

TOWARDS A FRAMEWORK FOR SUPPORTING AND ASSESSING TEACHING QUALITY AT CU-BOULDER

[The University of Colorado Boulder should] enhance efforts to upgrade the prestige, respect and reward structure for excellence in the scholarship of teaching; . . . Develop frameworks in which teaching excellence and dedication are evaluated with a level of scrutiny comparable to how research and creative work is scrutinized.

*Recommendation 7
Academic Affairs Persistence Committee
Co-Chairs: M. Grant & J. Cox*

Executive Summary

The University of Colorado Boulder requires that “[d]ossiers for comprehensive review, tenure, or promotion must include multiple measures of teaching” (J. Cox, 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 faculty growth in their teaching in a systematic way.

This paper outlines 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, including the work of Bernstein and colleagues (2002, 2010) and Glassick and colleagues (1997). This framework defines teaching as a scholarly activity like research. It assesses teaching in terms of six core components of scholarly activity—clear goals, adequate preparation, appropriate methods, significant results, effective presentation, and reflective critique—through the use of three “voices” —those of a faculty member, his or her students, and his or her peers. The framework also supports improved teaching, by providing mechanisms for assessment to help faculty to improve in their practices. These framework categories are held constant across all departments; however, the interpretation of these categories and their relative weights would be defined at a department-by-department level, thus specifying in a clear way what is meant by “multiple measures.” This would provide the university with a common approach to assessment while preserving disciplinary identity and specificity.

In addition to presenting this framework, we suggest a strategy for implementation that will lead to its campus-wide adoption. This strategy is not a top-down mandate. Instead, it 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 in a relatively low-stakes context (merit raises, rather than tenure and promotion). 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. Finally, we present a set of examples of ways in which teaching can be assessed that aligns with the framework in order to provide context to the reader.

1. Why do we need a Teaching Quality Framework?

As suggested by the recommendation quoted at the beginning of this document, a renewed focus on teaching excellence is timely given that improved teaching has the capacity to support many pressing campus initiatives. For instance, the CU administration has made numerous calls to improve the persistence and retention of students, and student success is a leading goal for the campus. This aligns with the needs of faculty, who have a growing interest in being able to develop and assess their teaching in more robust ways. This focus also aligns with national calls, such as those to increase retention of students in STEM (science, technology, engineering, and mathematics; PCAST, 2012). The interactions that students have with their instructors are a major factor in whether or not they persist in college, so an emphasis on teaching excellence is fundamental to improving persistence (Seymour & Hewitt, 1997). At present, the lack of a clear definition of multiple measures of teaching effectiveness makes it difficult to adequately assess and support improved teaching. Thus, a Teaching Quality Framework can improve student persistence by supporting faculty in their growth as teachers. As a part of our framework, we provide tools to help improve the assessment of teaching as a learning process.

Beyond supporting persistence efforts, defining teaching excellence has a number of other benefits for different campus constituencies:

- Tenure-Track Faculty & Instructors: By looking at faculty growth over time, rather than just absolute achievement, it is possible to promote faculty learning, rather than just a one-shot evaluation.
- Contingent Faculty: The peer review and reflection processes that are part of the teaching quality framework are mechanisms to support the development of all faculty, not just those seeking tenure or with long-term stability at the institution.
- Departments: Having a suite of well-designed community-supported assessment tools will make it easier to more accurately and consistently evaluate faculty teaching for the purposes of merit raises and potentially tenure and promotion. A framework for teaching quality can also be used to assess whether the student experience is comparable in courses that have multiple sections that are taught by different faculty, graduate students, and learning assistants. This will help reduce the current overreliance on FCQs for the evaluation of faculty.
- Administration: Having a better definition of “multiple measures” can signal (internally and externally) that the university values education while simultaneously enhancing its ability to improve education on campus. It can support the accreditation process. It also allows for clearer messaging to outside constituencies about the types of innovative scholarship of teaching and learning that are taking place on campus.

