

Final report of ad-hoc committee on student surveys of teaching

April 3, 2019

Committee members

Chris Elles
Heidi Hallman
Ravi Shanmugam
Holly Storkel
Doug Ward, chair
Ketty Wong

Contents

I. Overview and summary	1
II. Context for these recommendations	2-3
III. Increasing participation in student surveys	4-7
1. Have students complete surveys in class	
2. Remind students several times	
<i>Table: Classes with at least 80% participation rates (2015-18)</i>	
3. Add midterm surveys	
4. Give extra credit for completing surveys	
IV. Additional recommendations	7-9
1. Revise the questions and format of student surveys	
2. Revise the format of reports to instructors	
3. Adopt a more nuanced form of evaluating teaching	
V. Background.....	9
References	10
Appendix I: Benchmarks for Teaching Effectiveness	11-12
Appendix II: CLAS Guidelines for Evaluating Teaching Effectiveness	13-27
Appendix III: Updated survey questions from School of Health Professions	28-29

I. Overview and summary

Committee charge

Studying and recommending actions for increasing the completion rate of student surveys of teaching.

Summary of recommendations

No single action will magically improve student participation in end-of-semester teaching surveys, but there are several steps that faculty members can take to improve responses. These steps have proved effective with many instructors at KU. We see them as the easiest and most beneficial steps to take, although we identify additional possibilities below:

1. [Have students complete surveys in class](#)
2. [Remind students several times](#)
3. [Add midterm surveys](#)
4. [Give extra credit for completing surveys](#)

We also urge faculty governance and university administration to rethink the use and format of student surveys of teaching. (See [Additional recommendations](#).) The surveys have been grossly misused as an objective measure of an instructor's effectiveness. They *do not evaluate* the quality of instruction in a class, *do not correlate* with student learning, and *should not be used in isolation* to make decisions about promotion and tenure or merit raises. Reports from student surveys would be more valuable if they focused on developmental growth of instructors rather than providing a snapshot from a single class. Comparing instructors to a mean is a useless exercise because student surveys contain so many biases. For instance, instructors in STEM fields generally receive lower ratings than those in humanities. Those in large classes receive lower ratings than those in small classes. Those in required classes receive lower ratings than those in elective classes (Benton and Cashin, n.d.). The current system takes none of that into account and presents information in a way that suggests a neutral evaluation of individual instructors. That comparison against a mean should be eliminated.

We agree with university policy that student feedback about courses is important, but university policy also states that the surveys should be just one of multiple factors that determine an instructor's effectiveness. We recommend that the university embrace the work of the College of Liberal Arts and Sciences and the Center for Teaching Excellence in implementing a more thorough, fairer and nuanced evaluation process. Universities across the country are rethinking their approach to evaluating teaching, and the work of CTE and the College are at the forefront of that. Governance's support of this work would help make KU a national leader in this area.

We also recommend that a committee be appointed to revise the questions and format of the current student surveys to better reflect the multitude of course formats now in place (online, hybrid, short courses), the varied instructional roles, and the growing body of research about best practices in implementing student surveys. The surveys are best used as a means of formative feedback to instructors for improving future courses and should be structured in a way that emphasizes student comments. A five-point scale without comments provides little, if any, information about an instructor or a course and can be harmful to instructors and the teaching process.

II. Context for these recommendations

It is important to understand many contextual and cultural issues surrounding student surveys of teaching. These include the way surveys have been used and misused, and the recent history of the surveys.

Problems in the current process

The committee sees several problems with the evaluation process that hinder participation and diminish the validity of student surveys of teaching.

- **Decentralization and lack of consistency.** The evaluation process is decentralized and lacks coordination. The vice provost for faculty development sends a reminder about the student surveys of teaching each semester, but the surveys are handled by schools and departments, each of which has its own procedures and sometimes its own questions. Some departments solicit student comments and some don't, even though faculty members say the comments are often the most useful component of the surveys. That approach has created a hodgepodge of evaluation methods and a system with seemingly little oversight. For instance, the committee could find no one at the university level who oversees the process or who could provide university-wide data about the student surveys.
- **Lack of data.** The committee had only limited success in gathering university data about participation rates in student surveys of teaching. The Center for Online and Distance Learning was especially helpful, providing data about online evaluations for the past four years. That represents only a portion of surveys university-wide, though. Data about paper surveys apparently goes through the Office of Institutional Research and Planning. The vice provost for faculty development asked for that data on the committee's behalf but was told that gathering it would require a special algorithm to be written, something that the office had little interest in doing. Moving all departments to online surveys would provide a reliable dataset that would allow for better tracking in the future. Moving all surveys online would also save an enormous amount of staff time.
- **Perceived lack of value.** Many faculty members and students have low opinions of the end-of-semester surveys of teaching. Students don't have access to the class-by-class data – something they have sought repeatedly – and so don't see the process as benefiting them. The five-point scale on the surveys provide little relevant feedback for instructors, most of whom recognize the many biases in the evaluation process. The common name for the surveys – “evaluations of teaching” – exacerbates a lack of trust. The surveys allow students to provide feedback on classes and instructors, something that is important, but they provide little in the way of meaningful evaluation. That general lack of enthusiasm is one element of low response rates.
 - **Multiple surveys.** Students in lab courses are often asked to complete multiple evaluation forms for the same class. That creates confusion and can lead students to ignore some evaluations altogether. Exacerbating this, the surveys make no distinction between professors, GTAs, lab directors or other instructors, so questions are often irrelevant.

Biases in evaluations. Research over the past two decades has shown multiple biases in student surveys of teaching, making them an unreliable source for gauging teaching effectiveness. That plays a large role in instructors' lack of enthusiasm for the evaluations. Additionally, a recent meta-analysis of research into student surveys (Uttle, White and Gonzalez, 2017) found no correlation between student learning and scores on student surveys of teaching. That contradicts an earlier analysis by Wright and Jenkins-Guarieri (2012).

- **Fairness to instructors.** Many instructors teach courses that someone else has created, and they have little control over course content. That makes them vulnerable to negative reactions from students who dislike the course content or the course structure.
- **Fairness to GTAs.** Graduate teaching assistants and lab coordinators are evaluated in the same way as instructors, even though their duties and responsibilities are usually quite different.
- **Overuse of student surveys for P&T.** The committee certainly wants to see improved response rates for student surveys of teaching but also sees a need to minimize their use in the evaluation of instructors. The simplicity of the surveys has inflated their importance, with many departments using them as the primary tool for determining teaching effectiveness. They require little effort from faculty members or administrators and generate scores that seemingly allow comparison across a department. University guidelines call for multiple forms of evidence for evaluating teaching, but many, if not most, departments ignore those guidelines and use mean scores from student surveys of teaching combined with cursory class visits by peers. The College is implementing a more thorough evaluation process, and the Center for Teaching Excellence is piloting a rubric system that takes multiple forms of evidence into account in making judgments about teaching effectiveness. Those approaches show promise in making the evaluation process fairer and more meaningful while maintaining the student voice.

