

Establishing a Framework for Assessing Teaching Effectiveness

Shawn R. Simonson, Brittnee Earl & Megan Frary

To cite this article: Shawn R. Simonson, Brittnee Earl & Megan Frary (2021): Establishing a Framework for Assessing Teaching Effectiveness, College Teaching, DOI: [10.1080/87567555.2021.1909528](https://doi.org/10.1080/87567555.2021.1909528)

To link to this article: <https://doi.org/10.1080/87567555.2021.1909528>



© 2021 The Author(s). Published with
license by Taylor & Francis Group, LLC.



Published online: 03 May 2021.



Submit your article to this journal



View related articles



View Crossmark data

Establishing a Framework for Assessing Teaching Effectiveness

Shawn R. Simonson, Brittnee Earl, and Megan Frary

Boise State University

ABSTRACT

Most institutional practices to evaluate teaching are inadequate, inaccurate, neither improve teaching directly, nor incentivize teaching improvement. This is often because effective teaching is difficult to assess and most tools do not adequately or accurately do so and are often without established standards. Because of this, faculty may be hesitant to change or may not be aware of the need to change their teaching practice or how to effect such change. Here we establish a framework defining effective teaching and develop a tool that considers multiple facets of teaching and will accommodate different approaches, modes, and environments.

KEYWORDS

Teaching effectiveness; formative/summative assessment of teaching; learner-centred; scholarly teaching

Introduction

In the fall of 2019, a Roundtable on Systemic Change in Undergraduate STEM Education was convened by the National Academies of Science, Engineering, and Medicine and included a *Recognizing and Evaluating Science Teaching in Higher Education* workshop (National Academies of Sciences, Engineering, and Medicine 2020). “The goal of the workshop was to identify the questions, challenges, and levers for change that may be useful to consider in order to implement improvements in the teaching evaluation process, with the core mission of improving instruction and contributing to students’ success” (p. 2) (National Academies of Sciences, Engineering, and Medicine 2020). It was not only identified that the way in which teaching is evaluated must change, but so must the culture around teaching, assessing teaching, and the place of teaching in the hierarchical status of higher education (National Academies of Sciences, Engineering, and Medicine 2020).

Prior to this roundtable, a 2017 report from the Commission on the Future of Undergraduate Education, supported by the American Academy of Arts and Sciences, suggested that the undergraduate college experience must be improved to meet the demands of the evolving society in the United States (The Commission on the Future of Undergraduate Education 2017). Two of the three parts of the

proposed national strategy—ensuring that students have high-quality learning experiences and that institutions increase their overall completion rates and reduce inequities among student groups—relate directly to teaching (The Commission on the Future of Undergraduate Education 2017). The report indicates that undergraduate students need to more effectively learn and master the knowledge, skills, and dispositions that will lead to their success in our changing country (The Commission on the Future of Undergraduate Education 2017). The report goes on to say, “Ultimately, though, making undergraduate learning stronger and more rigorous will depend upon how undergraduate education invests in the teaching skills of its faculty and the kind of institutional and systemic commitment that is made.” (p. 22) (The Commission on the Future of Undergraduate Education 2017) The Commission recommends that colleges and universities need to invest more “in providing students with consistently good teaching.” (p. 75) (The Commission on the Future of Undergraduate Education 2017) Institutions of higher education need to know if faculty are teaching well (Fink 2008). Structures and mechanisms need to be in place to hold institutions and faculty accountable for their role in student learning. However, in many higher education contexts, current procedures to evaluate teaching are inadequate, nebulous, and

CONTACT Shawn R. Simonson  ShawnSimonson@BoiseState.edu  Department of Kinesiology, Boise State University, 1910 University Drive, Boise, ID 83725-1710, USA.

© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

inaccurate—many times relying solely on student course evaluations and peer observations, which paint a very incomplete picture—and neither improve teaching directly nor incentivize teaching improvement (Stupnisky et al. 2018; Shadle, Marker, and Earl 2017; Berman 2003; Fink 2008; National Academies of Sciences, Engineering, and Medicine 2020; Myyry et al. 2020).

The complexities of teaching and learning

According to Merriam-Webster, learning is “knowledge or skill acquired by instruction or study” and from a behaviorist standpoint it is “modification of a behavioral tendency by experience.” (Merriam-Webster, Inc 2018) Teaching, in turn, is defined as “to cause to know something,” “to guide the studies of,” and “to impart the knowledge of” (Merriam-Webster, Inc 2018). In other words, teaching is the process of helping others acquire knowledge or skill and/or to modify another’s behavior. Therefore, it would seem that holding faculty accountable for their teaching could be done by simply demonstrating that students have acquired the necessary knowledge or skill and/or that a desired behavioral change has occurred.

However, various circumstances might influence a student’s ability to learn such as their level of prior preparation or knowledge, previous experiences, social interactions within (and outside of) the classroom—including with the instructor, environmental interactions, cognitive development, instructional style, etc. (Karplus and Thier 1967; Tolman and Kremling 2017; National Academies of Sciences, Engineering, and Medicine 2020). Only a maximum of 59% of the variance in student performance can be attributed to the teacher and class (Alton-Lee 2003). As suggested by Chew and Cerbin:

... the everyday reality of education: teaching and learning are complex and hard. They are complex and hard because we don’t know the exact conditions in which student learning will occur. How people learn depends on multiple interacting factors that defy any one-size-fits-all solution. Yet we keep trying to find a simple solution to this complicated problem. (Chew and Cerbin 2017)

In addition to the complexities related to a student’s ability to learn, defining “good teaching” is problematic and might best be done using a multidimensional characterization that considers the complex tasks of teaching (Fink 2008; Gurney 2007; Benton and Young 2018). For example, faculty members engage in many activities as instructors: content and web development, course planning, instructional

activity design, course material construction, information presentation, assessing and evaluating learning, providing feedback and motivation, etc. (Franklin 2001; Arreola 2000; Seldin 2000). Thus, evaluating quality teaching, or one’s ability to help others acquire knowledge or skill and/or to modify another’s behavior is difficult and takes multiple perspectives to do so (Fink 2003; Berman 2003; Boyer 1990; McCabe and Layne 2012; Arreola 2000; Seldin, Miller, and Seldin 2010).

The complexities associated with teaching and learning in general are exacerbated in higher education due to the lack of sufficient faculty preparation for teaching, which may further cloud the interaction between teaching and learning (National Research Council 2012). It is generally assumed that when someone is hired within a college or university department they have the first level of professional teaching knowledge—content knowledge (Pallas, Neumann, and Campbell 2017; Fink 2008). The other two levels of professional teaching knowledge are general pedagogical and content pedagogical (Pallas, Neumann, and Campbell 2017). While the former represents the general teaching tools that are broadly applicable across disciplines, the latter requires an appreciation of students’ prior understanding, core concepts, threshold concepts, and how to help students think like subject-matter experts (Pallas, Neumann, and Campbell 2017; Wieman 2019). These two levels of professional knowledge are often missing as a result of the general lack of teaching preparation (Robinson and Hope 2013; Wieman 2019).

Challenges with assessing teaching

As stated previously, the current procedures and policies related to the assessment of teaching effectiveness in higher education are inadequate, nebulous, and inaccurate. In general, there is a substantial over-reliance on student course evaluations and in some cases student course evaluations are the sole source of information used to evaluate teaching (Seldin 1999b; Miller and Seldin 2014). Student course evaluations can provide the student voice in the evaluation process as well as the student perception of the teacher’s affect; however, recent research shows the use of student course evaluations or “student satisfaction surveys” is suspect (McCabe and Layne 2012; Franklin 2001; Bain 2004; Feldman 1978; Flaherty 2015; Richmond et al. 2014; Esarey and Valdes 2020). As noted earlier, faculty play many roles as instructors and student course evaluations only address a few of

those roles—in none of which students are experts (Franklin 2001; Seldin, Miller, and Seldin 2010; National Academies of Sciences, Engineering, and Medicine 2020; Boysen, Richmond, and Gurung 2015). Student course evaluations do not consider what pedagogical choices were made, why they were made, how they were implemented, and the factors that played a role in the successes and failures in the course (Seldin, Miller, and Seldin 2010). In addition, an issue with comparing course evaluations between classes, years, instructors, and against a standard is that the ratings include some error from mismarks, misreads, or less-than-optimum sampling and are not precise (Franklin 2001). Statistically significant differences can also not be identified as means are often being calculated on data from a Likert scale; it is rare for standard deviations, errors, or confidence intervals to be reported for the means, and sampling size can be an issue when comparing very small courses to very large (Franklin 2001). In addition, survey administration methods have changed at many institutions from in-class paper to out of class on-line and this tends to reduce response rate and scores (Franklin 2001; Flaherty 2015). The questions and analysis may have also changed over the years. Moreover, students tend to assign lower scores to courses heavy in science, those that are required, and to instructors using unfamiliar pedagogies (Franklin 2001; Bain 2004; Feldman 1978; Flaherty 2015). The instructor's gender, race, and identity can also influence student course evaluations (Flaherty 2015; Mitchell and Martin 2018; Boring, Ottoboni, and Stark 2016; Bavishi, Madera, and Hebl 2010; Schmidt 2015). The analysis of written comments is fraught with peril as comments are often out of context, open to interpretation, and given more weight if well-written (Franklin 2001). In addition, a 2016 study indicates that the relationship between student satisfaction surveys and the quality of learning, as measured by student performance in subsequent courses, may actually be inversely related (Kornell and Hausman 2016) and those instructors whom students rate the lowest, may, in fact, be the best teachers.

Another common practice in assessing teaching is requiring classroom observations. While classroom observations can provide insight into the teaching practices, there are drawbacks to relying on this data for evaluation or assessment purposes. First, conducting classroom observations on a large scale is time consuming and cumbersome, requiring significant resources (Durham et al. 2018; Teoh, Ming, and Khan 2016). Second, who is conducting the observations has

important implications related to the type and quality of information resulting from the observation (Durham et al. 2018; Teoh, Ming, and Khan 2016). For example, an external observer may overlook nuances in pedagogical approaches specific to a discipline, they may misinterpret classroom culture and student engagement, and they often lack expertise in the content area (Durham et al. 2018). Peer observations mitigate some of these challenges; however, the majority of faculty have had no formal education in education and lack the background to develop or participate in effective evaluations (Robinson and Hope 2013) or provide the critical feedback needed (Teoh, Ming, and Khan 2016). In addition, neither of these two observation protocols will optimally capture continuous pedagogical improvement, the use of multiple pedagogies, and sharing lessons learned about teaching (Richmond et al. 2014).