The development and rollout of a framework will require an initial investment of time and work, but use of the framework may not require additional resources in the long run. In fact, the proposed framework is anticipated to: make teaching more effective and efficient, use student

perceptions of their learning experience wisely, and provide ready-to-use resources that could be adapted by departments. In more detail:

- At present, a lot of effort is expended to evaluate faculty, but it is not used to support their development and the products of this effort are of dubious quality for assessment. By reimagining the ways in which students and faculty can contribute to the teaching assessment process, the framework has the potential to use existing resources more efficiently.
- Based on studies of faculty time (e.g., Boice, 1991) we know that early career faculty spend as much or more time on their teaching as on their research endeavors. Having toolkits for reflecting on, documenting, and improving teaching practices, along with established goal-posts (objectives for faculty), can support these earlier career faculty. We can shift normative teaching / educational practices from individual and private acts to collective and public acts, thereby reducing the burden to individual faculty and enhancing their capacities.
- The existing FCQs ask students to evaluate faculty using criteria that they are not well-positioned to judge; there is evidence from numerous studies that students are poor at assessing their own learning (Dunning et al., 2004) and that questions that ask students to rate courses or instructors in a generic way are susceptible to bias in terms of race, gender, and similar factors (Baldwin & Blattner, 2003). Instead, students' time should be used to generate data that are valuable. We have exemplars of how to collect such data from the University of Kansas (and other leading universities) that move away from asking students omnibus, biased questions and towards asking for observations of more concrete faculty behaviors.
- Part of the framework will be a bank of resources that departments can use as starting points for developing their own teaching assessment strategies (we provide a preliminary set in the appendices and have a sample of tools from many institutions). Providing these examples supports departments in implementing new measures of teaching that are appropriate for their context without having to reinvent the wheel.

2. Components of a Teaching Quality Framework

The framework that we propose breaks the teaching enterprise into six components that characterize scholarly activity and assesses these components through the contributions of three sources of data, or “voices.” Just as with the evaluation of research effectiveness, it is the individual faculty member who should be responsible for making the case for her or his educational effectiveness using the evidence provided by these voices. One plausible mechanism to do so is have faculty members create portfolios that address each of these components of scholarly activity, focusing on their growth over time. The portfolios could also be made public to other faculty, to serve as examples of reflective practice and to be an object for collective growth.

We note that a large amount of faculty and student resources are already allocated to the evaluation of teaching. Thus, a key goal of the framework is to make the evaluation of teaching

more straightforward. One manner in which we may enhance the evaluation process by providing clear goals for instruction and mechanisms to support faculty learning through the assessment process. This approach helps better utilize the resources that are already being allocated. We also offer suggestions for how to streamline the process, arguing that each participant in the evaluation process should only be asked to provide feedback that aligns with what he or she is uniquely positioned to assess and that faculty members can actually use to grow as teachers. Not only will this increase the quality of data gathered, it may reduce the amount of effort expended. By shifting the focus to an instructor's growth over time, not just absolute ratings of quality, it is possible to turn the evaluation process into a learning process.

2.1. Six Components of Scholarly Activity

Teaching and learning are complex social processes, studied by numerous professional communities (e.g., educational psychology, discipline-based education research (DBER), and the learning sciences). Researchers in these communities take teaching and learning as the focus of their scholarly activity, and through their research, we now know more than ever about how people learn. However, faculty within most academic disciplines at research extensive institutions are primarily hired for their scholarly expertise in research in a given domain that is not teaching and learning. Nevertheless, these are the individuals who are creating the learning experiences for students within their disciplines. Hence, we take it as a goal of an institution like CU to bring tools and structures from those fields in which education is a focus of attention to all other disciplines. This approach will encourage all faculty to take a more scholarly approach to teaching and its evaluation by reflecting on their teaching using skills similar to those they use in their research work. Encouraging this type of reflective practice is the reason for focusing the framework on the components of scholarly activity.