III. Increasing participation in student surveys of teaching

When student surveys of teaching move online, response rates generally decline (Nulty, 2008; Risquez, Vaughan and Murphy, 2015). The committee was unable to acquire data about surveys that were completed on paper, so it had no means to compare current rates to those from a few years ago. Even so, most instructors report that fewer students are completing the surveys.

The lower response rates aren't surprising. Paper surveys are distributed in class, and a portion of class time is set aside for students to complete them. For online surveys, students receive messages from the Center for Online and Distance Learning, an entity that most students don't know and never interact with otherwise. Students must act on their own to complete the surveys for each class they are enrolled in. Because the messages come during a busy time in the semester, they are easy for students to ignore, and many students do.

Despite the challenges, many instructors have continued to log response rates of 80% to 90% and above. We highlight some of those instructors' strategies below as we offer these recommendations:

1. Have students complete surveys in class

Returning surveys to class would provide an easy means to improve response rates. Completing the surveys outside class was necessary in the first three years of online surveys because students had to fill out the forms on Blackboard. The only way to do that was with a laptop or desktop computer. Online surveys are now handled through a system called Explorance Blue, which adapts easily to smartphones, tablets or computers. A vast majority of students have smartphones, so it would be easy for them to take the surveys of teaching in class. Instructors would need to give notice to students about bringing a device on survey day. Those who were absent or were not able to complete the surveys could still do so outside class.

- **Limitations.** Some instructors say that even when they set aside time in class, students ignore the surveys and simply check messages, play games on their phones or leave class early. To combat that, instructors might appoint monitors during the survey time to try to keep students on task. They should also adopt some of the additional strategies below.

2. Remind students about the surveys several times

As we said above, notices about the online surveys are sent by the Center for Online and Distance Learning, which also posts a link on Blackboard. Instructors who have had consistently high response rates send out multiple messages to students and speak about the surveys in class. They explain that student feedback is important for improving courses and that a higher response rate provides a broader understanding of students' experiences in a class.

- **Example.** Laurie Herrmann-Ginsberg has had response rates of 95% to 100% in her education classes of 20 to 25 students. She said her approach for getting students to respond was simple: "I'm a pest and keep reminding them." That approach will no doubt rankle some faculty members, but if instructors want higher response rates they must be invested in the process themselves.
 - **Limitations.** Courses with the smallest number of students – those in which students have the most interaction with instructors – have substantially higher participation rates than courses with large numbers of students. Research by Wachtel (1998) and others backs this up. So does data from online course surveys at KU. (See Table 1 below.) So to some extent, responses indicate the degree to which students feel a part of a class. If instructors have not engaged students or established that rapport during a semester, students are unlikely to fill out

surveys in large numbers. Even in classes where students feel engaged, a reminder or two from an instructor isn't enough. Rather, instructors should explain why the feedback from the surveys is important and how it is used to improve future classes. **An appeal that explains the importance and offers specific examples of how the instructor has used the feedback is more likely to get students to act than one that just reminds them to fill out the surveys.** Surveys can provide direction for future classes and future students, and current students need to hear that.

Table 1: Classes with at least 80% participation rates (2015-18)

Number of students in class	Total number of classes	# of classes with 80-100% participation	% of classes with 80-100% participation
5-9	3,243	692	21.3%
10-19	4,840	618	12.8%
20-25	1,570	213	13.6%
26-35	1,187	137	11.5%
36-60	846	81	9.6%
61-99	399	22	5.5%
100-199	315	11	3.5%
200 or more	155	5	3.2%
All classes	12,555	1,779	14.2%

Based on data provided by the Center for Online and Distance Learning

3. Add midterm surveys

A midterm survey helps instructors identify problems or frustrations in a class and signals to students that their opinions and experiences matter. This in turn helps motivate students to complete end-of-semester surveys. Many instructors already administer midterm surveys either electronically (via Blackboard or other online tools) or with paper, asking students such things as what is going well in the class, what needs to change, and where they are struggling. This approach is backed up by research from a training-evaluation organization called ALPS Insights, which has found that **students are more likely to complete later course surveys if instructors acknowledge and act on earlier feedback they have given.**

- **Example.** Roseann Pluretti, a journalism GTA, had 100% participation in 300-level course in 2016. The class had about 20 students. She gave her students informal surveys at midterm and also sent multiple reminders. Here's how she described her approach:
 - *First, I constantly stressed how their evaluations were super important to me and how I use them to improve my classes every year. I would give examples of comments given in past years that shaped their current class structure (like change in assignments, class time, etc). I would stress this at midterm time, and ask for informal midterm evaluations during class time, anonymously. Then I would the next week show them the results and what I was going to change and what I wasn't, with rationale for why I couldn't or wouldn't change some parts of the class for the rest of the semester.*

At the end of the year, I would remind them in class to do the final formal evaluations and I would again stress what I said at midterm time. I also sent emails every Sunday night before each week in the semester, of deadlines for the class etc., so at the end of the semester I would include reminders to do the evaluations in those emails as well.

- **Additional help for CODL.** The Center for Online and Distance Learning is already stretched thin with the end-of-semester course surveys done with Explorance Blue. If large numbers of instructors requested a midterm survey, CODL would need additional staffing to handle the workload.

4. Give extra credit for completing surveys

Instructors in large classes have found this an especially effective means of increasing student participation. Giving students as little as 1 point extra credit (amounting to a fraction of 1% of an overall grade) is enough to spur students to action. **This approach led to 89% participation in a general chemistry class of 511 students in Spring 2018.**

- **More extra credit can make a difference.** All but one of the 125 students in Ken Fischer's Mechanical Design course completed course surveys in Spring 2017. Fischer offered 3% extra credit for those who completed the survey. In retrospect, he said, that was probably too much. Next time, he will offer 1% to 2% extra credit. He generally has offered extra credit only in large classes but plans to do the same in small classes because participation rates have fallen.
- **Create a competition for completing surveys.** Some instructors have used participation in the surveys as a way for all students in the class to receive extra credit. Chris Gamblin has done this with Biology 454. All students receive 1 point extra credit if 80% of the class completes the surveys, 2 points for 90%, and 3 points for 100%. Participation rates rose from 50% in Spring 2016 to 100% in Spring 2018.

Additional possibilities

We see the four actions above as the most promising for improving response rates. Here are others that may work but that will require more time and in some cases may lead to alienation of students and instructors.

