our soft applied discipline, CAS, seems to be most aligned with and hence most comfortable with it. It is understandable that the language and disciplinary origins of assessment itself may alienate some disciplines while uniting others (Cain, 2014). As a university promotes an institutionwide assessment process, it is important to be flexible and adaptive in how the process is introduced to different disciplines. For instance, since assessment follows largely a social science

research model, focusing on the parallel between assessment and research on student learning may be welcoming to social science disciplines. For humanities faculty, weaving assessment into a continuing dialogue about how students develop through the curriculum might be perceived as less threatening. For art disciplines, connecting assessment with existing performance examinations could offer a natural bridge to overcome faculty unfamiliarity with and resistance to assessment.

The preference for objective measures in our hard pure discipline, physics, also led the faculty to question the redundancy between SLO assessment and grading. They accepted assessment after SLO assessment was reframed as an additional and fine-tuned way of examining student learning. Yet the assessment measures remain heavily reliant on embedded, standardized assignments and exam questions, which are viewed as objective measures by the faculty. Faculty from the hard applied discipline, civil engineering, also shared the preference for embedded, objective measures. The perceived "objectivity" of these measures enables the faculty to easily apply a standard grading rubric with reasonable consistency of responses. The preference for objective measures contrasts greatly with the two soft disciplines' case studies: history and CAS. Both disciplines employed multiple measures. In the case of history, different assignments in courses with multiple sections were used to accommodate the diverse background of the faculty. In the case of CAS, direct and indirect measures that mirror real-life practitioners' standards were employed. Compared to the CAS faculty, the history and physics faculty both are more reluctant to use indirect measures, a shared attitude that could be attributable to the "pure" nature of the disciplines. Civil engineering, also engaged with indirect measures that capture how student learning is perceived by the practitioners, is a reflection of its "applied" disciplinary culture.

These four case studies represent only a small proportion of the wide range of assessment measures on our campus, but the comparison made it clear that it would be impossible and unwise to require the disciplines to follow prescribed assessment methods. For science, technology, engineering, and mathematics disciplines, because having objective measures is often desired, engaging faculty in a discussion of developing objective exam questions or class assignments that could be used across multiple sections (i.e., independent of instructors) may be an effective way to stimulate faculty interest in assessment. When introducing indirect measures to these disciplines, it would be critical to provide convincing examples of how student perceptions could complement the objective, direct measures. On the other hand, for disciplines where subjectivity (e.g., student or faculty individual perspectives) are of more value, prioritizing standardized tests across sections is likely to alienate faculty from assessment. Instead, offering measurement options that could be tailored to individual contexts would be more effective.

The disciplinary differences in assessment approaches are not limited only to measures but are also seen in SLO development. To make assessment manageable, our university advises each program to have five to seven SLOs for a 5-year assessment cycle (which coincides with the duration of our strategic plan). However, CAS chose to develop 10 SLOs and has a successful assessment schedule. The seemingly large number of SLOs is necessary to reflect and engage the diverse perspective the faculty bring, as they come from a range of social science fields. On the other hand, our undergraduate engineering programs are required by ABET to assess 11 SLOs. These discipline-specific issues require assessment professionals to be willing to "bend" the official guidelines to address the specific needs of a discipline and provide innovative solutions to align assessment with the faculty's practices. In our case, had the university required CAS or engineering to reduce the number of SLOs to follow the university guidelines, the faculty would likely feel alienated and thus disengage from the assessment process.

Similarly, the four disciplines approached rubrics differently. Although both history and CAS regularly use rubrics in scoring student work, the former calibrates the rubric through a qualitative dialogue approach, whereas the latter engages in a quantitative research approach. It is interesting that although research (Neumann et al., 2002) suggests that soft pure disciplines are less open to collaborative work due to the individualistic nature of the inquiry, our history faculty work together in developing and using rubrics as effectively as their CAS counterparts. The physics group, a hard pure discipline that has a readiness to work cooperatively (Neumann et al., 2002), did not seem to embrace the unpacking and calibration of rubric criteria. One explanation could be that because the measures used in physics are deemed "objective" and less open to interpretation, faculty do not think there is much room for deviation in using the rubrics. Faculty in the other hard discipline, civil engineering, shared similar thoughts regarding grading rubric calibration. These varied responses to rubrics highlight the importance for assessment professionals to consider the faculty's level of familiarity with and understanding of the vocabulary of assessment when introducing rubrics (or any other assessment concepts) to a discipline. Discussing rubrics without mentioning interrater reliability to measurement-savvy faculty would make assessment seem "lacking of rigor"; yet for a discipline that does not routinely use rubrics, diving into calibration without explaining the basics may cause assessment to be viewed as overly cumbersome.