Ernest Boyer's publication, *Scholarship Reconsidered: Priorities of the Professoriate* (1990), has played a key role in broadening the perception of academic scholarship. Boyer defines four types of scholarship, including the scholarship of teaching. One of Boyer's major arguments is that all of these types of scholarship share much in common, so that one can use many of the same metrics that would be used to assess other types of scholarship (e.g., the scholarship of discovery) to assess teaching as well. Subsequent work, *Scholarship Assessed: Evaluation of the Professoriate* (Glassick, Huber, & Maeroff, 1997), has made great strides to operationalize the assessment of all forms of scholarship in terms of six components. We list these components and illustrate them with related questions 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 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? 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). We also note that these six components are illustrated in the most common educational context, the classroom; however these components of teaching scholarship can easily be applied to other educational endeavors, including: the development of new curriculum, new courses, and innovative classroom materials; the supervision of independent studies; the mentoring of undergraduates; the supervision of internships; and to field work

To make it easier to assess these six categories, Bernstein and colleagues have created a rubric with four levels: entry into teaching, basic skill, professional, and advanced. The clear delineation of levels of accomplishment in each of the six components makes growth and feedback designed to foster growth more likely. The rubrics are designed to help scholars aspire to the highest levels of professional practice in teaching. See Appendix A for an illustration of these six core components and four levels of accomplishment (from Bernstein et al, 2010).

2.2. Three Voices (Sources of Data)

To accurately capture evidence of scholarly teaching, it is crucial to use multiple data sources ("voices"). This use of multiple voices is consistent with the existing (but not well-defined) standard of using multiple measures of teaching effectiveness. When assessing a faculty member's teaching, there are three major sources of data from which we might draw: (1) students, (2) faculty peers, and (3) the individual instructor. Each of these voices is able to speak to specific aspects of the faculty member's scholarly teaching. For instance, students spend the most time with the faculty member in class sessions and office hours, so they have the clearest picture of what happens in class and are uniquely positioned to report on their perception of instructional time. However, given that the students are still novices in their field of study, they are not the most qualified to comment on the instructor's knowledge of the domain or choice of topics to include in the course; faculty peers would be much better qualified to do so. Peers who are knowledgeable in education research can also comment on whether the chosen methods of instruction are appropriate and as up-to-date as possible. 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 and learning, all of which are relevant in making an appropriate assessment of his or her teaching.

2.2.1. Students

There are a variety of studies that suggest the limitations and biases of traditional end-of-term student evaluations (e.g., Baldwin & Blattner, 2003). While students can provide valuable information about their educational experiences, it is important to keep the focus on aspects of teaching that students can accurately describe (e.g., the practices that occurred in the classroom or their own experience of how different teaching practices support their learning), rather than subjective perceptions (e.g., how much they “like” the professor or how knowledgeable they perceive her or him to be). By avoiding such omnibus questions and instead asking more targeted questions of students (like those in Appendix B), we should be able to document the educational context and how students perceive this context. In so doing, we may reduce the gender and racial biases that arise in current student-rating systems and thereby promote greater equity and diversity in the university faculty.

It is most accurate to think of the information given by students as either ratings or descriptions, not proper “evaluations.” In other words, students can be thought of as providing information about their experiences in a course with an instructor, but they lack the disciplinary knowledge, expertise, and perspective to properly evaluate their professors. This sort of information can be gained through faculty course questionnaires (FCQs), but the FCQs may need to be modified to best capture what students can accurately assess and eliminate the biases alluded to above. In adapting these student rating systems, instruments such as the Student Assessment of Learning Gains (SALG, Seymour et al., 2000) will certainly prove valuable. An example of a revised FCQ form along with reporting structure over time is provided in Appendix B.

Students can also provide their voices through formal observations of classrooms by training them to administer classroom observation protocols such as the Teaching Dimensions Observations Protocol (TDOP; Hora et al., 2013) or the Classroom Observation Protocol for Undergraduate STEM (COPUS; Smith et al, 2013). These protocols are designed to objectively capture the practices occurring in a classroom. Students may be best positioned to conduct these observations as their non-expert perceptions are more likely to be aligned with what other students actually perceive rather than the perceptions of faculty members. A sample list of observational codes from the TDOP is given in Appendix C.