The assessment of teaching is complicated further by being within a system that requires consensus building around almost all decisions—especially those surrounding job performance and job security (Gray 2016, Teaching Quality Framework Initiative, Center for STEM Learning 2017). In higher education, job performance and security are directly related to the tenure and promotion process and policies. These policies often prioritize research over teaching leading to ineffective teaching as faculty focus the majority of their efforts on research endeavors and neglect teaching related tasks (National Academies of Sciences, Engineering, and Medicine 2020; Miller and Seldin 2014). In *Scholarship Reconsidered*, Boyer asserts, “For teaching to be considered equal to research, it must be vigorously assessed, using criteria that we recognize within the academy, not just in a single institution” (p. 37) (Boyer 1990).

The problems outlined above are succinctly highlighted in the Commission on the Future of Undergraduate Education's report, “Widespread inattention to teaching quality in the preparation, selection, and assessment of faculty is a major obstacle to improved undergraduate student learning” (p. 22) (The Commission on the Future of Undergraduate Education 2017). The Commission's recommendation is:

Institutions must make a systemic commitment to the improvement of college teaching, a commitment that acknowledges and rewards good teaching practices that are grounded in the learning sciences and an understanding of the variety of experiences and learning styles students bring to campuses. (p. 23) (The Commission on the Future of Undergraduate Education 2017)

Table 1. Teaching evaluation best practices (Benton and Young 2018).

Assessment usefulness depends on what the instructor does with the information.
Assessment system is fair, valid, reliable, and practical.
Upholds ethical and legal principles to protect those being evaluated.
Institutional culture values teaching assessment and supports the process with appropriate policies and procedures in place.
Institutional culture is clear and consistent with standards and definitions of teaching expectations.
Standardization enhanced by including formal and informal measures to allow for individual differences and situational factors.
Flexible assessment schedule to allow for variety in career trajectories and development.
Balanced assessment system that considers student voice, self-assessment, and class design and implementation.
Authentic measures of student performance demonstrate what students can do.
Assessment process helps instructors improve teaching and results in improved student learning.
Supports and rewards a mastery approach to teacher development.
Sensitive to diversity, equity, and inclusion and the roles these play in teaching, learning, and assessment.
Appropriately generates and uses statistical measures.

Developing a framework for assessing teaching in higher education

In a recent IDEA paper, Benton and Young identify a number of best practices for evaluating teaching (Table 1) (Benton and Young 2018). Essentially, these 13 best practices recognize that evaluation is only as valuable as the work that goes into it, both from those being evaluated to those doing the evaluation. An institution-wide system that includes student, instructor, and relevant others voices which is accepted, accurate, authentic, and enforced is most likely to succeed (Benton and Young 2018). Such a system that also considers the diversity of instructors and students can engender a growth mindset, challenge teachers to be better, support their development, and enhance the value of teaching and assessment to achieve the ultimate goal of improving student learning (Benton and Young 2018).

While there are frameworks for assessing teaching used in primary and secondary education, there are few consistently used in post-secondary teaching (Danielson 2014, Gray 2016, Teaching Quality Framework Initiative, Center for STEM Learning 2017; Fink 2008). Thus, we undertook the task of creating a framework that defines effective teaching and allows for formative and summative assessment of college-level teaching. It is suggested that evaluating teaching should consider both student learning as well as the teacher's humanity and consider both quantitative and qualitative data (North 1999). Because teaching and learning are so complex, North goes on to recommend the use of peer evaluation, student

achievement and input, and personal reflection to analyze the individual teacher (North 1999). The American Sociological Association, endorsed by 21 other organizations, suggest a "holistic assessment that includes peer observations, reviews of teaching materials, and instructor self-reflections" (p. 2) (ASA 2019). The Society for the Teaching of Psychology recommends using evidence regarding "training, instructional methods, assessment process, syllabi, content, and student evaluations of teaching" (p. 282) (Richmond et al. 2014). In addition, enough is now known about teaching that evidence-based measurements of teaching can be made from research-based guidelines (Stains et al. 2018; North 1999).

The first step in developing the framework for assessing teaching effectiveness (FATE) was creating a definition of effective teaching. As Robert M. Pirsig argued, quality is difficult, if not impossible, to define (Pirsig 1974); thus, we choose to pursue effective teaching over the challenge of defining quality. So then, what is effective teaching? As previously stated, simply demonstrating student learning may not be enough (Chew and Cerbin 2017). In 1987, Chickering and Gamson identified seven principles of good teaching that still hold true: it should 1) encourage contacts between students and faculty, 2) develop reciprocity and cooperation among students, 3) use active learning techniques, 4) give prompt feedback, 5) emphasize time on task, 6) communicate high expectations, and 7) respect diverse talents and ways of learning (Chickering and Gamson 1987). More recently, Wieman agreed with Stains and suggested that enough controlled, validated, and peer-reviewed research has been conducted that we can now define teaching expertise based on cognitive psychology, brain research, learning science and discipline-based educational research (Wieman 2019). Between the two of them, Seldin and Bain identified the 18 behaviors and practices of effective teachers found in Figure 1 (Seldin 1999b; Bain 2004). Arreola (2000) and Fink (2008) conclude that teaching involves four fundamental tasks: 1) content and knowledge of the subject matter, 2) delivery and interacting with students, 3) course and material design, and 4) management of the course, students, and materials. Fink goes on to place these tasks in four "key dimensions" of teaching: 1) learning experience design, 2) student/teacher interaction quality, 3) extent and quality of learning, and 4) the teacher's improvement efforts (Fink 2008). Palmer, in *The Courage to Teach*, agrees and suggests that effective teachers 1) take teaching seriously, 2) carefully design the course, 3) identify and address

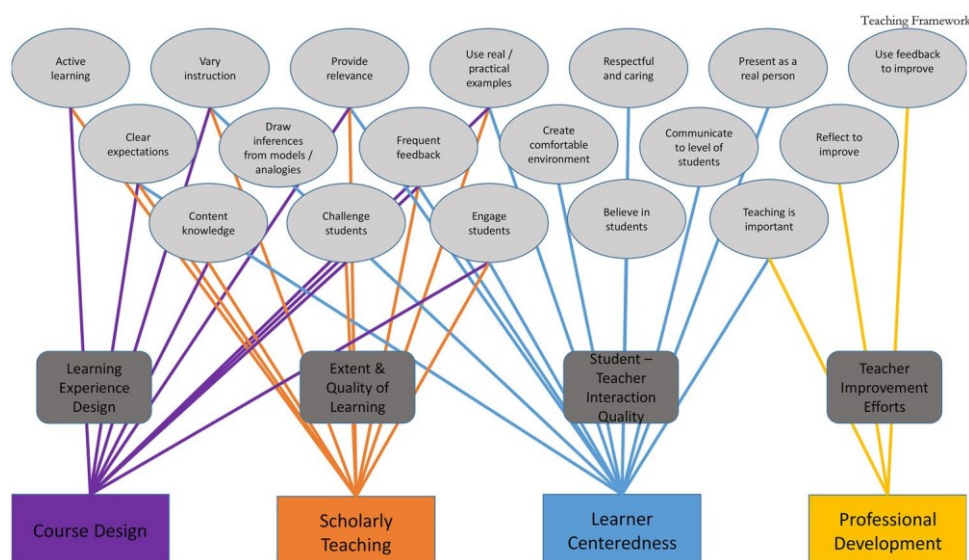


Figure 1. Interaction of the beliefs, traits, tasks, dimensions, and criteria of effective teaching (Seldin 1999a; Bain 2004; Arreola 2000; Fink 2008; Palmer 2017).

course issues, 4) learn from mistakes, and 5) share their knowledge of teaching to help others (Palmer 2017). These behaviors, practices, tasks, dimensions, and questions are classified into criteria in Figure 1.

When looking at Figure 1, it becomes evident that no single metric can measure teaching effectiveness. Thus, the evaluation of teaching should take into consideration the multiple dimensions of teaching, sources of relevant information, and criteria and standards (Fink 2008; Seldin, Miller, and Seldin 2010; National Academies of Sciences, Engineering, and Medicine 2020; Benton and Young 2018; Esarey and Valdes 2020). Because this is an assessment of teaching, an approach similar to planning a degree or course curriculum was used; the path for developing a teaching framework started with the outcomes and moved backward to identify assessments, then a guide was developed for successfully completing the assessments and achieving the outcomes (Arreola 2000; Wiggins and McTighe 2005; Fink 2008). In the present work, Arreola's eight steps were also implemented (Arreola 2000).

Comprehensive teaching evaluation considers the students from who they are to what they are doing (Cornelius-White 2007). The effectiveness of the tools and pedagogies used are evaluated (Richlin 2001). The course makes sense in that the outcomes, assessments, and activities are aligned and interrelated (Fink 2003; Alton-Lee 2003). In addition, because knowledge, pedagogies, students, and teachers change, the evaluation of teaching considers the efforts of the teacher for continued improvement (Brancato 2003; Mrig, Fusch, and Cook 2014). Thus, we have created a framework that focuses on these four elements of effective

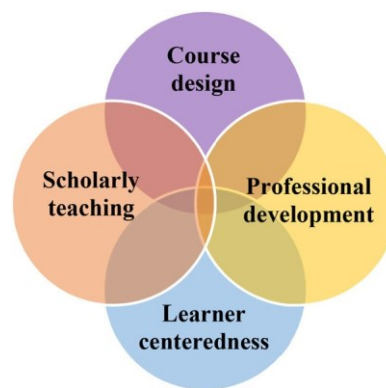


Figure 2. Model of the elements of optimized teaching.

teaching (Figure 2): learner-centered, scholarly teaching, course design, and reflective professional development, which at the intersection of these four criteria, student learning is optimized (Fink 2003; Berman 2003; Boyer 1990; McCabe and Layne 2012; Arreola 2000; Danielson 2014).

