Evaluating teaching in a scholarly manner:
A model and call for an evidence-based, departmentally-defined approach to enhance teaching evaluation for CU Boulder

Noah Finkelstein, Joel C. Corbo, Daniel L. Reinholz, Mark Gammon, and Jessica Keating

Abstract:

The University of Colorado Boulder requires that “[d]ossiers for comprehensive review, tenure, or promotion must include multiple measures of teaching” (Guidelines, 2007). However, at present we do not have a well-documented framework to guide individuals or departments in the selection and interpretation of such measures, which makes it difficult to assess teaching quality and support systemic faculty growth in teaching. In this white paper, we outline a framework for supporting and assessing teaching quality that is grounded in the scholarship of higher education. Such a framework will advance individual educational efforts as well as support the alignment of campus resources to enhance education.
Introduction

This white paper calls for an evidence-based, departmentally-defined approach to enhancing teaching evaluation at CU Boulder. It outlines a framework for supporting and assessing teaching quality that can be developed with all instructors across all departments on campus, aligning resources and enhancing educational practices. The framework brings together two essential elements: a structure of scholarly categories of educational practice for evaluation and a strategic process for enacting a scholarly approach to evaluation that draws from models of institutional change. This approach is grounded in the scholarship of higher education, including the work of Bernstein (2002, 2010), Glassick (1997), and colleagues. This framework defines teaching as a scholarly activity analogous to research, assessing teaching in terms of six core components of scholarly activity—clear goals, adequate preparation, appropriate methods, significant results, effective presentation, and reflective critique. It uses the perspectives of faculty members, their students, and their peers as key sources of data. While the framework categories are held constant across departments, the interpretation of these categories and their relative weights are defined at a department-by-department level, thus providing the university with a common approach to assessment, while preserving disciplinary identity and specificity. In addition to allowing for consistent and transparent assessment, the framework also supports improved teaching by providing mechanisms for feedback to help faculty improve as teachers.

This white paper also provides a strategy for implementation that is grounded in theories of organizational change and is designed to lead to campus-wide adoption. It is not a top-down mandate. Instead, it focuses on bringing together key faculty leaders and departments and providing them with a structure to co-create, test, and evaluate the framework in the context that makes the most sense to them. Simultaneously, this community designs the processes of implementation and identifies the resources necessary to succeed. It is an opt-in model, with pilot departments choosing to engage and become leaders in this process. Thus, this strategy empowers the community to voluntarily engage in the exploration of new ways of assessing and supporting teaching and to adopt the framework because they see its value.

Why do we need a Teaching Quality Framework?

Whether on the local or national level, there are increasing calls for attention to the quality of teaching in higher education as a means to improve student outcomes, including increased retention, graduation rates, and diversity (e.g., President's Council of Advisors on Science and Technology, PCAST, 2012; Seymour & Hewitt, 1997). These calls also align with the needs of faculty, who are being called upon to develop and assess their teaching in more robust ways (e.g., Berlinerblau, 2017; Davidson, 2017). Fortunately, given decades of research on scholarly and professional approaches to teaching (e.g., Beach et al, 2016; Boyer, 1990; Fairweather, 2002), the growth of work within the disciplines on more effective educational approaches (PCAST, 2012; National Research Council, 2012), models for institutional and organizational change applied to higher education (e.g., Corbo et al., 2016; Kezar, 2013; Reinholz et al., in press), and access to new forms of data on educational practices, higher
education is now positioned to address these calls in a scholarly and evidence-based manner. Nevertheless, there are still numerous challenges to improving teaching quality, which are often created by the ways in which teaching is evaluated. A framework for assessing teaching quality that is shared across a campus but contextualized on a department-by-department basis can help mitigate these challenges and serve both needs, improving student outcomes by supporting faculty in their growth as teachers.

**What is a Teaching Quality Framework?**

The Teaching Quality Framework (TQF) is defined as consisting of two elements:

1. a campus-wide, cross-disciplinary *structure* for defining teaching quality and associated tools for assessing teaching quality, and
2. a *process* for contextualizing the structure to each disciplinary unit and enacting it across campus.