2.2.2. Faculty Peers

Faculty peers can provide forms of evaluation that students cannot; they have much deeper knowledge of a discipline, its professional culture, and the range of possible instructional designs in the field. This evaluation should involve more than just observation of teaching events. The peer should also examine syllabi, assignments, samples of student work, trajectories of uncurved student scores over time, and the instructor’s claims about their

intentions and plans. Looking at the practice of teaching in this holistic way shifts it from being perceived as merely performance and acknowledges the many factors that go into creating a productive educational experience.

Additionally, typical peer evaluations are conducted as one-shot events and therefore fail to capture growth and change over time (i.e. they are typically focused on evaluation of outcomes rather than supportive feedback and growth). Thus, a key paradigm shift for the peer voice is to move away from monolithic evaluations of outcomes that judge instructors on an absolute scale and to include feedback that judges instructors based on their trajectories over time (Bernstein, 2008). Given that the process we propose is more extensive than typical peer feedback, one would not expect it to take place every semester, but on a longer timescale (perhaps every few years). While faculty members would likely only engage in this extensive portfolio creation process every few years, they might also engage in less extensive reflections, documenting work in their portfolio less formally, on an ongoing basis, say every semester or year, to support ongoing review. Furthermore, appropriate reallocation of faculty time and potential additional resources (such as undergraduate observers graduate student writing partners, or teaching postdocs) would ensure that there is little impact on the overall allocation of effort dedicated to peer observations.

The current approach to peer classroom observations do not typically use well-designed protocols, instead relying on each observer to decide which aspects of teaching are or are not important. Often, there is no formal process for the instructor to see the report generated by his or her peer. In this way, a large amount of faculty resources are used, but faculty growth is not meaningfully supported and the quality of evidence generated is questionable. By shifting towards a model that emphasizes holistic growth over time and gives immediate feedback to the instructor, it is possible to use these resources more effectively to promote faculty development. This is one area where educational experts (e.g., DBER faculty within a department) may play a particularly powerful role.

Finally, reframing “peer evaluation” from judging to supporting growth can help mitigate some of the difficulty faculty have of evaluating their peers in the same department who may have complicated professional relationships with each other. Instead, faculty can provide critical and constructive feedback to support one another, without fear that they might be jeopardizing their colleagues’ future with negative evaluations. It may also be useful to allow faculty to choose at least some of the peers who will evaluate them.

An example and description of this revised approach to peer observation for faculty development and evaluation is provided in Appendix D.

2.2.3. Instructor

The individual faculty member also contributes valuable perspective to her or his teaching. Some of the characteristics of scholarly teaching are not necessarily associated with an individual course or a particular class session, so it is the individual instructor, not students or

peers, who must provide certain forms of evidence. For example, the instructor is best able to describe his or her process of reflecting on teaching and the rationale behind her or his lesson planning and design. The instructor can also provide a sense of growth over time, by reflecting on and comparing multiple offerings of the same course. The instructor can provide evidence of student learning and of their own professional development. This process of data collection not only contributes to providing evidence for the evaluation of teaching, but also provides feedback to support the instructor's growth. Because the instructor has a choice of which materials to focus on and include, this approach also makes the individual reflections a learning process. This approach is one mechanism to empower faculty to continue to improve in their practices. An example of such reflective practice and the associated portfolios produced by faculty may be found in Appendix E.

In terms of support, one mechanism to help faculty create their portfolios would be by having writing partners. A writing partner could be a graduate student in the same academic field as the faculty member. The role of the partner is to interview the faculty member and help in the reflective write-up process. Having a real audience will help faculty in the reflection process, and also free up some of their time and resources to make this less of a burden. This partnership also supports graduate students, who can learn a lot through this process and better prepare the students for future careers. As noted before, this reflective practice would not be mandatory of any faculty, but it would be a mechanism for those who were aiming to achieve excellence. This strategy has been implemented effectively at the University of Kansas. While we note that the institutional context is different at CU Boulder, we believe it is an approach that can be adapted successfully at any type of institution where scholarly educational practice is valued. We also note that these portfolios are not something that would be created every semester, but every few semesters.