In what follows, the theoretical underpinnings related to each of the four elements of the framework are described. While there is obviously some overlap as indicated in Figure 1, we have tried to put each behavior, belief, task, and action in the element in which it most strongly fits. These elements come together as the criteria for effective teaching in a rubric (Figure 3), where each criterion is further defined, levels of achievement (exemplary, proficient, developing, missing) are established, and suggestions are provided about the type of evidence that one might include to demonstrate their efforts and effectiveness in each area. It is not the intent of this framework and evaluation system to force faculty in to a

CRITERION 1: Course Design: Designs course materials in alignment with course learning outcomes Exemplary teachers will design their courses around appropriate course learning outcomes, design a variety of summative and formative assessments, which effectively measure student achievement of those outcomes, and create course activities which support students in reaching the course learning outcomes.					
Possible sources of evidence	Category	Exemplary	Proficient	Developing	Missing
Syllabus Course assignments Student work samples Other course materials Course design table	1.1. Course learning outcomes (LOs) guide course design process	Well-developed course learning outcomes (more info) are present for all courses and guide the course design process.	Course learning outcomes are present for all courses and guide the course design process.	Course learning outcomes are present, are not well-developed, and/or do not guide course design process.	Course learning outcomes are absent.
	1.2. Alignment of assessments	Assessments obviously align with course learning outcomes.	Most assessments align with course learning outcomes.	Assessments are present and do not appear to match the course learning outcomes.	Without course learning outcomes, alignment of assessments cannot be determined.
	1.3. Alignment of course activities	Course activities are consistently aligned with, and thus support students in working toward, course learning outcomes and assessments; alignment is explicitly established and communicated.	Course activities are somewhat aligned with, and thus support students in working toward, course learning outcomes and assessments; alignment not explicitly established or communicated.	Course activities are not clearly aligned with course learning outcomes and assessments.	Without course learning outcomes, alignment of activities cannot be determined.
	1.4. Course design and learning outcomes encourage discipline-specific ways of thinking	Course activities teach students to think about and use the subject like a practitioner in the discipline, consistent with the students' background and level.	Some development of discipline-specific ways of thinking is evident, but it is not clear that this is successful or that it is consistent with the students' background or level.	Few course activities appear to support discipline-specific ways of thinking or this process is not demonstrated in a meaningful way.	Activities do not appear to help students develop discipline-specific ways of thinking.
	1.5. Student achievement of course learning outcomes	Ensures that students are achieving course LOs by <i>reflecting on student work</i> . Student work samples demonstrate substantial achievement of course learning outcomes.	Monitors student achievement of course LOs. Student work samples demonstrate achievement of course learning outcomes.	Student work samples present a tenuous link to course learning outcomes.	Student work samples do not appropriately demonstrate student success OR student work samples are absent.
CRITERION 2: Scholarly Teaching: Implements evidence-based practices Exemplary teachers will implement a variety of evidence-based instructional practices (EBIPs) in their daily teaching and assessments in order to best support student learning and students' development as learners. Note: Instructor does not need to cite the literature regarding EBIPs, but can refer to EBIPs from this list ; use of additional EBIPs is also encouraged.					
Possible sources of evidence	Category	Exemplary	Proficient	Developing	Missing
Examples of course activities or other teaching materials Examples of summative and formative assessments (more info) Feedback from peer teaching observation Course design table Class observation with COPUS, RTOP, or similar tool (more info)	2.1. Situational factors considered	Instructional choices are clearly guided by a thoughtful examination of all five categories of situational factors (described here).	Instructional choices are guided by some of the situational factors (described here) OR there was a deep examination but not clear implementation of what that meant.	Briefly considers how the situational factors (described) around the course and/or student prior knowledge affect choice of activities.	Does not consider how the situational factors (described here) around the course and student prior knowledge affect choice of activities.
	2.2. Relationship between instructional practices and learning outcomes	Provides a strong rationale /reflection linking the instructional practices with the learning outcomes.	Provides a rationale/reflection linking the instructional practices with the learning outcomes.	Rationale/reflection tenuously links instructional practices with the learning outcomes.	Does not provide a rationale or reflection linking the instructional practices with the learning outcomes.
	2.3. Implementation of EBIPs	Frequently implements a variety of evidence-based instructional practices (EBIPs) as appropriate for the course and diversity of situational factors .	Implements a more limited variety of EBIPs as appropriate for the course and diversity of situational factors.	Only occasionally implements a narrow variety of EBIPs without evidence of consideration for the course and diversity of situational factors.	Does not implement EBIPs or EBIPs appear to be inappropriate for the course and diversity of situational factors.
	2.4. Assessments follow good practices	Assessments, both formative and summative , are authentic , varied, and offer students choices. (more information)	Assessments, both formative and summative, are authentic, varied or offer students choice.	Assessments, both formative and summative, lack variety or student choice.	Assessments are limited.
	2.5. Assessment criteria are effectively communicated.	Criteria for evaluation are explicitly defined and clearly communicated.	Criteria for evaluation are occasionally defined and/ or are inefficiently communicated.	Criteria for evaluation are poorly defined and/or are poorly communicated.	Criteria for evaluation are not defined.

Figure 3. The FATE rubric for evaluating evidence of teaching effectiveness.

ITERION 3: Learner-Centered: Uses an inclusive, learner-centered approach

Exemplary teachers will design courses and course materials that focus on learning and the learner, rather than the instructor, and implement inclusive teaching practices which reach all learners and provide students opportunities for success.

Possible sources of evidence	Category	Exemplary	Proficient	Developing	Missing
Syllabus Examples of inclusive teaching practices [exampleshere] Course design table Examples of course materials Peer observation and/or MAP Student surveys Student course evaluations Class observation with COPUS, RTOP, or similar tool [more info]	3.1. Student engagement during class	During the majority of class, students are actively engaged with the course content, the instructor, and each other.	During class, students are actively engaged with the course content, the instructor, and each other.	During class, students are only occasionally engaged actively with the course content, the instructor, and/or each other.	During class, students are not actively engaged with the course content and do not interact with each other.
	3.2. Learning activities	Learning activities are consistently authentic , engaging, varied, and appropriate for students.	Learning activities are engaging, varied, AND appropriate for students.	Learning activities are engaging, varied, OR appropriate for students.	Learning activities are not obviously engaging, varied, or appropriate for students.
	3.3. Student-centered approach in course materials	Course materials (e.g., texts, presentation, movies, readings, etc.) consistently communicate an inclusive, student-centered approach (defined here) AND considers situational factors [described hfill] .	Course materials (e.g., texts, presentation, movies, readings, etc.) communicate an inclusive, student-centered approach (defined here) OR consider situational factors [described hfill] .	Course materials imply some effort has been made to adopt an inclusive, student-centered approach (defined here) with no evidence of consideration of situational factors [described hfill] .	Course materials do not communicate an inclusive or student-centered approach.
	3.4. Instructor behaviors	The instructor supports student learning by providing timely feedback, communicating effectively, and being trustworthy and appropriately available to students.	The instructor makes efforts to support student learning by providing timely feedback, communicating effectively, and being trustworthy and appropriately available to students; there is room for improvement .	The instructor minimally supports student learning by providing timely feedback, communicating effectively, or being trustworthy and available to students.	The instructor does not provide timely feedback, communicate effectively, engender trust, or make themselves available to students.
	3.5. Classroom climate	Teaching practices support a classroom climate which promotes a sense of belonging, values diverse contributions, respects individual differences, and encourages motivation, cooperation, and engagement [examples] .	Teaching practices support a classroom climate which mostly promotes a sense of belonging, values diverse contributions, respects individual differences, and encourages motivation, cooperation, and engagement [examples] .	Teaching practices support a classroom climate which somewhat promotes a sense of belonging, values diverse contributions, respects individual differences, and encourages motivation, cooperation, and engagement [examples] .	Teaching practices do not support a classroom climate which promotes a sense of belonging, values diverse contributions, respects individual differences, and encourages motivation, cooperation, and engagement [examples] .

Criterion 4: Practices reflective teaching to drive continuous improvement of teaching

Exemplary teachers will be reflective practitioners who use feedback from a variety of sources (students, peers, CTL, department, self) to seek a variety of approaches to continuously improve as teachers.

Possible sources of evidence	Category	Exemplary	Proficient	Developing	Missing
- Student course evaluations - Reflection on and response to course evaluations - List of professional development activities - Continuous improvement plan and reflection - Mid-semester assessment (MAP) report - and responses - Feedback from peer teaching observation - End-of-semester course reflection - Teaching log - Class observation with COPUS, RTOP, or similar tool [more info]	4.1. Professional development	Engages frequently with professional development opportunities (e.g., three or more per year).	Engages occasionally with professional development opportunities (e.g., one or two per year).	Engages infrequently with professional development opportunities (e.g., once every other year).	Does not engage with professional development opportunities.
	4.2. Self-reflection	Demonstrates a high level of self-reflection around teaching broadly, <i>objectively describing their strengths and weaknesses, consistent with evidence of teaching practices</i>	Demonstrates self-reflection around many aspects of teaching, objectively describing their strengths and weaknesses, consistent with evidence of teaching practices.	Demonstrates a limited amount of self-reflection around teaching, for example, by not identifying strengths and weaknesses or considering too narrow of a focus, or evidence is not sufficiently aligned with reflection.	Does not demonstrate self-reflection around teaching.
	4.3. Continuous improvement plan	Develop s, implements, and updates continuous personal improvement plan related to teaching.	A continuous improvement plan relative to teaching is present, but there are gaps in its implementation or adaptation.	Some evidence of a continuous improvement plan is present, but it is not well developed, implemented, or updated.	No continuous improvement plan related to teaching.
	4.4. Incorporates feedback	Consistently implements changes to teaching <i>as a result of reflection on multiple</i> sources of feedback.	Consistently implements changes to teaching as a result of reflection on limited sources of feedback.	Occasionally makes changes to teaching or solicits feedback about teaching.	No evidence of how feedback is collected or incorporated in teaching.
	4.5. Shares lessons learned about teaching with others	Demonstrates leadership as related to sharing lessons learned about teaching and/ or learning.	Sustained engagement in sharing lessons learned about teaching and/or learning.	Participates in sharing lessons learned about teaching and/ or learning.	Does not share lessons learned about teaching and/or learning.

Figure 3. Continued.