Thus, the TQF provides a common approach to enhanced teaching evaluation that is usable across a campus while supporting disciplinary specificity and variation. The TQF *structure* draws from several decades of research on the evaluation of teaching as a scholarly activity (see Boyer 1990, and its derivatives such as Bernstein and colleagues, 2002, 2010, and Glassick and colleagues, 1997). The TQF *process* builds upon recent work in institutional and organizational change in higher education to create holistic, sustainable, and scalable transformations (Corbo et al, 2016; Kezar, 2013; Reinholz et al., 2015).

There are several goals that guided the development of the TQF structure and process. One critical goal is to create a method for evaluating teaching that is both summative—for example, in evaluating tenure and promotion (T&P) cases—and formative—to support faculty reflection on their teaching and to align institutional resources to such end. Another goal is to more efficiently use the resources (mostly time) that are already allocated to teaching evaluation. For example, the TQF focuses on gathering data about faculty members’ teaching from their peers, their students, and themselves, just as in our current system. However, it only uses each source to provide feedback that aligns with what that perspective is uniquely positioned to assess. Thus, the development and rollout of the TQF will require an initial investment of time and work but may not require additional resources in the long run. A third goal is to leverage the collective efforts of multiple departments to create a shared pool of resources. As a campus implements a TQF, it will develop a bank of evaluation tools generated by pilot departments. Eventually, other departments will be able to use these starting points for developing their own teaching assessment strategies without having to repeat work. Finally, the most important goal of the TQF is to improve undergraduate education by providing faculty members with incentives and support to become better teachers. By establishing a discipline-specified set of criteria, this process externalizes the value-system and goals of education which then serve as an anchor around which the substantial institutional resources to enhance teaching can align.
Teaching Quality Framework Structure

The TQF defines teaching as a scholarly activity, first by establishing essential components and assessing them via different perspectives (or sources of data): those of the instructors, their students, and their peers. Just as in the evaluation of research effectiveness, it is the responsibility of individual faculty members to make the case for their educational effectiveness using the evidence provided by these perspectives. A proposed approach for constructing these evaluations is to have faculty members create portfolios that address each of the components, with a focus on their growth over time. This approach is consistent with research on faculty growth and adult learning through reflection (e.g., Boud & Walker, 1991; O’Meara, Terosky, & Neumann, 2008). The portfolios could also be shared with other faculty, to serve as examples of reflective practice and to encourage collective growth.

A Sample Structure: Six Components of Scholarly Activity

Ernest Boyer’s publication, Scholarship Reconsidered: Priorities of the Professoriate (1990), significantly broadened the understanding of academic scholarship. In his work, Boyer defines four types of scholarship, including the scholarship of teaching. One of Boyer’s major arguments is that all types of scholarship share much in common, including methods of assessment. Therefore, one can assess the scholarship of teaching using many of the same approaches that one would use to assess other types of scholarship (such as the scholarship of discovery). Subsequently, Scholarship Assessed: Evaluation of the Professoriate (Glassick, Huber, & Maeroff, 1997) operationalized the assessment of all forms of scholarship in terms of six components. These components (adapted from Scholarship Assessed) are illustrated by questions relevant to the scholarship of teaching:

1. Clear goals: Does the instructor state the goals of the course/learning experience clearly? Are these goals realistic and achievable? Do they relate to important questions in the relevant field of study?
2. Adequate preparation: Does the instructor have an understanding of the scholarship of teaching and learning in his or her field? Has he or she practiced the necessary skills and gathered the necessary resources to allow for successful learning?
3. Appropriate methods: Does the instructor choose teaching methods appropriate to achieve the learning goals, and does he or she apply them effectively? Does the instructor modify these methods in response to changing circumstances in the classroom?
4. Significant results: Does the instructor achieve his or her goals? Do students show evidence of learning and growth? Does the instructor’s work in the classroom add to the knowledge of teaching in his or her field or open up new areas for exploration?
5. Effective presentation: Does the instructor externalize and communicate the results of his or her teaching to peers using suitable style, effective organization, appropriate forums, and clarity and integrity?
6. Reflective critique: Does the instructor critically evaluate his or her teaching, using an appropriate breadth of evidence? Does the instructor use this evaluation to improve their teaching?
Each of these six components is described in more detail in *Scholarship Assessed*. Note that these are not the only components one could use to construct a TQF; other categorizations with more and fewer components exist. Nevertheless, there are strong overlaps among all such categorizations. Moreover, the questions above illustrate these components in the context of a classroom, the most common educational setting. However, they can be easily applied to other educational contexts, including the development of new curricula, new courses, and innovative classroom materials; the supervision of independent studies; the mentoring of undergraduates; and the supervision of internships or field work.