- **Create a pop-up announcement on Blackboard.** The Center for Online and Distance Learning already adds a survey link to Blackboard at the end of each semester, and that is how most students access the surveys, according to Laura Diede, director of CODL. The center could add a pop-up announcement, as well, including a reminder to complete the course surveys and a link to the survey site. This could be done in one of two ways. The first option would allow students to close the reminder box, although the box would show up again every time they logged in. The second option would keep the box open in Blackboard and would not allow students to close it until they had completed the online survey.
 - **Limitations and challenges.** This approach would have to be used for all students. It could not be done differently by schools or departments. A pop-up would no doubt annoy students, leading them to fill out surveys hastily or angrily, benefiting no one.
- **Create incentives for instructors.** There are no consequences for instructors when response rates to student surveys are low or incentives for having higher rates. That can lead to apathy among instructors about the surveys. To remedy that, the university could create either a rewards system for high response rates or penalties for low rates. For instance, instructors with response rates of 75% or above (or whatever rate was agreed on) could be entered in a drawing for a prize of some sort. This would need to be divided by class, size, though, as it is much easier to get higher participation in smaller classes. A

penalty for low rates would be much more difficult to implement and would no doubt alienate many instructors. It could also unfairly punish instructors in large classes, among others.

- **Give incompletes to students who don't complete surveys.** At least one department at the Medical Center does this, and response rates have been consistently high. Most of those classes are relatively small, though. Individual instructors give the incompletes, and students get a note saying they will get their grades after they complete the survey. It is up to the instructor to decide how long to hold out if the student refuses to complete the survey. We don't see this as a viable option for the Lawrence campus. It adds additional work for the instructor and on a wide scale would create confusion and anger.
- **Make grades available earlier for students who complete surveys.** The upside of this is that it would reward students for completing the surveys. The downside is that it would require additional administrative steps.
- **Make surveys available earlier.** The last two weeks of class are generally hectic for students and instructors. Making the surveys of teaching available a week or two earlier would give students additional time and might improve participation rates. Everspring has tried this in its classes, though, and has had little success. Students would also be filling the surveys out even though as much as a quarter of the semester remains.
- **Open survey results to students.** Students never see the results of teaching surveys, although they have asked the university to make the surveys open. Many universities do this. Students' argument is that they must make choices about which classes to enroll in but lack information to make informed choices. Their point is valid, and making survey results available would create an incentive for participating. Faculty at KU have resisted that, though, and taking this step would require a considerable change in culture. This could be a viable option, though, if the structure and content of the surveys were changed. (See below.)

IV. Additional recommendations

Much has changed since the survey questions were last revised in 2007-08. Online and hybrid courses have increased in number, more four- and eight-week courses have been created, more adjuncts and GTAs are teaching courses, core requirements have changed, and student attitudes toward education have become much more transactional. The [Association of American Universities](#), among other organizations, has pushed universities to improve teaching and to elevate the importance of teaching in the rewards systems. A fairer, more nuanced means of evaluation is an important part of that.

In addition, mounting evidence has challenged the validity and fairness of student surveys of teaching. For instance, Uttle, White and Gonzalez (2017) concluded in a meta-analysis of research that there was no association between student learning and scores on student surveys of teaching. Mitchell and Martin (2018) joined a growing number of researchers saying that the surveys are biased against women and instructors of color. At Ryerson University in Ontario, [an arbitrator ruled against the university's use of student surveys](#) as a means of determining instructors' teaching effectiveness.

This, along with other issues (see below), makes it imperative that KU revisit the structure and use of student surveys of teaching, and to add more nuance to the evaluation of teaching. Fortunately, other models are available, and adopting a more robust system for evaluating teaching would benefit instructors, departments and the university. We recommend the following:

1. Revise the questions and format of student surveys

The university should have a broad discussion about what it wants to achieve with student surveys and how they should be used. It needs to modernize the approach to the surveys and take into account the many biases inherent in them. Feedback from students is an important component of teaching and learning. It helps instructors make changes that benefit students and helps students feel invested in their learning. Unfortunately, the student surveys have become equated with teaching quality and have been misused – abused, even – in the promotion and tenure process. A crucial step in repairing this broken process is changing the questions and approach to student surveys of teaching.

- **Emphasize student comments.** Instructors overwhelmingly say that student comments are the most valuable component of student surveys of teaching. The comments point to specific areas where students have concerns or where a course is working well. Rather than emphasizing and encouraging those comments, though, the current surveys emphasize a five-point ratings scale that has been the source of many problems. Changes in 2007-08 improved the questions, focusing them on areas that students are qualified to answer, but the scale has persisted. We recommend that revisions eliminate the scale and encourage students to comment about a course and instruction. This would maintain the important feedback that surveys can provide but cut down on abuse of them as a single measure of teaching quality. **We suggest that FacEx appoint a separate committee to revise the student surveys.**
- **Add questions about student engagement.** These would provide information about such things as whether students bought books and other course materials, how much time they spent studying outside class, and how often they completed out-of-class work. The National Survey of Student Engagement contains similar questions that could be considered.
- **Combine survey results with other data.** Use of Explorance Blue provides an opportunity for adding student data so that administrators can better gauge student learning and success. Out of privacy concerns, that data should not be added to individual reports to faculty. Data about such things as gender, race, ethnicity, first-generation status, outside activities, and other characteristics would help administrators, deans and department chairs better understand the needs of students and more easily spot trends or problems across courses.
- **Differentiate among roles in courses.** Current surveys are one size fits all, even though instructors have many different roles. We recommend a separate survey for GTAs, adjuncts, lab assistants and instructors who do not design courses. That would provide more meaningful feedback to those instructors and would make the surveys more relevant to students. The School of Health Professions at the Medical Center has created separate questions for instructors, GTAs, lab instructors, and online and hybrid courses. It would be a good place to start in updating survey questions for the Lawrence campus. Jeff Radel, associate dean for academic and student affairs at the school, provided a copy of the questions. [See Appendix III.](#)

2. Revise the format of reports to instructors

Reports from student surveys would be more valuable if they focused on developmental growth of instructors rather than providing a snapshot from a single class. Comparing instructors to a mean is a useless exercise because student surveys contain so many biases. For instance, instructors in STEM fields generally receive lower ratings than those in humanities. Those in large classes receive lower ratings than those in small classes. Those in required classes receive lower ratings than those in elective classes (Benton and Cashin, n.d.). The current system takes none of that into account and presents information in a way that suggests a neutral evaluation of individual instructors. That comparison against a mean should be eliminated.

Providing substantive guidelines for a revised reporting format was outside the scope of this committee, but we see a clear need for a reporting structure that focuses on providing meaningful information to instructors for use in improving courses.