Table 2. The four major curriculum theories (Smith 1996, 2000; Bali 2018).

Curriculum as	
<i>Knowledge</i>	to be transmitted
<i>Products</i>	to be produced (outcomes to be met)
<i>Process</i>	what happens in the classroom and pertains to the three-way interaction between teachers, students, and knowledge
<i>Praxis</i>	informed and committed action based on an iterative interaction between planning, learning, and evaluation

particular pedagogy; the framework does not bias one philosophy of learning over another, and it is inclusive enough to capture multiple perspectives, disciplines, and approaches to teaching. A comprehensive framework that considers all four major curriculum theories (Table 2) should adequately address the philosophies of the instructors and the needs of the learners at multiple levels (Smith 1996, 2000; Krathwohl 2002; Bali 2018; Fink 2008).

Effective course design

Not only does the effective teacher have the base professional skills and knowledge of content expertise, practice/clinical skills, research techniques, and strategies for keeping current, they also have the skills and knowledge in instructional design and delivery (Arreola 2000). Effective course design integrates learning goals, scaffolded teaching and learning activities, feedback and assessment, and situational factors (Fink 2003; Gurney 2007; Fink 2008; Richmond et al. 2014; Alton-Lee 2003). Exemplary teachers design their courses around disciplinary and student appropriate course learning outcomes, design a variety of summative and formative assessments that effectively measure student achievement of those outcomes, and create course activities that support students in reaching and demonstrating completion of the course learning outcomes (Gurney 2007; Fink 2008; Richmond et al. 2014; Alton-Lee 2003). Pedagogical and assessment choices align with the learning outcomes, student diversity, and situational factors (Fink 2003; Alton-Lee 2003). The instructor monitors this process using formative assessments and provides goal-referenced, actionable, appropriate, consistent, non-threatening, ongoing, tangible and transparent, and timely feedback (Wieman 2019; Wiggins 2012; Alton-Lee 2003). Effective course design includes designing learning activities that requires students to make decisions using discipline-specific knowledge, skills, and reasoning (Wieman 2019). Lastly, students actually learn in the course as evidenced by success in and after the course (Fink 2008; Gurney 2007). Success

Table 3. Characteristics of effective course design and alignment (Alton-Lee 2003; Fink 2003, 2008; Gurney 2007; Richmond et al. 2014; Seldin 1999a; Wieman 2019; Wiggins 2012).

Course has a clear purpose within the overall curriculum.
Assessments and assignments are appropriate for and aligned with course learning outcomes.
Content is successfully connected to student abilities and interests.
Amount of material is appropriate to allotted time and student level.
Appropriate instructional strategies are used for particular learning tasks.
Course design and learning outcomes encourage discipline-specific ways of thinking.
Students achieve course learning outcomes.

after the course might be measured in subsequent courses, on standardized tests, and by adding value to student lives (Fink 2008; Alton-Lee 2003). Course design skills may not be readily observable to students or observers, so the characteristics of effective course design and alignment need to be documented as found in Table 3 (Berman 2003):

Scholarly teaching

Scholarly teaching is making evidence-based decisions about what will be taught and how. Bernstein *et al.* define the scientist-educator as someone who takes a scholarly approach to teaching, as they do their research, and defines learning objectives, designs learning activities to meet those objectives, and assesses the outcomes of those learning activities (Bernstein et al. 2010). Scholarly teachers identify a problem or a question about the teaching and learning connection (Richlin 2001). They use their familiarity with the educational literature to find and read the pertinent works and can determine which factors should be considered when selecting and implementing the appropriate pedagogical tools (Richlin 2001). Scholarly teachers then choose and apply a suitable intervention and systematically observe, document, and analyze the results of that intervention and make decisions about the effectiveness of the implementation (Richlin 2001; Richmond et al. 2014). They also invite external review of their classroom to obtain a different perspective (Richlin 2001). Exemplary teachers implement a variety of evidence-based instructional practices (EBIPs) in their daily teaching and assessments to best support student learning and students' development as learners (Richmond et al. 2014; Alton-Lee 2003). Table 4 summarizes the attributes of scholarly teaching.

Learning science has identified four critical components for learning success: retrieval practice, distributed practice, practice and transfer, and active learning (Cavanagh 2019). The use of EBIPs that incorporates these four components leads to improved learning, development of academic and cognitive

Table 4. Characteristics of scholarly teaching (Fink 2003, 2008; Bernstein et al. 2010; Richlin 2001; Richmond et al. 2014; Alton-Lee 2003).

Implementing a variety of evidence-based instructional practices as appropriate for the course and diversity of students.
 Providing a strong rationale linking the instructional practices with the learning outcomes.
 Assessments, both formative and summative, are authentic, varied, and offer students choices.
 Criteria for evaluation are explicitly defined and effectively communicated.
 Consideration of how all five categories of situational factors (1. specific context of the teaching/learning situation, 2. general context of the learning situation, 3. nature of the subject, 4. characteristics of the learners, and 5. characteristics of the teacher) around the course and student prior knowledge affect choice of activities.

skills, and increased retention (Freeman et al. 2014; National Research Council 2012; Wieman 2019; Cavanagh 2019). EBIPs also benefit students from groups historically underrepresented in STEM (National Research Council 2012). The 2012 National Research Council report on discipline-based educational research recommends that “current faculty should adopt evidence-based teaching practices to improve learning outcomes for undergraduate science and engineering students (p. 198) (National Research Council 2012).

Activities/information presentation that build on prior knowledge and experience enhance connections that aid retention (Tanner 2010; Tanner and Allen 2005; Wieman 2019). Effective EBIPs encourage students to develop disciplinary expertise by requiring that they use the material, tools, and ways of thinking within the discipline to make evidence-based decisions (Wieman 2019). Solving significant problems by using the newly acquired knowledge and tools enhances motivation, understanding, and provides an opportunity for students to develop expert ways of thinking (Gulikers, Bastiaens, and Kirschner 2004; Fink 2003). Students are motivated to learn when they believe that they can be successful and that the material is presented in an interesting and meaningful manner (Wieman 2019; Dweck and Leggett 1988). In addition, learning is fundamentally a social activity and the use of peer groups increases metacognition and social skill attainment (Vygotsky 1930; Grinnell 2000; Wieman 2019). Lastly, the use of EBIPs makes learning visible to the instructor and improves the instructor’s understanding of students’ learning and can guide their teaching to continually advance student learning (Wieman 2019).

Learner centeredness

In general, faculty perceptions about teacher-centered vs. learner-centered teaching need to be shifted

Table 5. Characteristics of a learner-centered course (Gurney 2007; Fink 2008; del Carmen Salazar, Norton, and Tuitt 2010; Alton-Lee 2003; Wieman 2019).

Students are actively engaged with the course content, the instructor, and each other during the majority of class.
 Learning activities are authentic, interesting, and motivating for students.
 Students’ prior knowledge and varied backgrounds are respected by including a variety of perspectives.
 Course materials (e.g., texts, presentation, movies, readings, etc.) communicate an inclusive, learner-centered approach and consider situational factors.
 The instructor supports student learning by providing timely feedback, communicating effectively, and being available to students.
 Teaching practices support a classroom climate, which promotes a sense of belonging, values diverse contributions, respects individual differences, and encourages motivation, cooperation, and engagement.

toward learners (Gibbons et al. 2018). Learner-centered instructors address the distinct needs of students, employ a variety of educational methods, encourage students to actively participate in the construction of knowledge, and also recognize that learning is a social process; therefore, attention is also paid to peer and student-instructor interactions, student collaboration, and communication (Cornelius-White 2007; Gurney 2007; Richmond et al. 2014; Alton-Lee 2003). Exemplary teachers design courses and course materials that focus on learning and the learner, rather than the instructor, and implement inclusive teaching practices that reach all learners and support all students’ success (Alton-Lee 2003). Table 5 summarizes the characteristics of a learner-centered course.

Effective learner-centered instruction entails distributed and deliberate practice with appropriate and timely feedback (Wieman 2019; Richmond et al. 2014). Learners must practice using their knowledge to make decisions and reflect on the feedback as to the “correctness” and/or effect of those decisions; and it is the instructor’s responsibility to guide and support this process (Wieman 2019). In a learner-centered environment, students focus on the meaning of what is learned, how it ties to their preexisting knowledge and experience, and how to organize the information (Trigwell, Prosser, and Waterhouse 1999). This improves students’ motivation, learning, achievement, and knowledge retention (National Research Council 2012; Freeman et al. 2014; Armbruster et al. 2009; Mazur 2009; Wieman 2019).

Inquiry-based learning enhances depth of knowledge, retention, and the ability to apply the content (Prince 2004; Lo and Mendez 2019; Lewis and Lewis 2005). Misconceptions are more readily identified and corrected when students construct their own knowledge and this process is visible to the instructor (Tanner and Allen 2005). Metacognition, or self-regulated learning, is improved in the learner-centered classroom when students are made aware of their

thinking, believe in their abilities, and practice meta-cognitive skills (Schraw 1998; Dweck and Leggett 1988; Baird 1986; Alton-Lee 2003).

The use of inclusive practices creates a positive, learner-centered course climate that allows all students to be active and welcome participants in the learning environment and benefits all students (del Carmen Salazar, Norton, and Tuitt 2010; Bell, Goodman, and Ouellett 2016; Marchesani and Adams 1992; Alton-Lee 2003). There are several proposed frameworks for inclusivity; the five-dimensional model proposed by Bell, Goodman, and Ouellett (2016) considers multiple perspectives of social justice education and provides a broad footing for course design, course implementation, facilitation, and interactions amongst students and between students and teachers (Bell, Goodman, and Ouellett 2016). The five dimensions are (Bell, Goodman, and Ouellett 2016; Marchesani and Adams 1992): 1) Instructor: Reflecting on one's academic socialization, social and cultural backgrounds, values, and beliefs, and increasing awareness of the interactions of instructor's background with students' backgrounds (Marchesani and Adams 1992). 2) Students: Understanding of how the students' academic socialization, social and cultural backgrounds, values, and beliefs affect their classroom and academic experience and working to eliminate alienation, isolation, and injury of all students (Marchesani and Adams 1992). 3) Curriculum: Developing curriculum that provides students with a multicultural perspective of the course content, course materials, and sources of knowledge (Marchesani and Adams 1992). 4) Pedagogy: Utilizing multiple teaching strategies and delivery modes to meet the diverse needs and talents of the students (Marchesani and Adams 1992). 5) Classroom climate and group dynamics: Affected by and affecting the other four dimensions, this entails creating an inclusive environment in which all participants feel heard and respected and can constructively engage with each other, the instructor is believed to be trustworthy and knowledgeable, and the content and pedagogy are relevant and culturally inclusive (Bell, Goodman, and Ouellett 2016).