These six components have been further operationalized by others. For example, Bernstein and his co-authors (2010) added clarity in assessing these six components by creating a rubric that specified four levels of accomplishment for each: entry into teaching, basic skill, professional, and advanced (see Appendix A, pg. 2). The clear articulation of these levels makes both formative feedback and growth more likely and provides scholars with a roadmap to excel in the professional practice of teaching.

To accurately capture evidence of scholarly teaching, it is crucial to use multiple data sources (or “perspectives”). This use of multiple perspectives is consistent with standards that already exist at many universities and with the evaluation policies of the American Association of University Professors that propose using “various measures of the effectiveness of [teaching] efforts” (AAUP, 2006). While these “various measures” are often poorly-defined, assessments of faculty members’ teaching typically rely on three major sources of data: (a) students, (b) faculty peers (internal or external to the specific department), and (c) the instructors themselves. Each of these perspectives illuminates specific aspects of the faculty members’ teaching with respect to the components discussed above. Students are uniquely positioned to report on their perception of instructional time (component 3). However, given that students are still novices in their field of study, they are insufficiently qualified to comment on the instructor’s knowledge of the domain or choice of topics to include in the course (component 1); faculty peers, however, are qualified to do so. Furthermore, faculty peers who are knowledgeable in education research can also comment on whether the chosen methods of instruction are appropriate and up-to-date (components 2 and 3). Finally, the instructor who is being evaluated has privileged information about his or her goals, design process, professional development, observations of student progress, and engagement with the scholarship of teaching, all of which are relevant in making an appropriate assessment of teaching quality (all 6 components).

It is also expected that there will be differentiation across disciplines in how each of these perspectives contributes to the overall evaluation of teaching effectiveness. Finally, it is worth reiterating that a large amount of faculty and student time is already allocated to the evaluation of teaching. Thus, a key goal of the TQF is to make the evaluation of teaching more straightforward, efficient, and useful.
Teaching Quality Framework Process

The proposed strategy for implementation is the second element of the Teaching Quality Framework. The process is built on a gradual, opt-in strategy that allows for as much ownership of the framework by the faculty as possible. This element of the TQF draws from the acknowledgement that knowing what to do is insufficient, and establishing a strategic process for institutional change is as essential as providing tools and goals (AAU, 2014; Corbo et al., 2016; Kezar, 2013; Reinholz et al., in press). We note that a pilot version of this process is currently underway at University of Colorado Boulder, supported by CU and NSF.

The strategic process element of the TQF draws from particular models of institutional change that are applied to the relatively unique societal enterprise of higher education systems (AAU, 2014; Corbo et al., 2016; Kezar, 2013). The TQF draws from studies (Association of American Colleges and Universities, 2014) that identify the departmental unit as the sustained unit of educational change in larger research universities.

The process proposed here involves working across levels of the institution simultaneously, creating communications mechanisms, and honoring the independence of academic units. At a campus-wide level the process includes a campus-wide taskforce with the charge of developing and localizing a framework based on a bank of example materials from other institutions, such as the six components from Scholarship Assessed described above. In parallel, individual academic units opt in and contextualize the framework for their own uses, with what are referred to as rubrics and the associated resources needed to implement enhanced measures to teaching. It will be up to departments to delineate their best application of their rubrics. A third layer of process is to convene cross-departmental working teams to share materials and processes for contextualizing the components. As departments document the promise of this approach, administrative proclamations of support for the framework and for pilot departments, raising them up as exemplars for others to follow, will provide incentive, recognition, and institutional cover for participating departments. Finally, a centralized resource is necessary to manage this process, facilitate the department and cross-institutional conversations, archive resources, and advance the effort within and beyond the campus.