3. Adopt a more nuanced form of evaluating teaching

We urge faculty governance to embrace new models for evaluating teaching put forward by the Center for Teaching Excellence and the College of Liberal Arts and Sciences. CTE is part of a \$2.8 million National Science Foundation Grant intended to help academic departments provide a richer evaluation of faculty teaching and, ultimately, expand the use of practices that have shown to improve student learning. CTE is working with three other universities and nine departments at KU to test a framework that guides instructors and committees on gathering and evaluating evidence of effective teaching. The system uses a rubric to identify multiple sources of evidence from students, peers, and instructors. ([See Appendix I.](#)) This project is at the forefront of a North American movement to change the way teaching is evaluated. A related project at CLAS expands on the work of CTE and others, providing guidelines for an evaluation system that is fairer and more meaningful. ([See Appendix II.](#)) The Provost's Office has supported both of these efforts. Adding the support of faculty governance would further boost these efforts and help the university become a model for effective evaluation and effective teaching.

V. Background

Some history helps put the committee's work in context.

[In 2007](#), a Faculty Senate task force revised the form for student surveys of teaching, doing away with generic questions and focusing on such elements as clarity of course objectives and whether the instructor was clear, organized and respectful. That form was first used in 2008, although departments were still free to create their own evaluation forms.

Much has changed since then. A larger number of instructors are teaching in online and hybrid courses, the number of part-time and adjunct instructors has grown, a growing body of research has called the validity of student surveys of teaching into question, and a growing number of departments and schools have moved the surveys online.

In 2014, the Center for Online and Distance Learning began offering online evaluations, first for online classes and then for in-person classes. It used Blackboard for those evaluations until 2017, when it adopted a system called Explorance Blue. The new system provides much more flexibility in the way surveys can be taken (computers, phones, tablets) and makes the surveys much easier to administer and analyze. The online evaluation system also saves hundreds of hours of staff time each year. For example, Ally Smith, an administrative assistant in environmental studies, geology, geography, and atmospheric sciences, estimates that staff time needed to prepare data and distribute results for those four units has dropped from 51.5 hours with paper surveys to 4 hours with online surveys.

In 2015, Student Senate issued a report calling for all student surveys of teaching to be moved online, for the surveys to include additional questions, and for the results to be made available to students. No action was taken on that report.

References

- Benton, Stephen L., and William E. Cashin. "Student Ratings of Teaching: A Summary of Research and Literature." IDEA Paper 50, IDEA Center.
https://www.ideaedu.org/Portals/0/Uploads/Documents/IDEA%20Papers/IDEA%20Papers/PaperIDEA_50.pdf
- Doerer, Kristen (2019). "Colleges Are Getting Smarter About Student Evaluations. Here's How," *Chronicle of Higher Education* (13 January).
<https://www.chronicle.com/article/Colleges-Are-Getting-Smarter/245457/>
- Farr, Moira (2018). "Arbitration decision on student evaluations of teaching applauded by faculty," *University Affairs* (28 August).
<https://www.universityaffairs.ca/news/news-article/arbitration-decision-on-student-evaluations-of-teaching-applauded-by-faculty/>
- Mitchell, Kristina, and Jonathan Martin (2018). "Gender Bias in Student Evaluations," *PS: Political Science and Politics* 51:3 (July): 648-652.
- Nulty, Duncan D. (2008). "The adequacy of response rates to online and paper surveys: what can be done?" *Assessment and Evaluation in Higher Education* 33:3 (June): 301-314.
- Risquez, Anglica, Elaine Vaughan and Maura Murphy (2012). "Online student evaluations of teaching: what are we sacrificing for the affordances of technology?" *Assessment and Evaluation in Higher Education* 40:1 (February): 120-134.
- Uttle, Bob, Carmela A. White and Daniela Wong Gonzalez (2017). "Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related," *Studies in Educational Evaluation* 54 (September): 22-42.
<https://doi.org/10.1016/j.stueduc.2016.08.007>
- Wachtel, Howard K. (1998). "Student Evaluation of College Teaching Effectiveness: A Brief Review," *Assessment & Evaluation in Higher Education*, 23:2, 191-212.
<https://doi.org/10.1080/0260293980230207>
- Wright, Stephen L. and Michael A. Jenkins-Guarnieri (2012). "Student evaluations of teaching: combining the meta-analysis and demonstrating further evidence for use." *Assessment and Evaluation in Higher Education* 37:6 (September): 683-699.

Appendix I

Benchmarks for Teaching Effectiveness, Center for Teaching Excellence

This framework specifies seven dimensions of teaching practice, organized as a **rubric** to generate a comprehensive view of an instructor's teaching contributions. The framework draws on two decades of scholarship on peer review and evaluation of teaching. It can facilitate department committees' evaluation of faculty teaching, or guide a peer review or mentoring process that prompts reflection and iterative improvement. The category labels are designed to align with categories used in KU's P&T and merit systems. The scholarship on peer review and the evaluation of teaching suggests these essential components to teaching evaluation:

1. **Multiple dimensions of teaching:** The seven rubric dimensions are designed to capture the teaching endeavor in its totality, including activities outside of the classroom (e.g., identification of learning goals, assignment design, reviewing student work), and contributions to individual courses and the curriculum.
2. **Multiple sources** of information to speak to teaching effectiveness, including:
 - *The faculty member* - including course materials, evidence of student learning and reflections on it (often described in a narrative or portfolio)
 - *Peers* -including class visits, observations from team teaching, review and evaluations of course materials, and discussions with the instructor
 - *Students* – student course evaluations, additional feedback or student responses gathered by the instructor
3. **Adaptability:** The rubric can be adapted by departments to fit particular disciplinary expectations and to weight areas in ways most meaningful to the discipline. When completing the rubric, evaluators should consider the *types of courses taught* (required or elective, major or non-major, lecture or discussion, team taught or individual, size and level of class) and the *stage of the faculty member's career* (tenure track, tenured, instructor, adjunct). Departments may focus on various facets of the rubric at various stages in a faculty member's career, but at all times, evidence of student learning should be paramount.

How to use this rubric:

- **Review and Adapt Language-** Department members review the form and suggest modifications to make it appropriate for their discipline and department. This includes coming to a consensus about the questions and criteria in each category,
- **Assign Weights-** Identify the weights to be assigned to each category.
- **Identify Sources-** For each category identify appropriate sources of evidence. This should include a framework for how to read student evaluations of teaching and where they will be used as evidence within the rubric.
- **Use** to structure peer review or your own teaching narrative, or to integrate information from multiple sources.

BENCHMARKS FOR TEACHING EFFECTIVENESS



The Center for Teaching Excellence has developed a framework called Benchmarks for Teaching Effectiveness to support better methods of reviewing, documenting, and evaluating teaching. The framework is organized around a multidimensional rubric for reviewing faculty teaching. Seven rubric dimensions (below) have been designed to capture teaching in its totality. The rubric includes guiding questions and defined expectations for each dimension (see reverse). Departments are encouraged to adapt the rubric to fit disciplinary expectations and to weight areas most meaningful to the discipline.