Professional development

Professional development is critical for teaching faculty to prepare for both the opportunities and challenges of the present and future (Brancato 2003; Mrig, Fusch, and Cook 2014; Richmond et al. 2014). Exemplary teachers are reflective practitioners who use feedback from a variety of sources to continuously improve their teaching abilities and expertise (King 2004; Brancato 2003; Gurney 2007). Part of this process should be the review of schema

surrounding teaching and learning and the opportunity for reflection (Brancato 2003). Faculty base their pedagogical decisions on their beliefs about teaching, learning, and their effectiveness using particular strategies (Gibbons et al. 2018). New teaching styles are more likely to be incorporated when faculty are confident in their ability to use that style and to successfully interact with the content and students as required by that style (Gibbons et al. 2018). Faculty who participate in relevant professional development opportunities are better able to make informed choices about what to teach and how to teach it and this can increase confidence, teaching motivation, the use of more effective strategies, and improve student learning (Gibbons et al. 2018; Stains et al. 2018; Stupnisky et al. 2018; Fink 2008). Lastly, exemplary teachers share what they learn with their colleagues through informal and/or formal discussions and presentations (Palmer 2017; Fink 2008; Richmond et al. 2014). Table 6 presents the components of professional development.

Just like in the content areas, knowledge about teaching and learning is evolving and requires a continuous commitment to maintain relevance and skill (Richmond et al. 2014). In addition, professional development in teaching is supported when one considers the reports as to what teachers do in the undergraduate classroom. Despite considerable evidence supporting active learning and more learner-centered approaches, passive lecture is still the predominant pedagogy used in STEM (National Research Council 2012). Effective professional development can change teaching behaviors when it is distributed across over time, provides evidence-based strategies, requires reflection, and intentionally encourages faculty reconsidering conceptions about teaching and learning (National Research Council 2012).

Discussion: Using the framework to assess teaching

Evaluation is assigning a value to the subjective interpretation of a collection of measurements to determine how well a desirable condition is achieved (Arreola 2000). It should be valid, fair, and guide improvement efforts (Wieman 2019). Evaluations

Table 6. Components of professional development (Gurney 2007; Fink 2008; Palmer 2017; Seldin 1999a; Richmond et al. 2014).

Frequently engaging in professional development opportunities.
A high level of self-reflection around all aspects of pedagogical practice.
A high level of reflection about one's own and one's students' positionalities and analyzing the interactions between these.
Developing, implementing, and updating a continuous teaching improvement plan.
Frequent feedback from others with reflection and change as a result.
Demonstrating leadership in sharing lessons learned about teaching and learning.

result in judgements as to whether the desirable condition is achieved (positive or good) or not achieved (negative or poor) (Arreola 2000). With judgements come consequences and, when those judgements are about human performance, emotional responses; therefore, equitable evaluation systems should provide valid and reliable data on which to base decisions (Arreola 2000; Benton and Young 2018; Myyry et al. 2020). Evaluation is inherently subjective to some degree; however, the data and tools used to appraise the data should lead multiple evaluators to the same conclusions (Arreola 2000; Benton and Young 2018). Objectivity is enhanced with what Arreola terms *controlled subjectivity*—"the consistent applications of a consensus-based set of values in the interpretation of measurement data" (p. xix) (Arreola 2000). Thus, the establishment of appropriate data and shared values is critical in developing effective evaluations and consists of a number of steps (Arreola 2000).

1. Identify the activities to be evaluated.
2. Define observable achievements, products, and/or performances
3. Identify sources of information
4. Identify the information collection process
5. Design objective, reliable, and valid evaluation form

Steps one through three were completed by developing the framework. Within such a framework, it is evident that, as in the classroom, a compilation of evidence (or a portfolio) is more appropriate than attempting to use a single source of information to evaluate learning (Angelo and Cross 1993; Barkley and Major 2016; Seldin 2000; Fink 2008; Seldin, Miller, and Seldin 2010; Richmond et al. 2014; Esarey and Valdes 2020). The expectation here is that the information collection process will result in a teaching portfolio that can be used for both formative and summative assessments and requires the teacher to provide evidence demonstrating achievement

(Zubizarreta 1999; Seldin, Miller, and Seldin 2010). It is recommended that faculty document their teaching practice by including multiple sources of teaching-related evidence that illustrate teaching effort and effectiveness, including student work, course materials, rationales for instructional choices, summaries of mid-semester evaluations, etc. (Seldin 2000; Zubizarreta 1999; Fink 2008; Seldin, Miller, and Seldin 2010). Undertaking this process aids the teacher in self-assessment by encouraging analysis of various data around their teaching, reflection about their teaching, articulation of successes, and providing evidence to support their self-assessment (Zubizarreta 1999; Fink 2008; Seldin, Miller, and Seldin 2010).

Berman, Boyer, and Arreola provided suggestions for different types of evaluative data and who can effectively provide each type of data (Table 7) (Berman 2003; Boyer 1990; Arreola 2000). These materials can be used in the teaching portfolio and/or tenure and promotion dossier (Berman 2003; Boyer 1990; Seldin, Miller, and Seldin 2010; National Academies of Sciences, Engineering, and Medicine 2020). Selected evidence to support a "factual description" (p. 4) (Seldin, Miller, and Seldin 2010) of teaching philosophy, instructional activities, performance, outcomes, situational factors, goals, and reflections is included in the portfolio to present a complete, representative sample of teaching effectiveness (Seldin, Miller, and Seldin 2010; National Academies of Sciences, Engineering, and Medicine 2020; Esarey and Valdes 2020).

Teaching portfolios are not new in higher education and the majority of institutions require them for promotion and tenure, but their effective use is likely limited (De Rijdt et al. 2006). Concern about using portfolios may be related to the faculty choosing what goes into them. However, recent research suggests that portfolios constructed around guidelines tend to be accurate and are an acceptable mechanism for documentation (Gibbons et al. 2018; Smith et al. 2014; Seldin 2000; Drinkwater, Matthews, and Seiler 2017).

Table 7. Sources of evaluative data for teaching activity. (Berman 2003; Boyer 1990; Arreola 2000; Fink 2008; Miller and Seldin 2014).

Type of teaching-related data	Students	Peers	Department review	External review	Self
Assessing learning	Maybe	Maybe	Maybe	Maybe	Maybe
Content expertise	No	Yes	Yes	Yes	Yes
Content and course materials design	Maybe	Yes	Maybe	Yes	Yes
Course management	No	No	Yes	Yes	Yes
Course design	No	Maybe	Yes	Yes	Yes
Faculty development	No	No	Yes	Yes	Yes
Information presentation	Yes	Yes	Maybe	Yes	Yes
Instructional activity design	Yes	Yes	Maybe	Yes	Yes
Providing feedback and motivation	Yes	No	No	No	Yes
Student learning	Maybe	Maybe	Maybe	Yes	Yes
Student/Teacher interactions	Yes	Maybe	Maybe	No	Yes

Resistance to using portfolios may also be related to perceptions that faculty lack time to properly construct one, increasing bureaucracy, and an overall lack of value (De Rijdt et al. 2006). Another criticism is that portfolios may be a nice way to document teaching but they are unconnected to the evaluation process and, therefore, of little value. The FATE framework and accompanying rubric should ease these concerns by providing a mechanism for evaluating a portfolio and directly connecting it to the performance evaluation process. In addition, the majority of faculty report that they can effectively document their teaching behaviors and activities and more experienced teachers can provide more evidence (Boysen, Richmond, and Gurung 2015). As previously stated, a benefit to portfolios is that most institutions already use something similar to this in the promotion and tenure dossier; thus, FATE and the accompanying rubric build on an existing tool and are simply standardizing the evaluation of the dossier.

Arreola's fifth step in creating an evaluation tool is to design an objective, reliable, and valid evaluation form (Arreola 2000). Objectivity, or fairness, protects the teacher from personal interpretations and agendas and is enhanced by the use of multiple sources of evidence and a standard evaluation tool (Benton and Young 2018; Arreola 2000). A reliable measure is one that consistently leads different evaluators to the same results or conclusions and is critical when the evaluation affects a teacher's employment and compensation (Benton and Young 2018). Reliability is enhanced when multiple sources of evidence are considered and evaluators are appropriately trained to use the same criteria (Benton and Young 2018). A valid tool will measure what it is intended to (Benton and Young 2018). While student satisfaction surveys give students a voice, they cannot provide insight into why an instructor chose a particular pedagogy like a narrative can and syllabi can demonstrate alignment between learning objectives and assessments while copies of student work can demonstrate the effectiveness of this alignment; thus, multiple measures can enhance validity (Benton and Young 2018).

The need for an objective, reliable, and valid evaluation tool led to the creation of the FATE rubric for the assessment of the teaching portfolio. The use of the rubric with an existing portfolio may reduce the time and cost of evaluating faculty, as there will be clear evaluation criteria, and less time will be spent discussing nebulous measures and subjective perceptions. Establishing clear success criteria and reducing the reliance on nebulous measures will increase

transparency and engender trust in the system and reduce the inherent stress in the evaluation and promotion and tenure process (Arreola 2000; Benton and Young 2018).

As with many types of evaluation, the evaluation of teaching should serve two different purposes: formative and summative assessment (Fink 2008; Benton and Young 2018). An effective formative and summative evaluation system reviews multiple sources of data and uses a system that is: valid and comprehensive, reliable, explicit and public, flexible, time and cost effective, encouraging of periodic self-evaluation, linked to formative evaluation, linked to planned change strategies, and supported by the highest level of administration (Berman 2003).