Campus-wide Taskforce

The first step is to develop a task-force of stakeholders endorsed by the provost, to create a framework (for example, working from the one shared above) for teaching quality. This framework needs to be designed for the local context, and it should have enough specificity so as to be understandable and applicable to individual departments, while having enough flexibility to allow for appropriate variation across campus (according to the ‘cultural’ perspective). The categories of the framework described above, along with examples from other institutions, should serve as a starting point. In addition to establishing a prototype framework, the task-force members consider the processes for the support and adoption of the newly-specified multiple measures of teaching. This approach might include coordinating with faculty governance, working with and educating the college and campus-level evaluation committees, and
establishing a communication strategy. Ultimately, creating centralized resources for evaluation, sharing of materials, and showcasing effective departmental-strategies will help institutionalize the use of quality measures of teaching. This cross-campus team can surface the various elements that need be involved and addressed when considering holistic, sustained change.

To succeed in its mission, the campus-wide taskforce needs appropriate support. At a minimum, this requires administrative sanction and support. Endorsement from the highest levels, usually the provost or chancellor/president, will convey the importance of this activity. The first essential resource is personnel time for someone to facilitate the process. Given the number of departments involved, a full-time staff person who is dedicated to facilitating this process may be essential. Key attributes of the facilitator are an understanding of institutional change, scholarly approaches to teaching, organizational skills, and sensitivity to different disciplinary cultures.

**Departmental Teams**

In parallel to and in conjunction with the task-force that creates a framework, individual departments contextualize the framework for their own use. This process begins with a goal of 5-15 departments self-identifying (i.e. opting in) for participation in the pilot version. These departments serve as exemplars to help other departments successfully follow their lead. Of course, as this occurs in parallel with the campus task-force, the framework and how it is contextualized will be iterative, and both the framework and the rubrics within departments should be living and evolving documents.

In each pilot department, a Departmental Action Team (DAT; Reinholz et al., in press)—a new type of faculty team that the authors have developed to support departmental change—are established. This DAT, a team of 3 to 6 faculty members, focuses on defining localized measures of teaching quality (i.e. turning the framework into a departmentally-specific rubric). It directs efforts within its own department and communicates progress with groups outside of the department. The primary roles of the departmental committee are to contextualize the framework to their units and help develop standards and processes for use. This process may involve mapping the three perspectives identified above to the six components of scholarly activity previously identified, or to any framework that is developed by the campus-wide committee. Having contextualized the framework, the committee may also benefit from having some of its members actually create portfolios and engage in a pilot version of the peer review process. These are suggested possibilities, but it will ultimately be the role of the departmental committee to decide how they proceed.

There should also be incentives for departments to engage in this process. Rather than trying to define these incentives from the outset, it is more productive for departments to have agency in defining what resources they may need. This allows the process to be flexible with respect to the specifics of the departments who are engaged in it. Initial incentives for membership on the task-force may be assignment of service duties and the opportunity to participate in a community with like-minded peers and administrative sanction. In our experience, the norm of providing food at all meetings helps build community.
Cross-Departmental Meetings: Intersection of the Teams

To support these teams, a working group that combines representatives from faculty member teams from each of the initial pilot departments is also formed. This working group would resemble a mix of a Faculty Learning Community (FLC; Cox, 2004) and a Departmental Action Team (Reinholz et al., in press). This group of 10 to 28 faculty works together to figure out how to roll out the framework in their departments, to compare strategies across departments, and to support each other through successes and setbacks. In parallel, the deans and administrative committees responsible for review of tenure and promotion cases are included to advise and frame the productive application of this framework.

Institutional Support: One way to support broader participation in the use of the framework is by having a public proclamation from the administration in support of departments doing so. This proclamation ought focus on a celebration of the good work that the pilot departments are doing, to show that this work adds value to the life of the departments, rather than as a mandate for engagement. In parallel, listening tours and eliciting public input about the teaching quality framework are essential throughout the process.

Centralized Resources: In addition to the support structures for each of the layers (task-force, departmental, cross-departmental), which largely amounts to dedicated staff time, centralized resources are necessary. While not substantial, especially when compared to the potential outcomes, having the dedicated time of individuals and modest funding for materials and food provides key resources for advancing the project. This modest allocation of resources also signals administrative support and sanction.