GOALS,
CONTENT, &
ALIGNMENT



TEACHING
PRACTICES



ACHIEVEMENT
OF LEARNING
OUTCOMES



CLASSROOM
CLIMATE &
STUDENT
PERCEPTIONS



REFLECTION
& ITERATIVE
GROWTH



MENTORING
& ADVISING



INVOLVEMENT IN
TEACHING SERVICE,
SCHOLARSHIP
OR COMMUNITY

Benchmarks Goals and Objectives

1. Broaden faculty perspectives on and build consensus on effective teaching
2. Encourage the use of multiple sources of information to evaluate teaching (instructor, peers, and students)
3. Improve synthesis and representation of this information at the department or school level.

EXPLORING APPLICATIONS OF THE FRAMEWORK

CTE has received funding from the National Science Foundation for a 5-year-project that supports department-level adaptation and use of the Benchmarks framework. With assistance from CTE, participating departments are having conversations about what effective teaching is and how it should be evaluated. As they do this, they are adapting the rubric and identifying materials that that could provide information for each category. They are sharing their efforts with colleagues in other departments and with colleagues at the University of Colorado, Boulder and the University of Massachusetts, Amherst, which have created similar programs. The goal is to develop models that can be applied in other departments and other institutions.

Benchmarks Contact Information

If you have any questions or if you would like more information, please contact:

Andrea Greenhoot
CTE Director
agreenhoot@ku.edu
(785) 864-4193

Doug Ward
CTE Associate Director
dbward@ku.edu
(785) 864-7637

Kaila Colyott
Project Manager
kcolyott@ku.edu
(785) 864-7637

WHY WE ARE DOING THIS

Most evaluations focus on a narrow range of teaching practice and prioritize a limited source of evidence. Often, teaching is measured either through student evaluations, which contain inherent biases, or peer observations of a single class period. The Benchmarks framework provides a comprehensive, balanced view of faculty teaching contributions by broadening the types of activities that are reviewed and the sources of information on those activities. Thus, the Benchmarks aligns with KU policy, which requires multiple sources in teaching evaluation and specifies students, peers, and the faculty member as required sources in promotion and tenure and progress-toward-tenure processes.

*See reverse for complete rubric

This material is based upon work supported by the National Science Foundation under Grant Number DUE-1726087. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.



Rubric for Faculty Teaching Effectiveness (*department should modify as needed*)

*Aligned with KU Progress-toward-Tenure and Promotion & Tenure rating scales.	*Below Expectations: 1 - 2 <i>Poor (1): Consistently at this level</i> <i>Marginal (2): Some teaching at this level</i>	Meets Expectations: 3 <i>Competent</i>	Exceeds Expectations: 4 - 5 <i>Professional (4): Some teaching at this level</i> <i>Advanced (5): Consistently at this level</i>
Goals, content, and alignment <i>What are students expected to learn from the courses taught? Are course goals appropriately challenging? Is content aligned with the curriculum?</i>	<ul style="list-style-type: none"> Course goals are unclear, inappropriate, or marginally related to curriculum Content and materials are outdated or unsuitable for students in the courses Range of topics is too narrow or too broad Content is not clearly aligned with curriculum or institutional expectations 	<ul style="list-style-type: none"> Course goals are articulated and appropriate for curriculum Content is current and appropriate for topic, students, and curriculum Course topics include an appropriate range Standard, intellectually sound materials 	<ul style="list-style-type: none"> Course goals are well-articulated, high quality, and clearly connected to program or curricular goals Content is challenging and innovative or related to current issues and developments in field Topics are of appropriate range and depth, with integration across topics High quality materials, well-aligned with course goals
Teaching practices <i>How is in-class and out-of-class time used? What assignments, assessments, and learning activities are implemented to help students learn?</i>	<ul style="list-style-type: none"> Teaching practices are not sufficiently planned or organized, or are poorly implemented Practices are not well executed; little development in methods despite evidence of need Students lack opportunities to practice the skills embedded in course goals Student engagement is variable 	<ul style="list-style-type: none"> Teaching practices are well planned and organized Standard course practices carried out; follows conventions within discipline and institution Students have some opportunities to practice skills embedded in course goals Students consistently engaged 	<ul style="list-style-type: none"> Activities are well planned, integrated, and reflect commitment to providing meaningful assignments and assessments Uses effective, high-impact or innovative methods to improve understanding In- and out-of-class activities provide opportunities for practice and feedback on important skills and concepts Students show high levels of engagement
Achievement of learning outcomes <i>What impact do these courses have on learners? What evidence shows the level of student understanding?</i>	<ul style="list-style-type: none"> Insufficient attention to student learning – quality of student learning is not described or analyzed with clear standards Evidence of poor student learning; low level of skill/understanding is required or achieved without clear attempts to improve 	<ul style="list-style-type: none"> Clear standards for evaluating the quality of student understanding Typical student achievement for courses at these levels 	<ul style="list-style-type: none"> Standards for evaluating student understanding are connected to program or curriculum expectations, or use authentic assessments Efforts to support learning in all students Quality of learning supports success in other contexts (e.g., subsequent courses or non-classroom venues), or is increasing over successive offerings
Classroom climate and student perceptions <i>What are the students' views of their learning experience? How has student feedback informed the faculty member's teaching?</i>	<ul style="list-style-type: none"> Classroom climate does not promote civility or discourages student motivation and engagement Consistently negative student reports of teacher accessibility, interaction skills Poor sense of learning among students Little attempt to address concerns voiced by students 	<ul style="list-style-type: none"> Classroom climate promotes civility No consistently negative student ratings of teacher accessibility, interaction skills Most students indicate progress with their learning Instructor articulates some lessons learned through student feedback 	<ul style="list-style-type: none"> Evidence that classroom climate is respectful, cooperative, and encourages motivation and engagement Student feedback on teacher accessibility, interaction skills is generally positive Students perceive that they are learning important skills or knowledge Instructor is responsive to student feedback in short- and long-term
Reflection and iterative growth <i>How has the faculty member's teaching changed over time? How has this been informed by evidence of student learning?</i>	<ul style="list-style-type: none"> No indication of having reflected upon or learned from prior teaching or feedback 	<ul style="list-style-type: none"> Continued competent teaching, possibly with minor reflection based on input from peers and/or students Articulates some lessons learned from prior teaching and feedback 	<ul style="list-style-type: none"> Regularly makes adjustments to teaching based on reflections on student learning, within or across semesters Examines student performance following adjustments Reports improved student achievement of learning goals based on past course modifications
Mentoring & advising <i>How effectively has the faculty member worked individually with UG or graduate students?</i>	<ul style="list-style-type: none"> No indication of effective mentoring or advising students (but expected in department) 	<ul style="list-style-type: none"> Some evidence of effective advising and mentoring (<i>define as appropriate for discipline</i>) 	<ul style="list-style-type: none"> Evidence of exceptional quality and time commitment to advising and mentoring (<i>define as appropriate for discipline</i>)
Involvement in teaching service, scholarship, or community <i>In what ways has the instructor contributed to the broader teaching community, both on and off campus?</i>	<ul style="list-style-type: none"> No interaction with broader community about teaching, including involvement with teaching-related committees No evidence of keeping up with reports on effective teaching Practices and results of teaching are not shared with others Actions have negative impact on teaching culture in department or institution 	<ul style="list-style-type: none"> Some involvement in teaching-related committees, or engagement with peers on teaching (e.g., teaching-related presentations or workshops) Participates in department-level curriculum decisions 	<ul style="list-style-type: none"> Regular involvement in teaching-related committees, engagement with peers on teaching (e.g., teaching-related presentations or workshops) Occasional (or more) local or external presentations or publications to share practices or results of teaching Contributes to department or university curricular planning or assessment Advanced—Scholarly publications or grant applications related to teaching