Formative assessments provide feedback for improvement and are most effective when the expectations are clearly identified, evaluation is objective, appropriate feedback and support for continued growth are provided, and reflection is intentional (Angelo and Cross 1993; Arreola 2000; Baird 1986; Fink 2008; Benton and Young 2018). Teachers at all levels can benefit from self-evaluation and reflection to improve their teaching practice (Boysen, Richmond, and Gurung 2015). Using clearly established procedures and feedback for formative assessment and creating a supportive culture encourages a growth mindset in regards to teaching and can help instructors try new things and work toward mastery of the various skills of teaching (Benton and Young 2018). As part of formative assessments, the teacher can use the framework and rubric to create improvement action plans and monitor progress made over time (Benton and Young 2018).

Formative assessments can also lead to improved summative assessment, especially when the criteria are similar to the summative tool and a clear link can be established to the expected outcomes (Angelo and Cross 1993; Crooks 2001). Summative assessments should provide an objective and accurate representation of performance over time (Arreola 2000; Benton and Young 2018). Summative assessments often provide data for performance judgment and personnel decisions (Arreola 2000; Crooks 2001; Angelo and Cross 1993; Fink 2008).

Thus, with the data suggested by Berman, Boyer, and Arreola and the systemic requirements outlined above in mind, the FATE framework and a rubric provide faculty members with guidelines for and examples of how to document their teaching process and effectiveness (Berman 2003; Boyer 1990; Arreola 2000) for both formative and summative assessment

purposes. Formative assessments can be made during annual evaluations as faculty select particular pieces of the framework to focus on and work to improve those aspects of their teaching. They can focus on and accumulate evidence for the entire criteria over a period of years and make the building of their portfolio an incremental process. Summative assessments of teaching then occur in 4–6 year intervals and consider the entirety of the framework and portfolio. The complete portfolio will be compared to FATE using the accompanying rubric and a completed rubric and brief summary of evaluators' overall comments are included in the portfolio and the promotion/review dossier.

The use of this framework and accompanying rubric does not eliminate the use of existing teaching evaluation practices, student course evaluations, peer observation, etc., and the teacher still determines what goes into their portfolio; however, the framework provides a more holistic assessment in which each data point plays its own role in telling the story about a person's teaching. The teacher's story is tied together by including a narrative describing teaching effectiveness and how the evidence supports this. The teacher will draw a map or write a discussion that guides the evaluator from the evidence to the conclusions. Whether the evidence supports the conclusions or not will then be up to the evaluators. This process is analogous to writing a manuscript discussion and peers reviewing it for accuracy and validity.

An important aspect of the FATE framework is that it considers multiple pieces of evidence that come from a variety of sources in order to allow for flexibility given the complex nature of teaching and learning and does not rely solely on one, or a few, limited measures. Because of this, it can be used to evaluate any instructional process (i.e., to work for those teaching courses at all levels, of all sizes, across all disciplines, with all different approaches to teaching, and in all different instructional roles). Teachers themselves, department heads, annual review committees, and/or promotion and tenure committees can conduct the portfolio evaluations and arrive at similar conclusions.

As with the use of any framework and rubric, users must be appropriately trained to use the same criteria (Benton and Young 2018). Since teachers and evaluators may not possess the appropriate background knowledge in teaching, assessment, or content, there will be a need for training of all individuals involved in the teaching and the evaluation of teaching. Workshops and resources for building and assessing teaching portfolios, the framework and the accompanying rubric, and teaching improvement will need to

be provided. Portfolio evaluators must also practice using the rubric and calibrate their assessments to institutional standards and previously calibrated portfolios and rubrics. (A panel initially calibrates the rubric to established institutional standards. Evaluators then learn to use the rubric by assessing sample portfolios and are coached to meet these same standards.) This will initially result in an increased workload; however, as this process grows and becomes normalized at an institution, familiarity will increase and the training requirements will decrease.

It is evident that by increasing the amount of evidence required to evaluate teaching, the time and effort required will also increase—both for the teacher and those doing the evaluation (National Academies of Sciences, Engineering, and Medicine 2020). One suggestion is to break the framework and rubric into sections and limited portions, or particular aspects of teaching, to be focused on in the short-term (Benton and Young 2018). A comprehensive portfolio can be built over time and the complete portfolio might only be reviewed every four to six years or when a teacher is applying for promotion and/or tenure (Fink 2008).

Creating a framework and a good evaluation system is not enough; a culture that expects, supports, and rewards good teaching is critical (Fink 2008; National Research Council 2012; National Academies of Sciences, Engineering, and Medicine 2020; Myyry et al. 2020). Evaluators will have to hold teachers to a high standard, look beyond academic freedom, and require “sufficient evidence” to support the assertions made (Fink 2008; Seldin, Miller, and Seldin 2010). The evaluation of teaching must be linked to a meaningful incentive system that recognizes and rewards success in each of the four criteria so that instructors are motivated to spend the time and effort (Fink 2008; National Research Council 2012; National Academies of Sciences, Engineering, and Medicine 2020; Myyry et al. 2020). And, there needs to be a strong support system to help low performing faculty identify areas of improvement, create plans to improve, and provide tools for making those improvements (Fink 2008).

Finally, leadership can use the resulting data to determine faculty needs, allocation of resources, and strategic initiatives that can result in large-scale improvements in student success (Fink 2008; Zubizarreta 1999; Seldin, Miller, and Seldin 2010). For example, support units, such as centers for teaching and learning can use aggregated data to determine and prioritize programming to better support faculty in their teaching endeavors. The use of the framework

and rubric may also improve institutional reporting during annual reviews and the accreditation process with use of clear criteria and documentation of meeting those success measures. In addition, the data will make teaching and career achievements more clear and improve the presentation and meaning of teaching awards.

Conclusion

In order to provide students a quality education and the skills necessary to meet the demands of the evolving society, institutions need to invest in structures and mechanisms that hold faculty accountable for the quality and effectiveness of their teaching (National Academies of Sciences, Engineering, and Medicine 2020). Overall, the current methods of assessing teaching neither improve teaching directly nor incentivize teaching improvement (Stupnisky et al. 2018; Shadle, Marker, and Earl 2017; Berman 2003; National Academies of Sciences, Engineering, and Medicine 2020; Myyry et al. 2020). As previously intimated by Fink (Fink 2008), it is anticipated that the application of the FATE framework and rubric will help faculty engage in reflection and formative assessment of their own teaching resulting in a more intentional teaching approach. Using this framework and rubric will improve the robustness and enhance effectiveness of teaching evaluations and identify those who really are excellent teachers (Fink 2008). As teaching improves, learning will improve, and long-term student retention and graduation rates will increase (National Academies of Sciences, Engineering, and Medicine 2020; Benton and Young 2018; Miller and Seldin 2014). Providing evidence of these practices and resulting improvements through a portfolio and reflection process will make this work visible to others (Gibbons et al. 2018; Smith et al. 2014; Seldin 2000; Drinkwater, Matthews, and Seiler 2017; Zubizarreta 1999; Seldin, Miller, and Seldin 2010). This will lead academic units to embrace the framework and view the rubric as a valuable tool to for evaluating teaching effectiveness. This in turn will lead to institutional transformations as the importance of teaching is increased and the campus climate around teaching improves as the evaluation process is more transparent, we measure what is actually valued, and teaching efforts are appropriately accounted for in performance reviews.

The next steps are to adopt the framework, test the rubric, refine the process and collect data to determine the effect of implementing FATE.

Acknowledgement

The authors would like to thank the numerous teachers and those who support teaching at our institution who have provided valuable input and critique.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding details

This work was supported by the NSF under Grant DUE 1856653.

References

- Alton-Lee, Adrienne. 2003. *Quality Teaching for Diverse Students in Schooling: Best Evidence Synthesis*, edited by Ministry of Education. Wellington, New Zealand: New Zealand Ministry of Education. http://intra.bay.net.nz/learning_objects/datas/quality.pdf.
- Angelo, Thomas A., and K. Patricia Cross. 1993. *Classroom Assessment Techniques: A Handbook for College Teachers*. 2nd ed. San Francisco, CA: Jossey-Bass Publishers.
- Armbruster, Peter, Maya Patel, Erika Johnson, and Martha Weiss. 2009. "Active Learning and Student-Centered Pedagogy Improve Student Attitudes and Performance in Introductory Biology." *CBE Life Sciences Education* 8 (3): 203–13. doi: [10.1187/cbe.09-03-0025](https://doi.org/10.1187/cbe.09-03-0025).
- Arreola, Raoul. 2000. *Developing a Comprehensive Faculty Evaluation System: A Handbook for College Faculty and Administrators on Designing and Operating a Comprehensive Faculty Evaluation System*. 2nd ed. Bolton, MA: Anker Publishing Company.
- ASA. 2019. "Statement on Student Evaluations of Teaching." ASA.net.org: American Sociological Association.
- Bain, Ken. 2004. *What the Best College Teachers Do*. Cambridge, MA: Harvard University Press.
- Baird, John R. 1986. "Improving Learning through Enhanced Metacognition: A Classroom Study." *European Journal of Science Education* 8 (3):263–82. doi: [10.1080/0140528860080303](https://doi.org/10.1080/0140528860080303).
- Bali, Maha. 2018. "Assessing Teaching: A Curriculum Theory Approach." *ProfHacker*, Feb. 1.
- Barkley, Elizabeth F., and Claire Howell Major. 2016. *Learning Assessment Techniques: A Handbook for College Faculty*. Hoboken, NJ: John Wiley and Sons.
- Bavishi, A., J. M. Madera, and M. R. Hebl. 2010. "The Effect of Professor Ethnicity and Gender on Student Evaluations: Judged before Met." *Journal of Diversity in Higher Education* 3 (4):245–56. doi: [10.1037/a0020763](https://doi.org/10.1037/a0020763).
- Bell, Lee Anne, Diane J. Goodman, and Mathew Ouellett. 2016. "Design and Facilitation." In *Teaching for Diversity and Social Justice*, edited by Maurianne Adams and Lee Anne Bell, 3–93. New York, NY: Routledge, Taylor & Francis.