A Key Opportunity and Next Steps

With initial support from the Association of American Universities, the Bay View Alliance, and the National Science Foundation, the authors along with others from the Center for STEM Learning, the Office of Information Technology, and campus administrative units have launched a pilot implementation of this Teaching Quality Framework at the University of Colorado Boulder. Pilot work is underway with 11 departments across the College of Arts and Sciences and the College of Engineering and Applied Sciences to demonstrate the viability of this approach.

For more information on this project, see our website: https://www.colorado.edu/teaching-quality-framework/
References


Appendix A

TEACHING QUALITY FRAMEWORK (TQF)

The University of Colorado Boulder requires that "[d]ossiers for comprehensive review, tenure, or promotion must include multiple measures of teaching." (Guidelines 2007) However, at present we do not have a well-defined framework to guide individuals or departments in the selection and interpretation of such measures, which makes it difficult to assess teaching quality and support systemic faculty growth in teaching. In this project, we outline a framework for supporting and assessing teaching quality for all instructors across all departments on campus that is grounded in the scholarship of higher education. Such a framework will advance individual educational efforts as well as support the alignment of campus resources to enhance education.

The Framework

The goal of the framework is to support improved teaching by providing faculty members with feedback that they can use to improve as educators and to provide better mechanisms for assessing teaching quality for tenure, promotion, and merit.

This framework defines teaching as a scholarly activity (like research) and assesses core components of such scholarship. One example we draw from is: Glussick et al., Scholarship Assessed: Evaluation of the Professoriae, ’97 (more on Fg 2).

1. clear goals,
2. adequate preparation,
3. appropriate methods,
4. significant results,
5. effective presentation, and
6. reflective critique.

This assessment of framework criteria is made through the use of the three standard "voices" (data sources): the faculty member, the students, and peers. These framework categories are held constant across all departments; however, the definition and interpretation of these components of the framework (making them specific) and their relative weights would be defined at the unit level. Thus departments specify in a clear way what is meant by "multiple measures" locally, but using common categories across campus. This approach provides the university with a common framework while preserving disciplinary identity and specificity.

The Process

The implementation of the TQF that is not a top-down mandate, but instead focuses on bringing together key faculty leaders and departments and providing them with a structure to help them co-create, test, and evaluate the framework. This is an opt-in model, with pilot departments choosing to engage and become leaders in this process. Thus, this strategy empowers the community to voluntarily engage in the exploration of new ways of assessing teaching and to adopt the framework because they see its value.

Departmental TQF Teams:
- 9 Departments in A&S and CEAS committed
- 3-4 leads in each department
- Tasked with contextualizing the elements of the framework to the discipline and deciding what resources and process are required for implementation in their department.
- Following a Dept. Action Team model.

Campus / Cross-Unit TQF Dialog:
- Wide participation (departmental representatives, deans, VC-level & other key stakeholders)
- Defining the TQF & including changes proposed by the departmental teams.
- Communicate with campus T&F committees, non-pilot departments, etc.

A facilitator will support multiple departmental TQF teams and act as a communication channel across the departmental teams. Additionally, we expect the departmental teams to generate lists of required resources necessary to make the implementation of the TQF feasible given limitations on faculty time.

Contacts: Noah Finkelstein, Jessica Keating, Joel Corbo, & Mark Gammon (CU), Dariel Reinholz (SDSU), Daniel Bernstein (Univ. of Kansas) Center for STEM Learning www.colorado.edu/csi

University of Colorado Boulder
Ernest Boyer’s publication, *Scholarship Reconsidered: Priorities of the Professorate* (1990), has played a key role in broadening the perception of academic scholarship. Boyer defines four types of scholarship, including the scholarship of teaching. Subsequent work, *Scholarship Assessed: Evaluation of the Professorate* (Glassick, Huber, & Maeroff, 1997), has made great strides to operationalize the assessment of all forms of scholarship in terms of six components. Adapted from *Scholarship Assessed*.