Conclusion

Nearly three decades have passed since the official notion of SLO assessment, as well as the tie between assessment and accreditation, was introduced to higher

education (Ewell, 2002). However, the integration of assessment into an institution's regular teaching and learning practices has yet to take place on most campuses. Hutchings (2011) highlighted that one major reason for this slow change is the lack of assessment efforts situated in specific disciplines: "Assessment's focus on cross-cutting outcomes makes perfect sense, but it has also meant that the assessment of students' knowledge and abilities within particular fields, focused on what is distinctive to the field, has received less attention. And that's too bad" (p. 36). We echo Hutchings' concern and argue that disciplinary differences do not have to be ignored in order to accomplish uniform accountability goals.

Our case study demonstrated that four disciplines, although operating under the same universitywide assessment process, managed to adapt the process to their unique "flavor." Every step of the assessment process is enacted differently between the disciplines yet maintains the integrity of the process itself. As such, each discipline consistently developed a successful model of assessment that engages its faculty and, at the same time, allows the university to effectively and efficiently gather institutionwide information on student learning. Given that faculty are the most critical members of the assessment process, we hope that our case study presents a convincing argument for the importance of respecting disciplinary differences in implementing a universitywide assessment process (Driscoll & Wood, 2007). It is our belief that only by doing so can we generate valid, reliable, and meaningful data through assessment to enhance student learning across the disciplines.

References

Alise, M. A. (2008). *Disciplinary differences in preferred research methods: A comparison of groups in the Biglan classification scheme* (Doctoral dissertation). Retrieved from http://etd.lsu.edu/docs/available/etd-02222008-085519/

Allen, M. J. (2004). *Assessing academic programs in higher education*. San Francisco, CA: Jossey-Bass.

Arum, R., Roksa, J., & Cook, A. (2016). *Improving quality in American higher education*. San Francisco, CA: Wiley.

Baker, G., Jankowski, N., Provezis, S., & Kinzie, J. (2012). *Using assessment results: Promising practices of institutions that do it well*. Retrieved from http://learningoutcomesassessment.org/UsingAssessmentResults.htm

Banta, T. W. (2007). *Assessing student learning in the disciplines: Assessment update collections*. San Francisco, CA: Jossey-Bass.

Banta, T. W., & Associates. (2002). *Building a scholarship of assessment*. San Francisco, CA: Jossey-Bass.

Barr, R. B., & Tagg, J. (1995). From teaching to learning: A new paradigm for undergraduate education. *Change*, *27*(6): 12–25.

Becher, T. (1989). *Academic tribes and territories*. Buckingham, UK: Society for Research into Higher Education and Open University Press.

Biglan, A. (1973). The characteristics of subject matter in different academic areas. *Journal of Applied Psychology*, *57*(3), 195–203.

Bresciani, M. J. (2006). *Good practices in outcomes-based assessment program review*. Sterling, VA: Stylus.

Cain, T. R. (2014). *Assessment and academic freedom: In concert, not conflict* (NILOA Occasional Paper No. 22). Retrieved from http://learningoutcomesassessment.org/occasionalpapertwentytwo.html

Driscoll, A., & Wood, S. (2007). *Outcomes-based assessment for learner-centered education*. Sterling, VA: Stylus.

Ewell, P. (2002). An emerging scholarship: A brief history of assessment. In T. Banta (Ed.), *Building a scholarship of assessment* (pp. 3–25). San Francisco, CA: Jossey-Bass.

Ewell, P. (2012). *The Lumina Degree Qualifications Profile (DQP): Implication for assessment*. Champaign, IL: National Institute for Learning Outcome Assessment.

Ewell, P., Paulson, K., & Kinzie, J. (2011). *Down and in: Assessment practices at the program level*. Champaign, IL: National Institute for Learning Outcome Assessment.