- Benton, Stephen L., and Suzanne Young. 2018. "Best Practices in the Evaluation of Teaching." In *IDEA Papers*. Manhattan, KS: IDEA.
- Berman, Elena. 2003. "A Short Guide to Evaluating Teaching." Tucson, AZ: University of Arizona. Accessed October 4, 2016. http://www.tltaskforce.uconn.edu/docs/resources/A_Short_Guide_to_Evaluating_Teaching.pdf.
- Bernstein, Daniel J., William Addison, Cindy Altman, Debra Hollister, Meera Komaraju, Loreto Prieto, Courtney A. Rocheleau, and Cecilia Shore. 2010. "Toward a Scientist-Educator Model of Teaching Psychology." In *Undergraduate Education in Psychology: A Blueprint for the Future of the Discipline*, edited by Diane F. Halpern, 29–45. Washington, DC: American Psychological Association.
- Boring, Anne, Kellie Ottoboni, and Philip B. Stark. 2016. "Student Evaluations of Teaching (Mostly) Do Not Measure Teaching Effectiveness." *ScienceOpen Research* 1–11. doi: [10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1](https://doi.org/10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1).
- Boyer, Ernest L. 1990. *Scholarship Reconsidered: Priorities of the Professoriate*. Princeton, NJ: The Carnegie Foundation for the Advancement of Teaching.
- Boysen, Guy A., Aaron S. Richmond, and Regan A. R. Gurung. 2015. "Model Teaching Criteria for Psychology: Initial Documentation of Teachers' Self-Reported Competency." *Scholarship of Teaching and Learning in Psychology* 1 (1):48–59. doi: [10.1037/stl0000023](https://doi.org/10.1037/stl0000023).
- Brancato, Vera C. 2003. "Professional Development in Higher Education." *New Directions for Adult and Continuing Education* 2003 (98):59–65. doi: [10.1002/ace.100](https://doi.org/10.1002/ace.100).
- Cavanagh, Sarah Rose. 2019. "Honeybees, Social Neuroscience, and Team Dynamics: Lessons for Building Community." POD Network Annual Conference, Pittsburgh, PA, November 17.
- Chew, Stephen L., and William J. Cerbin. 2017. "Teaching and Learning: Lost in a Buzzword Wasteland." *Teaching and Learning*, December 5. https://www.insidehighered.com/views/2017/12/05/need-theory-learning-opinion?utm_source=Inside+Higher+Ed&utm_campaign=df16c5da25-DNU20171205&utm_medium=email&utm_term=0_1fcb04421-df16c5da25-197786989&mc_cid=df16c5da25&mc_eid=2dae384bc1.
- Chickering, Arthur W., and Zelda F. Gamson. 1987. "Seven Principles for Good Practice in Undergraduate Education." *AAHE Bulletin* (March):2–6.
- Cornelius-White, Jeffrey. 2007. "Learner-Centered Teacher-Student Relationships Are Effective: A Meta-Analysis." *Review of Educational Research* 77 (1):113–43. doi: [10.3102/003465430298563](https://doi.org/10.3102/003465430298563).
- Crooks, Terry. 2001. "The Validity of Formative Assessments." British Educational Research Association Annual Conference, University of Leeds.
- Danielson, Charlotte. 2014. *The Framework for Teaching Evaluation Instrument*. 2013 ed., 105. Chicago, IL: The Danielson Group.
- De Rijdt, Catherine, Eva Tiquet, Filip Dochy, and Maurice Devolder. 2006. "Teaching Portfolios in Higher Education and Their Effects: An Explorative Study." *Teaching and Teacher Education* 22 (8):1084–93. doi: [10.1016/j.tate.2006.07.002](https://doi.org/10.1016/j.tate.2006.07.002).
- del Carmen Salazar, Maria, Amanda Stone Norton, and Franklin A. Tuitt. 2010. "Weaving Promising Practices for Inclusive Excellence into the Higher Education Classroom." *To Improve the Academy* 28 (1):208–26. doi: [10.1002/j.2334-4822.2010.tb00604.x](https://doi.org/10.1002/j.2334-4822.2010.tb00604.x).
- Drinkwater, Michael J., Kelly E. Matthews, and Jacob Seiler. 2017. "How Is Science Being Taught? Measuring Evidence-Based Teaching Practices across Undergraduate Science Departments." *CBE - Life Sciences Education* 16 (1):1–11. doi: [10.1187/cbe.15-12-0261](https://doi.org/10.1187/cbe.15-12-0261).
- Durham, Mary F., Jennifer K. Knight, Emily K. Bremers, Jameson D. DeFreece, Alex R. Paine, and Brian A. Couch. 2018. "Student, Instructor, and Observer Agreement regarding Frequencies of Scientific Teaching Practices Using the Measurement Instrument for Scientific Teaching-Observable (MISTO))." *International Journal of STEM Education* 5 (1):31. doi: [10.1186/s40594-018-0128-1](https://doi.org/10.1186/s40594-018-0128-1).
- Dweck, Carol S., and Ellen L. Leggett. 1988. "A Social-Cognitive Approach to Motivation and Personality." *Psychological Review* 95 (2):256–73. doi: [10.1037/0033-295X.95.2.256](https://doi.org/10.1037/0033-295X.95.2.256).
- Esarey, Justin, and Natalie Valdes. 2020. "Unbiased, Reliable, and Valid Student Evaluations Can Still Be Unfair." *Assessment and Evaluation in Higher Education* 45:1–15. doi: [10.1080/02602938.2020.1724875](https://doi.org/10.1080/02602938.2020.1724875).
- Feldman, Kenneth A. 1978. "Course Characteristics and College Students' Ratings of Their Teachers: What We Know and What We Don't." *Research in Higher Education* 9 (3):199–242. doi: [10.1007/BF00976997](https://doi.org/10.1007/BF00976997).
- Fink, L. Dee. 2003. *Creating Significant Learning Experiences: An Integrated Approach to Designing College Courses, Higher and Adult Education Series*. San Francisco, CA: Jossey-Bass.
- Fink, L. Dee. 2008. "Evaluating Teaching: A New Approach to an Old Problem." *To Improve the Academy* 26 (1): 3–21. doi: [10.1002/j.2334-4822.2008.tb00497.x](https://doi.org/10.1002/j.2334-4822.2008.tb00497.x).
- Flaherty, Colleen. 2015. *AAUP Committee Survey Data Raise Questions on Effectiveness of Student Teaching Evaluations*, 3. Washington, D.C.: Inside Higher Ed.
- Franklin, Jennifer. 2001. "Interpreting the Numbers: Using a Narrative to Help Others Read Student Evaluations of Your Teaching Accurately." *New Directions for Teaching and Learning* 2001 (87):85–100. doi: [10.1002/tl.10001](https://doi.org/10.1002/tl.10001).
- Freeman, Scott, Sarah L. Eddy, Miles McDonough, Michelle K. Smith, Nnadozie Okoroafor, Hannah Jordt, and Mary Pat Wenderoth. 2014. "Active Learning Increases Student Performance in Science, Engineering, and mathematics." *Proceedings of the National Academy of Sciences of the United States of America* 111 (23): 8410–5. doi: [10.1073/pnas.1319030111](https://doi.org/10.1073/pnas.1319030111).
- Gibbons, Rebecca E., Sachel M. Villafane, Marilyne Stains, Kristen L. Murphy, and Jeffrey R. Raker. 2018. "Beliefs about Learning and Enacted Instructional Practices: An Investigation in Postsecondary Chemistry Education." *Journal of Research in Science Teaching* 55 (8):1111–33. doi: [10.1002/tea.21444](https://doi.org/10.1002/tea.21444).
- Gray, Tara. 2016. "New Mexico State University Teaching Academy." New Mexico State University, Last Modified 2016, Accessed May 22. <https://teaching.nmsu.edu/teaching-effectiveness/>.