1. **Clear goals**: Does the instructor state the goals of the course/learning experience clearly? Are these goals realistic and achievable? Do they relate to important questions in the relevant field of study?
2. **Adequate preparation**: Does the instructor have an understanding of the scholarship of teaching and learning in the field? Has he or she practiced the necessary skills and gathered the necessary resources to allow for successful learning?
3. **Appropriate methods**: Does the instructor choose teaching methods appropriate to achieve the learning goals, and do he or she apply them effectively? Does the instructor modify these methods in response to changing circumstances in the classroom?
4. **Significant results**: Does the instructor achieve his or her goals? Does the instructor’s work in the classroom add consequentially to the knowledge of teaching in his or her field or open up new areas for exploration?
5. **Effective presentation**: Does the instructor communicate with his or her students using suitable style, effective organization, appropriate forums, and clarity and integrity? Does the instructor communicate the results of his or her teaching to peers using the same set of criteria?
6. **Reflective critique**: Does the instructor critically evaluate his or her teaching, using an appropriate breadth of evidence? Does the instructor use this evaluation to improve the quality of future work?

Each of these six components is elaborated in more detail in *Scholarship Assessed* and has been further operationalized by others (Bernstein et al., 2010). A sample rubric is below; though many exist to draw from.

**Figure 2: Rubric for Assessing Teaching as Scholarly Activity (from Bernstein)**

<table>
<thead>
<tr>
<th>Components</th>
<th>Entry into teaching</th>
<th>Basic Skill</th>
<th>Professional</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goals of the course or other learning activity</td>
<td>Course/activity goals are absent, unclear, or inappropriate.</td>
<td>Course/activity goals are well articulated and appropriate to the course and to the curriculum.</td>
<td>Course/activity goals identify intellectually challenging and enduring targets and/or are especially well matched to students.</td>
<td>Course/activity goals identify levels of performance that represent excellence and are of interest to many stakeholders.</td>
</tr>
<tr>
<td>Preparation for the course or learning activity</td>
<td>Teacher is not adequately knowledgeable and/or has no background in teaching.</td>
<td>The teaching is based on prior scholarship in its area, including current content as well as pedagogical methods and conceptual frames.</td>
<td>The teacher’s preparation includes broad synthesis of prior work in content as well as practice in pedagogical methods and conceptual frames.</td>
<td>The teacher acquires and integrates knowledge and skills drawn from the literature of multiple disciplines, both in content and pedagogy.</td>
</tr>
<tr>
<td>Methods used to conduct the teaching</td>
<td>No apparent rationale for teaching methods is used; there is no instructional design.</td>
<td>The work follows the conventions of teaching practices within its domain of discipline and institution.</td>
<td>The teaching takes full advantage of effective methods discussed within its discipline.</td>
<td>The work generates new practices that will enable others to improve or enhance their teaching.</td>
</tr>
<tr>
<td>Evidence gathered to demonstrate the impact of the teacher’s work</td>
<td>There is no measure of student learning, or assessment method does not match expressed goals.</td>
<td>There is evidence linking students’ performance to expressed goals.</td>
<td>Student performances indicate that deep and/or broad learning is taking place.</td>
<td>The learning demonstrated is exemplary in either depth of learning and/or in breadth of student success.</td>
</tr>
<tr>
<td>Communication of teaching results to others</td>
<td>The practice and results of teaching are kept private.</td>
<td>The teacher’s work and students’ performance are publicly accessible for others to use, to build on, and to review critically.</td>
<td>The teacher’s reflective work has been read and adjustments in practice have arisen through the public discourse.</td>
<td>The teacher’s work has had an impact on the practices and inquiry of many others and has contributed to related conceptual frameworks.</td>
</tr>
<tr>
<td>Reflection on the teaching and its impact on student learning</td>
<td>The teacher provides no indication of having reflected on or learned from prior teaching.</td>
<td>The teacher articulates lessons learned from reflecting on prior teaching.</td>
<td>The teacher has examined the impact on students’ performance within a conceptual framework and adjusted practices based on reflection.</td>
<td>Enhanced achievement of learning goals results from reflection on evidence within a conceptual framework, or the teacher revises the conceptual framework based on student learning outcomes.</td>
</tr>
</tbody>
</table>