Appendix II

CLAS Guidelines for Evaluating Teaching Effectiveness *Executive Summary*

One of the priorities of the college is to increase the success of all our students through high quality teaching that fosters enduring understanding. As such, teaching and learning is a primary function of the college and its faculty. Thus, it is crucial that teaching effectiveness be rigorously evaluated when making personnel decisions, especially promotion and/or tenure decisions. Evaluation of teaching is challenging because objective measures and external evaluations, like those used in evaluating research, are not available or possible. For this reason, it is important to regularly collect three sources of information on teaching effectiveness: (1) the candidate's reflection and activities; (2) student evaluations; (3) peer evaluations. Each source of information "can supply unique information, but also is fallible, usually in a way different from other sources" (Berk, 2005, p. 49). Using multiple methods and multiple perspectives allows "the strength of each source" to "compensate for weaknesses of the other sources" (Berk, 2005, p. 49). This also means that equal attention and care needs to be taken in developing each source of information and not reduce the evaluation of teaching effectiveness solely to student evaluations, which have known limitations.

1. ***Candidate's Activities & Reflections***: Candidates can provide evidence of their command of teaching practices and documentation of development of teaching practices. To accomplish this, candidates need to regularly document development of teaching practices over time both in terms of general skills (e.g., participating in training activities to learn new pedagogical approaches) and in terms of specific courses (e.g., tracking how their approach to a given course has changed over time). This documentation should serve as a foundation for crafting the teaching statement and should provide complementary detail to the brief teaching statement.
2. ***Student Evaluations***: Students can provide an understanding of their experience in a course. Student evaluations are regularly collected at the end of each course using a standard set of questions. Because of inherent biases in student evaluations, departments need to be thoughtful in their interpretation and use of these data, especially in comparing ratings across instructors and courses. Student comments (and other data sources) may be valuable in contextualizing numerical ratings.
3. ***Peer Evaluations***: Peers can provide valuable insights about a candidate's content knowledge, quality of teaching methods, fit of courses within the overall curriculum, and fit of courses' learning objectives within the learning objectives of the degree program. However, effective peer evaluations require structure and coaching of peer evaluators to provide substantive feedback.

Effective evaluation of teaching requires a [systematic approach](#) throughout the probationary or promotion period. It is critical to plan ahead to ensure that relevant data are regularly collected and discussed with the candidate. Each source of information needs to be collected at least annually.

Evaluation committees need to [integrate](#) the three data sources to accomplish the charge: "The evaluation of classroom teaching should consider evidence of the candidate's command of the subject matter at undergraduate and graduate levels, effective communication, commitment to student learning, development of teaching practices over time, student teaching evaluations, etc. The evaluation of teaching should be based on 'multiple sources of information and evidence

about the intellectual aspects of teaching and student learning” ([Instructions from Faculty Senate Rules and Regulations](#)). A rubric may be useful in evaluating and integrating multiple sources and in maintaining objectivity and consistency across candidates and over time. In some cases, [substantive action](#) may need to be taken prior to high stakes evaluations to ensure progress towards development of effective teaching practices.

The following recommended procedures provide more details about the value and shortcomings of each data source as well as best practice guidelines and options. In addition, numerous resources are provided to facilitate adoption of these best practices. The main goal of this document is to help departments be intentional and thoughtful in their approach to evaluating teaching effectiveness and to provide a variety of options to spark discussion around evaluation of teaching.

CLAS Guidelines for Evaluating Teaching Effectiveness *Recommended Procedures*

[Instructions from Faculty Senate Rules and Regulations](#) indicate that “The evaluation of classroom teaching should consider evidence of the candidate’s command of the subject matter at undergraduate and graduate levels, effective communication, commitment to student learning, development of teaching practices over time, student teaching evaluations, etc. The evaluation of teaching should be based on ‘multiple sources of information and evidence about the intellectual aspects of teaching and student learning’ as stated in [FSRR Article 6.2.2.](#)”

The instructions further indicate that three sources of information should be used to evaluate teaching effectiveness: (1) the candidate’s reflection and activities; (2) student evaluations; (3) peer evaluations. This multi-method, multi-perspective approach to evaluation of teaching effectiveness is considered best practice. Each source of information “can supply unique information, but also is fallible, usually in a way different from other sources” (Berk, 2005, p. 49). Using multiple methods and multiple perspectives allows “the strength of each source” to “compensate for weaknesses of the other sources” (Berk, 2005, p. 49). This also means that *equal attention and care needs to be taken in developing each source of information*. Guidance is provided to help candidates and departments gather strong evidence, over time, in all three areas to support a fair, comprehensive, and effective evaluation of teaching.

Candidate’s Activities & Reflection

Potential measures:

- (1) Development of teaching practices as evidenced by activities listed on the CV. Example activities include (but are not limited to):
 - a. Attending training activities related to teaching at KU or elsewhere
 - b. Observing or team teaching with others to develop teaching practices
 - c. Engaging in course or curricular revision activities
 - d. Applying for and receiving grants to support teaching-related activities
 - e. Developing course materials, such as textbooks or websites
 - f. Teaching related presentations at KU or elsewhere
 - g. Organizing training activities related to teaching at KU or elsewhere
 - h. Publishing articles related to teaching in peer-reviewed or non-peer-reviewed local or national outlets
 - i. Mentoring others (colleagues, GTAs) in teaching through peer observation, trainings, supervision (e.g., of GTAs) or other activities
 - j. Teaching awards
- (2) [Teaching statement](#) to demonstrate development of teaching practices and commitment to student learning
- (3) [Course portfolio](#) to document development of teaching practices, command of subject matter, and commitment to student learning
- (4) [Course syllabi](#) to show command of subject matter
 - a. Can be combined with [brief documentation](#) of course changes over time to demonstrate development of teaching practices
- (5) Self-ratings on a [teaching rubric](#) (areas tapped will depend on rubric)
- (6) [Reflection](#) on student or peer evaluations (areas tapped will depend on the nature of the evaluations)
- (7) Electronic copies of teaching presentations, posters, articles, and/or internal or external funding applications (areas tapped will depend on the topics discussed)

Value of this data source:

- (1) Establishes engagement in developing teaching practices
- (2) Demonstrates candidate's knowledge of teaching practices
- (3) Documents candidate's approach to teaching specific classes
- (4) Reveals candidate's perception of their teaching effectiveness
- (5) Can demonstrate development of teaching practices over time if multiple time points are included

Shortcomings of this data source:

- (1) Candidate's reflections may be biased, being too favorable or too harsh. This primarily impacts self-rating and self-reflection measures, but has a lesser impact on objective descriptions of teaching practices. Student and peer evaluations should be consulted as a means of corroborating the candidate's reflections.