3. Strategy for Implementing a Teaching Quality Framework

Our proposed strategy for implementing a Teaching Quality Framework at CU is built on a gradual, opt-in process that allows for as much ownership of the framework by the faculty as possible. The initial step is to convene a taskforce with the charge of developing a framework based on a bank of example materials from other institutions. In parallel with the work of this taskforce, opt-in pilot departments will work to contextualize the framework for their own uses. (We suspect that the initial most productive use of such a contextualized framework, or departmentally based rubric of teaching quality, will be for evaluating "teaching" in the merit raise process; however it will be up to departments to delineate their best application of their rubrics). One plausible way to use the framework for merit raises would involve in-depth reflections every few years, supported by shorter reflections on an ongoing basis. It may be that the more robust updates to the portfolio would be used primarily for larger events, such as reappointment, tenure, and continued advancement post-tenure. In environments where we can document the promise of this approach, the administration can proclaim their support for the framework and hold up the pilot departments as exemplars for others to follow, not only in the merit process but also for tenure and promotion decisions.

3.1. Taskforce to Develop the Framework

Our first proposed step is to develop a taskforce of approximately 15 key stakeholders charged by the provost, chancellor, or appropriate senior administrator to create a framework for teaching quality. This framework needs to be designed for the local context at CU, and it should have enough specificity so as to be understandable and applicable to our departments, while having enough flexibility to allow for appropriate variation across campus. The Framework described above, along with examples from other institutions, would serve as a starting point.

The deans and other stakeholders will be asked to recommend people for membership on the taskforce. The taskforce will consist of key members of the faculty who are recognized as campus leaders; as knowledgeable about education, evaluation, and CU-Boulder culture; and as trustworthy by faculty across campus. The taskforce should consist of a mix of full professors, younger faculty, and members of key campus teaching initiatives to allow for a wide representation of views.

The team working on CU's Association of American Universities (AAU) STEM Education Initiative project may be able to help with the construction and operation of the taskforce. It may also be useful for the taskforce to have external advisers consisting of national leaders in higher education change who have experience higher education transformation (e.g., Pat Hutchings [Bay View Alliance], Mary Huber [Bay View Alliance], Dan Bernstein [University of Kansas and Bay View Alliance], and/or Linda Slakey [National Science Foundation, AAU, and Bay View Alliance]).

The timeframe for producing a framework is roughly one semester, with a target of Spring 2016.

3.2. Pilot Adoption

In parallel and conjunction with the taskforce that creates the framework, individual departments will contextualize the framework for their own use. This process will begin as opt-in, with a goal of 5 or 6 departments self-identifying for participation. These departments will serve as exemplars to help other departments successfully follow their lead. Of course, as this occurs in parallel with the campus taskforce, the framework (and how it is contextualized) will be iterative, and we anticipate both the framework and the rubrics within departments to be living (evolving) documents.

To identify potential departments, a lead from our AAU project team or other campus-appointed individual will meet with chairs from 6 to 8 departments. These initial meetings will allow us to get a sense of what departments are currently doing and what types of support they would need to participate. The departments that are brought into the pilot round would be those who have strong support from their department chair and have expressed willingness through a memorandum of understanding to actually use the framework. From work related to the AAU project, there is already significant evidence that several departments would be willing to

consider an approach aligned with our proposed framework and would value the opportunity to define tools and processes for more effectively evaluating the teaching enterprise.

In each pilot department, 2 to 4 faculty members will be identified to direct efforts within their own department and communicate progress with groups outside of their department. These faculty members would likely be drawn from or have connections to existing committees in charge of merit raises, coordination of faculty peer observations, or other departmental teaching evaluation functions. They will be provided with incentives to participate through various mechanisms such as course buyouts, service credit, and the explicit support of their department chair. Their explicit charge will be to develop and execute an inclusive and departmentally-valued process by which the framework can be contextualized and used in the merit review process as early as the 2017-2018 academic year. We advise that the initial focus is on merit review rather than tenure and promotion because: (1) merit review affects all faculty members in a department equally, including senior faculty who already have tenure and (2) focusing on promotion and tenure will put an undue amount of risk on the few individuals whose tenure and promotion cases happen to be up for review at that time. At the same time, it would be dependent upon the department, in consultation with the campus-wide taskforce, to determine the best course of action. Before enacting any change in evaluation criteria, senior administration (from faculty affairs to the deans offices) would commit to ensure adequate support within their offices and their promotion and tenure committees.