- Grinnell, Frederick. 2000. "The Practice of Science at the Edge of Knowledge." *FASEB News*, June 2–5.
- Gulikers, Judith T. M., Theo J. Bastiaens, and Paul A. Kirschner. 2004. "A Five-Dimensional Framework for Authentic Assessment." *Educational Technology Research and Development* 52 (3):67–86. doi: [10.1007/BF02504676](https://doi.org/10.1007/BF02504676).
- Gurney, Philip. 2007. "Five Factors for Effective Teaching." *New Zealand Journal of Teachers' Work* 4 (2):89–98.
- Karplus, Robert, and Herbert D. Thier. 1967. "A New Look at Elementary School Science." In *New Trends in Curriculum and Instruction*, edited by John U. Michaelis, 204. Chicago, IL: Rand McNally & Company.
- King, Helen. 2004. "Continuing Professional Development in Higher Education: What Do Academics Do?" *Planet* 13 (1):26–9. doi: [10.11120/plan.2004.00130026](https://doi.org/10.11120/plan.2004.00130026).
- Kornell, Nate, and Hannah Hausman. 2016. "Do the Best Teachers Get the Best Ratings." *Frontiers in Psychology* 7: 1–8. doi: [10.3389/fpsyg.2016.00570](https://doi.org/10.3389/fpsyg.2016.00570).
- Krathwohl, David R. 2002. "A Revision of Bloom's Taxonomy: An Overview." *Theory into Practice* 41 (4): 212–8. doi: [10.1207/s15430421tip4104_2](https://doi.org/10.1207/s15430421tip4104_2).
- Lewis, Scott E., and Jennifer E. Lewis. 2005. "Departing from Lectures: An Evaluation of a Peer-Led Guided Inquiry." *Journal of Chemical Education* 82 (1):135–9. doi: [10.1021/ed082p135](https://doi.org/10.1021/ed082p135).
- Lo, Stanley M., and Jonathan Mendez. 2019. "L: Learning—The Evidence." In *POGIL: An Introduction to Process Oriented Guided Inquiry Learning for Those Who Wish to Empower Learners*, edited by Shawn R. Simonson, 85–112. Sterling, VA: Stylus Publishing, LLC.
- Marchesani, Linda S., and Maurianne Adams. 1992. "Dynamics of Diversity in the Teaching-Learning Process: A Faculty Development Model for Analysis and Action." *New Directions for Teaching and Learning* 1992 (52):9–19. doi: [10.1002/tl.37219925203](https://doi.org/10.1002/tl.37219925203).
- Mazur, Eric. 2009. "Education. Farewell, Lecture?" *Science* 323 (5910):50–1. doi: [10.1126/science.1168927](https://doi.org/10.1126/science.1168927).
- McCabe, Kimberly A., and Leslie S. Layne. 2012. "The Role of Student Evaluations in Tenure and Promotion: Is Effective Teaching Really Being Measured?" *The Department Chair: A Resource for Academic Administrators* 22 (3):17–20.
- Merriam-Webster, Inc. 2018. *The Merriam-Webster Dictionary*. Springfield, MA: Merriam-Webster, Inc.
- Miller, Elizabeth, and Peter Seldin. 2014. "Changing Practices in Faculty Evaluation." In *Academe*. Washington, D.C.: American Association of University Professors.
- Mitchell, Kristina M. W., and Jonathan Martin. 2018. "Gender Bias in Student Evaluations." *PS: Political Science & Politics* 51 (03):648. doi: [10.1017/S1049096800001X](https://doi.org/10.1017/S1049096800001X).
- Mrig, Amit, Daniel Fusch, and Paul Cook. 2014. *The State of Professional Development in Higher ed*, 48. Higher Ed Impact.
- Myrsky, Liisa, Terhi Karaharju-Suvanto, Marjo Vesalainen, Anna-Maija Virtala, Marja Raekallio, Outi Salminen, Katariina Vuorensola, and Anne Nevgi. 2020. "Experienced Academics' Emotions Related to Assessment." *Assessment & Evaluation in Higher Education* 45 (1):1–13. doi: [10.1080/02602938.2019.1601158](https://doi.org/10.1080/02602938.2019.1601158).
- National Academies of Sciences, Engineering, and Medicine. 2020. "Recognizing and Evaluating Science Teaching in Higher Education: Proceedings of a Workshop in Brief." Edited by Susan J. Debad. Washington, DC: The National Academies Press. Accessed January 29. <http://nap.edu/25685>.
- National Research Council. 2012. "Discipline-Based Educational Research: Understanding and Improving Learning in Undergraduate Science and Engineering." Edited by Susan R. Singer, Natalie R. Nielsen, and Heidi A. Schweingruber. Washington, DC: The National Academies Press. Accessed January 29. <http://nap.edu/13362>.
- North, Joan DeGuire. 1999. "Administrative Courage to Evaluate the Complexities of Teaching." In *Changing Practices in Evaluating Teaching*, edited by Peter Seldin, 183–93. Bolton, MA: Anker Publishing Company, Inc.
- Pallas, Aaron M., Anna Neumann, and Corbin M. Campbell. 2017. *Policies and Practices to Support Undergraduate Teaching Improvement*. Cambridge, MA: American Academy of Arts and Sciences.
- Palmer, Parker J. 2017. *The Courage to Teach: Exploring the Inner Landscape of a Teacher's Life*. 20th ed. San Francisco, CA: Jossey-Bass.
- Pirsig, Robert M. 1974. *Zen and the Art of Motorcycle Maintenance: An Inquiry into Values*. New York: William Morrow and Company.
- Prince, Michael J. 2004. "Does Active Learning Work? A Review of the Research." *Journal of Engineering Education* 93 (3):223–31. doi: [10.1002/j.2168-9830.2004.tb00809.x](https://doi.org/10.1002/j.2168-9830.2004.tb00809.x).
- Richlin, Laurie. 2001. "Scholarly Teaching and the Scholarship of Teaching." *New Directions for Teaching and Learning* 2001 (86):57–68. doi: [10.1002/tl.16](https://doi.org/10.1002/tl.16).
- Richmond, Aaron S., Guy A. Boysen, Regan A. R. Gurung, Yvette N. Tazeau, Steven A. Meyers, and Mark J. Sciutto. 2014. "Aspirational Model Teaching Criteria for Psychology." *Teaching of Psychology* 41 (4):281–95. doi: [10.1177/0098628314549699](https://doi.org/10.1177/0098628314549699).
- Robinson, Terrell E., and Warren C. Hope. 2013. "Teaching in Higher Education: Is There a Need for Training in Pedagogy in Graduate Degree Programs?" *Research in Higher Education* 21:11.
- Schmidt, Benjamin M. 2015. "Gendered Language in Teaching Reviews." Ben Schmidt. http://benschmidt.org/profGender/#%7B%22database%22%3A%22RMP%22%2C%22plotType%22%3A%22pointchart%22%2C%22method%22%3A%22return_json%22%2C%22search_limits%22%3A%7B%22word%22%3A%5B%22funny%22%5D%2C%22department_id%22%3A%7B%22%24lte%22%3A25%7D%7D%2C%22aesthetic%22%3A%7B%22x%22%3A%22WordsPer
- Schraw, Gregory. 1998. "Promoting General Metacognitive Awareness." *Instructional Science* 26 (1/2):113–25. doi: [10.1023/A:1003044231033](https://doi.org/10.1023/A:1003044231033).
- Seldin, Peter. 1999a. "Building Successful Teaching Evaluation Programs." In *Changing Practices in Evaluating Teaching*, edited by Peter Seldin, 213–42. Bolton, MA: Anker Publishing Company, Inc.
- Seldin, Peter. 1999b. "Current Practices—Good and Bad—Nationally." In *Changing Practices in Evaluating*

- Teaching*, edited by Peter Seldin, 1–24. Bolton, MA: Anker Publishing Company, Inc.
- Seldin, Peter. 2000. "Teaching Portfolios: A Positive Appraisal." *Academe* 86 (1):36–44. doi: [10.2307/40252334](https://doi.org/10.2307/40252334).
- Seldin, Peter, J. Elizabeth Miller, and Clement A. Seldin. 2010. *The Teaching Portfolio: A Practical Guide to Improved Performance and Promotion/Tenure Decisions, the Jossey-Bass Higher and Adult Education Series*. San Francisco, CA: Jossey-Bass.
- Shadle, Susan E., Anthony Marker, and Brittnee Earl. 2017. "Faculty Drivers and Barriers: Laying the Groundwork for Undergraduate STEM Education Reform in Academic Departments." *International Journal of STEM Education* 4 (1):8. doi: [10.1186/s40594-017-0062-7](https://doi.org/10.1186/s40594-017-0062-7).
- Smith, Mark K. 1996, 2000. *Curriculum Theory and Practice. The Encyclopedia of Informal Education*. London: infed.org. Accessed February 1, 2018.
- Smith, Michelle K., Erin L. Vinson, Jeremy A. Smith, Justin D. Lewin, and MacKenzie R. Stetzer. 2014. "A Campus-Wide Study of STEM Courses: New Perspectives on Teaching Practices and Perceptions." *CBE Life Sciences Education* 13 (4):624–35. doi: [10.1187/cbe.14-06-0108](https://doi.org/10.1187/cbe.14-06-0108).
- Stains, Marilyne, Jordan Harshman, Megan K. Barker, Stephanie V. Chasteen, Renee Cole, Sue Ellen DeChenne-Peters, M. Kevin Eagan, Jr., Joan M. Esson, Jennifer K. Knight, Frank A. Laski, et al. 2018. "Anatomy of STEM Teaching in North American Universities." *Science* 359 (6383):1468–70. doi: [10.1126/science.aap8892](https://doi.org/10.1126/science.aap8892).
- Stupnisky, Robert H., Allison BrckaLorenz, Bridget Yuhas, and Frederic Guay. 2018. "Faculty Members' Motivation for Teaching and Best Practices: Testing a Model Based on Self-Determination Theory across Institution Types." *Contemporary Educational Psychology* 53:15–26. doi: [10.1016/j.cedpsych.2018.01.004](https://doi.org/10.1016/j.cedpsych.2018.01.004).
- Tanner, Kimberly D. 2010. "Order Matters: Using the 5E Model to Align Teaching with How People Learn." *CBE Life Sciences Education* 9 (3):159–64. doi: [10.1187/cbe.10-06-0082](https://doi.org/10.1187/cbe.10-06-0082).
- Tanner, Kimberly D., and Deborah Allen. 2005. "Approaches to Biology Teaching and Learning: Understanding the Wrong Answers-Teaching toward Conceptual Change." *Cell Biology Education* 4 (2):112–7. doi: [10.1187/cbe.05-02-0068](https://doi.org/10.1187/cbe.05-02-0068).
- Teaching Quality Framework Initiative, Center for STEM Learning. 2017. "Teaching Quality Framework Initiative." University of Colorado, Last Modified 2017, Accessed May 22. <https://www.colorado.edu/teaching-quality-framework/about-tqf>.
- Teoh, Siew Li, Long Chiau Ming, and Tahir Mehmood Khan. 2016. "Faculty Perceived Barriers and Attitudes toward Peer Review of Classroom Teaching in Higher Education Settings: A Meta-Cynthesis." *Sage Open* 6 (3):215824401665808. doi: [10.1177/2158244016658085](https://doi.org/10.1177/2158244016658085).
- The Commission on the Future of Undergraduate Education. 2017. *The Future of Undergraduate Education, The Future of America*. Cambridge, MA: American Academy of Arts and Sciences.
- Tolman, Anton O., and Janine Kremling, eds. 2017. *Why Students Resist Learning: A Practical Model for Understanding and Helping Students*. Sterling, VA: Stylus Publishing, LLC.
- Trigwell, Keith, Michael Prosser, and Fiona Waterhouse. 1999. "Relations between Teachers' Approaches to Teaching and Students' Approaches to Learning." *Higher Education* 37 (1):57–70. doi: [10.1023/A:1003548313194](https://doi.org/10.1023/A:1003548313194).
- Vygotsky, Lev S. 1930. "Mastery of Memory and Thinking." In *Mind and Society*, edited by Andy Blunden and Nate Schmolze, 25–41. Cambridge, MA: Harvard University Press.
- Wieman, Carl E. 2019. "Expertise in University Teaching and the Implications for Teaching Effectiveness, Evaluation, and Training." *Daedalus Fall* 148 (4):47–78. doi: [10.1162/daed_a_01760](https://doi.org/10.1162/daed_a_01760).
- Wiggins, Grant. 2012. "Seven Keys to Effective Feedback." *Educational Leadership* 70 (1):10–16.
- Wiggins, Grant, and Jay McTighe. 2005. *Understanding by Design*. 2nd ed. Alexandria, VA: Association for Supervision and Curriculum Development.
- Zubizarreta, John. 1999. "Evaluating Teaching through Portfolios." In *Changing Practices in Evaluating Teaching*, edited by Peter Seldin, 162–82. Bolton, MA: Anker Publishing Company, Inc.