Gilbert, E. (2015, August 14). Does assessment make colleges better? Who knows? *Chronicle of Higher Education*. Retrieved from http://www.chronicle.com/article/Does-Assessment-Make-Colleges/232371/

Halpern, D. F., & Hakel, M. D. (2003). Applying the science of learning to the university and beyond: Teaching for long-term retention and transfer. *Change*, *35*(4), 36–41.

Heiland, D., & Rosenthal, L. J. (2011). *Literary study, measurement, and the sublime: Disciplinary assessment*. New York, NY: Teagle Foundation.

Huber, M. T., & Morreale, S. P. (Eds.). (2002). *Disciplinary styles in the scholarship of teaching and learning: Exploring common ground*. Washington, DC: American Association for Higher Education and the Carnegie Foundation for the Advancement of Teaching.

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

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

Kelly, W. E. (2008). *Assessment in engineering programs: Evolving best practices*. Tallahassee, FL: Association for Institutional Research.

Kuh, G. D., Ikenberry, S. O., Jankowski, N. A., Cain, T. R., Ewell, P. T., Hutchings, P., & Kinzie, J. (2015). *Using evidence of student learning to improve higher education*. San Francisco, CA: Jossey-Bass.

Kuh, G. D., Jankowski, N., Ikenberry, S. O., & Kinzie, J. (2014). *Knowing what students know and can do: The current state of student learning outcomes assessment in US colleges and universities*. Urbana: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Lumina Foundation. (2014). *The Degree Qualifications Profile: A learning-centered framework for what college graduates should know and be able to do to earn the associate, bachelor's or master's degree*. Retrieved from http://degreeprofile.org/press_four/wp-content/uploads/2014/09/DQP-web-download.pdf

Madison, B. L. (2006). *Assessment of student learning in college mathematics: Towards improved programs and courses*. Tallahassee, FL: Association for Institutional Research.

Martell, K., & Calderon, T. (2005). *Assessment of student learning in business schools: Best practices each step of the way*. Tallahassee, FL: Association for Institutional Research.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis*. Thousand Oaks, CA: Sage.

Neumann, R. (2001). Disciplinary differences and university teaching. *Studies in Higher Education*, *26*(2), 135–146.

Neumann, R., Parry, S., & Becher, T. (2002). Teaching and learning in their disciplinary contexts: A conceptual analysis. *Studies in Higher Education*, *27*(4), 405–417.

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

Paretti, M. C., & Powell, K. M. (2009). *Assessment of writing*. Tallahassee, FL: Association for Institutional Research.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14.

Spradley, J. P. (1980). *Participant observation*. New York, NY: Holt, Rinehart and Winston.

Stokes, D. E. (1997). *Pasteur's quadrant: Basic science and technological innovation*. Washington, DC: Brookings Institution Press.

WASC Senior College and University Commission. (2013). *2013 handbook of accreditation*. Retrieved from https://www.wascsenior.org/resources/handbook-accreditation-2013.

Wright, B. D. (2007). The way to a faculty member's head is through the discipline. In T. W. Banta (Ed.), *Assessing student learning in the disciplines: Assessment update* (pp. 240–258). San Francisco, CA: Jossey-Bass.

Yin, R. K. (2014). *Case study research: Design and methods*. Thousand Oaks, CA: Sage.

Authors

SU SWARAT is the director of assessment and educational effectiveness at California State University, Fullerton. Her research interests include assessment and evaluation; science, technology, engineering, and mathematics education; student interest and motivation; and faculty development.

PAMELLA H. OLIVER is the interim associate vice president of academic programs and a professor in the Department of Child and Adolescent Studies at California State University, Fullerton. She is an associate director of the Fullerton Longitudinal Study, and her research involves family processes, temperament, children's adjustment, and academic success.

LISA TRAN is a professor of history at California State University, Fullerton. Her research focuses on women and the law in 20th-century China.

J. G. CHILDERS is an associate professor in the Department of Physics at California State University, Fullerton. His research interests vary from the reactions of low-energy electrons with atoms and molecules to algorithms for factoring large composite integers into prime numbers.

BINOD TIWARI is a professor of civil and environmental engineering at California State University, Fullerton. His research interests include geotechnical engineering, earthquake engineering, natural hazard mitigation, and engineering education. He also serves a program evaluator for ABET.

JYENNY LEE BABCOCK is a senior assessment and research analyst at California State University, Fullerton. Her research interests include the promotion of well-being, food environment, and the ancestral health movement.