To support these teams, we suggest the creation of a working group that combines the 2 to 4 faculty member teams from the 5 to 6 pilot departments. This working group would resemble a mix of a Faculty Learning Community (FLC; M. D. Cox, 2004) and a Department Action Team (DAT; Corbo et al., in press), the latter of which is a new type of faculty team that we have developed as part of the AAU project specifically to support departmental change. This group of 10 to 28 faculty would work 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. The AAU project team may be able to facilitate this group. In parallel those deans and administrative committees responsible for review of tenure and promotion cases would be included to advise and frame the productive application of this framework.

The primary role of the departmental committee will be to contextualize the framework to their units and help develop standards. This process may involve mapping the three voices identified above to some sort of framework, either the six components of scholarly activity previously identified, or any framework that is developed by the 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 committee to decide how they will proceed.

We envision this process lasting for two semesters: Fall 2016 and Spring 2017.

3.3. Public Proclamation of Support

Once the pilot departments work through their adoption of the framework, they will be able to serve as exemplars for other departments. 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 would 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. In parallel, we will consider listening tours and eliciting public input about the teaching quality framework.

Moreover, in developing this framework, CU would be a leader among research universities in conducting this work. Thus, the piloting of the framework could be connected to larger messaging campaigns on and beyond the campus.

While we suggest that the framework and the departmentally-specific evaluation rubrics will initially likely be used only in the evaluation of merit raises, it is ultimately the decision of the taskforce and departments to determine how they will use the framework. We anticipate that the rubrics and framework will be successful in this initial context and that it will pave the way for them to be used in tenure and promotion cases as departments become comfortable with their use. This transition will require communication with and participation by the relevant extra-departmental committees that are involved in promotion and tenure evaluation, in particular the Deans' and AVC-FA evaluation committees. The framework and the pilot departments' approaches would have to be shared with these committees, and they would have to be charged by the appropriate administrator to attend to, support, and review the cases for promotion and tenure they receive using the newly formed framework and rubrics.

4. Draft Timeline

Spring 2016

- Decide on taskforce membership and collect an initial set of resources for them to use.
- Identify potential departments to pilot the framework.
- Engage key campus committees (evaluation) and faculty assembly on the processes and goals of augmenting our current teaching evaluation practices.

Spring 2016-Fall 2016

- The taskforce uses the resources to develop a framework for teaching quality for CU.
- Departments pilot the framework, contextualizing to local use and provide feedback to the campus Taskforce.
- A cross-department faculty group is formed to support the rollout teams.
- Engage key campus committees (evaluation) and faculty assembly on the processes and goals of augmenting our current teaching evaluation practices.

Fall 2016 – Spring 2017

- The pilot departments and rollout teams work throughout the academic year to pilot the framework as contextualized in their departments. As departments demonstrate success in the creation of contextualized frameworks (rubrics), they may apply these in their merit (or promotion) processes, with appropriate support of their Deans and the AVC Faculty Affairs.
- Departments may also choose to implement the framework on a trial basis for 1-2 years without affecting merit raises to achieve greater faculty acceptance. Solicit feedback and engagement of key campus stakeholders (evaluation committees, faculty assembly, etc.).

Summer 2017

- Pilot department successes celebrated and shared widely by the administration.
- Additional departments encouraged to contextualize the framework for their own use.
- Pilot departments encouraged to use their evaluation rubrics for tenure and promotion.
- University tenure and promotion committees are educated in the existence and use of the framework and individual departmental rubrics.