Guidelines:

- (1) Teaching activities, beyond just a listing of taught courses, should be well documented on the CV
- (2) More than one measure of the candidate's reflection and activities should be included, to provide a broader view of teaching practices over the course of the evaluation period. Repeated reflection and documentation provides a scaffold for crafting the teaching statement that accompanies promotion and tenure dossiers.
- (3) In crafting the required [teaching statement](#), candidates should follow the prompt provided, which attempts to solicit a more objective description of how the candidate has developed and refined their teaching over time through 1 or 2 examples. Some candidates offer a more subjective statement of philosophy, rather than following the prompt. This subjective statement of philosophy is more open to bias and consequently may be less informative to evaluation committees. See the [sample statement](#) for a model of what the teaching statement should look like.
- (4) An overarching theme within the candidate's materials should be the connection between teaching and learning. High quality teaching practices by the candidate should facilitate strong learning by students. The items developed by the candidate should make this connection explicit (i.e., relate teaching practices to either intended or actual student learning).

Student Evaluation

Potential measures:

- (1) [Student evaluations](#) at the end of the course (e.g., ratings, open-ended comments), including ratings that are [proficiency based](#).
- (2) [Student evaluations](#) during the course (e.g., ratings, open-ended comments), including informal measures, such as the [muddiest point](#)
- (3) [Delayed student/alumni evaluations](#) that occur sometime after the course has been completed (e.g., to examine student-rated learning of taught skills and student-rated preparation for upper-level courses, work in the field, graduate school, etc)
- (4) Structured student/alumni interviews, which could incorporate open-ended versions of any of the linked survey questions above.

Value of this data source:

- (1) Reflects the students experience in the course
- (2) Validity and reliability are enhanced when students are asked to judge specific aspects of the course or instructor (e.g., "The instructor provided content and materials that were

useful”), rather than their overall impression of the course or instructor (e.g. “The instructor was effective”)

(3) Time efficient

Shortcomings of this data source:

- (1)** Students are less able to evaluate many important aspects of teaching (e.g., instructor’s content knowledge, quality of teaching methods, fit of the course within the overall curriculum), which are better evaluated by a peer.
- (2)** Can be biased and unreliable when global impressions are assessed. Thus, more specific prompts should be used as much as possible
- (3)** Ratings, even those tied to specific aspects of the course or instructor, may be influenced by situational factors (e.g., course level, student’s grade in the course), gender, race/ethnicity/nationality, self-presentation, proficiency in the language of instruction, and other factors, such as instructor personality or political perspectives, that may not relate to teaching effectiveness). Comparisons across courses and instructors should be approached with caution and numerical ratings should not be interpreted as highly precise measures (see Guidelines #1 below).

Guidelines:

- (1)** [Summaries of quantitative and qualitative data](#) for a particular course over time are especially helpful in understanding how teaching practices have changed over time.
- (2)** Numerical ratings should be interpreted in terms of their categorical meaning rather than in terms of quantitative differences. That is, small differences in mean ratings should not be interpreted as significant. For example, a mean rating of 4.25 and 4.50 on a 5-point scale (3 = neutral; 4 = agree; 5 = strongly agree) may seem different with 4.25 seeming “low” and 4.50 seeming “strong” but both mean ratings indicate general agreement with the rated statement. Both should be interpreted as a positive assessment of the instructor or course. In addition, the range of ratings on a particular item should be examined to determine the pattern of responses across students. Did the majority of students rate the course or instructor positively (e.g., majority of ratings clustered at 4 – agree and 5 – strongly agree)?
- (3)** Because situational factors make comparisons across instructors or across courses difficult, a better approach is to examine change in ratings for the same instructor over multiple offerings of the same (or similar) course to evaluate the trajectory of student satisfaction. That is, are the instructor or the course generally improving over time? Likewise, identifying and tracking themes in student comments may be helpful in documenting whether an instructor has been responsive to student feedback and is actively developing teaching practices around a specific course (or type of course).
- (4)** In examining the trajectory of a particular instructor in a particular course, it is important to consider the instructor’s role in the course (e.g., whether the course was developed by the instructor vs. developed by a departmental committee) and appropriately weight items that reflect the instructor’s primary role in the course. Likewise, the history of the course also should be considered in interpreting the trajectory. That is, changes in how the course is offered can lead to sudden decreases or increases in student ratings as the instructor develops their approach to the revised course. Instructors should not be penalized when courses are adjusted to better meet the needs of the student. Integrating student evaluations with other data sources should provide context for understanding the trajectory of an instructor in a particular course.
- (5)** Regardless of whether ratings are positive or negative, student comments or explanations of the ratings should be considered to contextualize the ratings. For example, an instructor may be rated less positively on an item like “What this instructor

expected of me was appropriately challenging” and the comments may reveal that the instructor gave rigorous assignments and had high standards for completion. The less positive rating may not be cause for concern because the program may appreciate rigor and high standards. Cross-referencing with peer evaluations also may be useful in establishing the appropriateness of the rigor and high standards.