Fall 2017 – Spring 2018

- Initial outcomes and successes from pilot efforts are shared internally across CU.
- Outcomes are also shared externally through national networks, including AAU, APLU, STEM Education Centers networks.
- Efforts are coordinated and celebrated with the Regents and state legislature

Authors & Suggested Citation

Noah Finkelstein, Daniel L. Reinholz, Joel C. Corbo
University of Colorado, Boulder

Daniel J. Bernstein
University of Kansas

Towards a teaching framework for assessing and promoting teaching quality at CU-Boulder (2015). Finkelstein, N., Reinholz, D. L., Corbo, J. C., & Bernstein D. J. Report from the STEM Institutional Transformation Action Research (SITAR) Project, Boulder, CO: Center for STEM Learning.

References

Baldwin, T., & Blattner, N. (2003). Guarding against Potential Bias in Student Evaluations: What Every Faculty Member Needs to Know. *College Teaching*, 51(1), 27–32.

Bernstein, D. (2002). Representing the intellectual work in teaching through peer-reviewed course portfolios. In *The teaching of psychology: Essays in honor of Wilbert J. McKeachie and Charles L. Brewer*. Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved from <http://kuscholarworks.ku.edu/handle/1808/7987>

Bernstein, D. J. (2008). Peer Review and Evaluation of the Intellectual Work of Teaching. *Change: The Magazine of Higher Learning*, 40(2), 48–51. <http://doi.org/10.3200/CHNG.40.2.48-51>

Bernstein, D. J., Addison, W., Altman, C., Hollister, D., Komarraju, M., Prieto, L., ... & Shore, C. (2010). Toward a scientist-educator model of teaching psychology. In D. F. Halpern (Ed.), *Undergraduate education in psychology: A blueprint for the future of the discipline* (pp. 29-45). Washington, DC: American Psychological Association.

Boice, R. (1991). New faculty as teachers. *Journal of Higher Education*, 62(2), 150–173.

Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the professoriate*. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching.

Corbo, J. C., Reinholz, D. L., Dancy, M. H., Deetz, S., & Finkelstein, N. (in press). *Sustainable change: A framework for transforming departmental culture to support educational innovation*. Physical Review Special Topics Physics Education Research. Retrieved from <http://arxiv.org/abs/1412.3034>

Cox, M. D. (2004). Introduction to faculty learning communities. *New Directions for Teaching and Learning*, 2004(97), 5–23. <http://doi.org/10.1002/tl.129>

Cox, J. (2007). Boulder campus guidelines for the gathering of information on teaching for comprehensive review, promotion, and tenure. http://www.colorado.edu/FacultyAffairs/atoz/Multiple_Measures_of_Teaching_November_2007.pdf

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed Self-Assessment: Implications for Health, Education, and the Workplace. *Psychological Science in the Public Interest*, 5(3), 69–106. <http://doi.org/10.1111/j.1529-1006.2004.00018.x>

Glassick, C. E., Huber, M. T., & Maeroff, G. I. (1997). *Scholarship Assessed: Evaluation of the Professoriate. Special Report*. San Francisco, CA: Jossey Bass Inc.

Hora, M. T., Oleson, A., & Ferrare, J. J. (2013). Teaching Dimensions Observation Protocol (TDOP) User's Manual. *Madison: Wisconsin Center for Education Research*. Retrieved from <http://tdop.wceruw.org/Document/TDOP-Users-Guide.pdf>

President's Council of Advisors on Science and Technology. (2012). Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics. Report to the President. *Executive Office of the President*.

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences* (Vol. 12). Boulder, CO: Westview Press.

Seymour, E., Wiese, D., Hunter, A., & Daffinrud, S. M. (2000). Creating a better mousetrap: On-line student assessment of their learning gains. In *Paper presented at the National Meeting of the American Chemical Society*. San Francisco, CA. Retrieved from <http://t.salgsite.org/docs/SALGPaperPresentationAtACS.pdf>

Smith, M. K., Jones, F. H. M., Gilbert, S. L., & Wieman, C. E. (2013). The Classroom Observation Protocol for Undergraduate STEM (COPUS): A New Instrument to Characterize University STEM Classroom Practices. *CBE-Life Sciences Education*, 12(4), 618–627. <http://doi.org/10.1187/cbe.13-08-0154>