Peer Evaluation

Key Resources: (1) Overview; (2) FAQ

Potential measures:

- (1) Peer observation of teaching performance (live or recorded classes), possibly using a [general rubric](#) or an observational framework focused on specific elements of teaching (e.g., [COPUS](#) or [GORP](#) to examine how instructor and students are spending class time)
- (2) [Peer review of teaching materials](#) (e.g., syllabus, blackboard, portfolio)
- (3) Peer review of measures of learning outcomes

Value of this data source:

- (1) Provides an evaluation of the instructor’s content knowledge, quality of teaching methods, fit of the course within the overall curriculum, fit of the course’s learning objectives within the learning objectives of the degree program
- (2) Validity and reliability are enhanced when a structured process is used (e.g., rubric, standard set of questions or prompts)
- (3) Peer feedback may be helpful in pinpointing specific areas for improvement and recommending campus resources to support that improvement

Shortcomings of this data source:

- (1) Can be biased and unreliable, especially when unstructured methods are used (e.g., observe and write a letter)
- (2) Time consuming for peer evaluators
- (3) Limited observations (e.g., one class observation) or review (e.g., review of one class assignment) may not be representative of overall teaching practices

Guidelines:

- (1) A structured review format (e.g., [rubric](#), [standard set of prompts/topics](#)) should be used to focus the review on areas where peers are best able to provide feedback (e.g., intellectual content, teaching practices, student learning). Ideally, the format will assist the peer reviewer in identifying strengths and areas for growth across a range of skills relevant to teaching effectiveness. It is important that peer reviewers be [encouraged to identify both strengths and areas for improvement](#) so that an authentic evaluation is achieved.
- (2) Peer reviews should be shared with the candidate. This sharing should include the completed written review along with a discussion of the review. The purpose of the discussion is to help the candidate think of ways to further develop and enhance their teaching. Any insights or plans resulting from this discussion should be documented as an addendum to the peer review, as noted in the candidate activities and reflection above.
- (3) Relatedly, the timing of peer review should be [systematically planned](#). When peer reviews occur too close to submission of materials to an evaluation committee, it place pressure on the peer reviewer to only make positive comments because the candidate has no opportunity to make changes in teaching practices prior to a high-stakes evaluation.

Multiple peer reviews provides an avenue for documenting growth in teaching practices over time.

- (4) Choice of peer reviewers should be discussed between the candidate and the unit. In some cases, appropriate peer reviewers may be external to the candidate's home unit. Likewise, non-tenure track faculty (e.g., teaching faculty) also may be appropriate peer reviewers. In selecting peer reviewers, both content expertise and teaching expertise should be considered and should guide the format for the review. That is, a reviewer with appropriate content expertise will be well qualified to comment on the instructor's content knowledge, fit of the course within the overall curriculum, and appropriateness of the learning objectives. However, if this same peer reviewer has not taught a course in a similar format (e.g., large class), then this peer reviewer may not be as well qualified to comment or provide guidance on teaching practices that are well suited to the class format. In this case, it may be appropriate for the peer reviewer to focus their review primarily on content. A mix of reviewers over time may be needed to effectively evaluate multiple dimensions of teaching.
- (5) Not all peer evaluations have to be used as part of the official dossier for promotion and/or tenure but an a priori system should be in place for determining this. For example, perhaps the first evaluation of a specific class is considered preliminary and informational to the candidate. Only second and later evaluations of a specific class are considered "final" and must be included in the dossier. There should not be "cherry-picking" of peer evaluations, such that only positive peer evaluations are included in the dossier. This undermines effective evaluation of teaching.
- (6) To ensure systematic peer evaluation of teaching, a departmental representative (e.g., department chair) should be in charge of tracking and implementing this process. Although candidates should have some input to the process (e.g., a voice in suggesting peer reviewers and specific semesters for peer review), it should not be the responsibility of the candidate to solicit peer reviewers.
- (7) Departments should develop a standard process for conducting peer review (e.g., rubric, standard set of prompts, differentiation of preliminary versus final peer evaluations, etc.) with clear expectations for participation as a peer reviewer (e.g., each tenured faculty member is expected to complete 2 reviews per year or certain faculty conduct peer review in lieu of other service commitments).

Timeline for Compiling Data Sources

Data sources need to be systematically compiled over the course of the evaluation period to accrue a representative sample of growth in teaching practices. Student evaluations are regularly collected for each course, although candidates and departments may wish to supplement that with additional sources of student information (as noted above). In addition, it is important to solicit student comments along with numerical ratings to contextualize ratings. Candidate reflection and peer evaluation also needs to be collected regularly, as shown in the table below. Given the schedule below, *candidate and peer measures should be incorporated into the annual faculty evaluation process* to ensure that they are being collected and discussed regularly. The overall goal is that multiple sources of data are regularly being used to evaluate teaching effectiveness and guide faculty development of teaching practices over time.

Guidelines for compiling candidate & peer data sources

	Y1-Y2		Y3	Y4-Y5		Y6	Y1-Y3	Y4-Y6	Y7	Y1-Y3	Y4-Y6	Y7
Candidate	x	x	Progress toward Tenure Review	x	x	Promotion & Tenure Review	x	x	Promotion to Full Review	x	x	Post-Tenure Review
Peer	x	X		x	x		x	x		x		

Note. x = at least one measure from this type of data source that can be included in the PTTR or PT dossier.

Integration of Data Sources

Not only is it important to have multiple sources of information documenting teaching effectiveness but it is critical to integrate these sources to gain a clear picture of teaching effectiveness. The candidate and committee should think carefully about the best way to integrate data sources. In many cases, it may make sense to integrate data sources at the level of the course. Most of the resources in this document assume that to be the case. Thus, a committee could easily examine the [candidate's synopsis of course changes by term](#) for a given course, the [summary of student evaluations](#) by course by term, and any peer evaluations for that same course. If the candidate irregularly teaches different courses then a different organization may make more sense, such as examining different courses that have a common element (e.g., grouping undergraduate service courses together).

In terms of interpreting different data sources, the candidate and peer information help to place the student evaluations in context. For example, new pedagogical practices (as documented in the candidate materials) may not be well accepted by students (as documented in the student evaluations). Also, implementation of new practices may not be optimal on the first attempt. As a result, student evaluations of teaching may be low but when placed within the context of an instructor who is trying something new to improve a perceived shortcoming of a course or to adopt a best practice method, low initial student evaluations may not be cause for concern. The evaluation committee can integrate these multiple data sources to better understand the instructor's overall development as a teacher.

This [rubric](#), in whole or in part, may be useful in evaluating and integrating multiple sources. Departments may wish to adapt this rubric or create their own to maintain objectivity and consistency in evaluating teaching effectiveness.

Actions for Improving Teaching Prior to High Stakes Evaluations

Ideally, any areas for improvement in teaching will be identified well in advance of high stakes evaluations (i.e., promotion and/or tenure review). If only minor improvements are needed (e.g., when recommended improvements enhance teaching that is already relatively effective), then typical annual review and feedback procedures can be used to make recommendations for enhancing teaching effectiveness and to track implementation of these recommendations. If more substantial improvements are needed to achieve effective teaching, then a teaching enhancement plan or a performance improvement plan may be needed.

Teaching Enhancement Plan. Teaching Enhancement Plans are a formal supportive action that can be triggered by a variety of processes, including annual review, any of the teaching evaluation procedures outlined in this document, faculty mentoring, and/or the request of the instructor. A Teaching Enhancement Plan can be created any time when a structured plan for improving teaching effectiveness is desired. Creation of a Teaching Enhancement Plan does not necessarily imply a failure to meet academic responsibilities.

Performance Improvement Plan. Performance Improvement Plans are a formal corrective action triggered by the annual review process. Each unit's [Faculty Evaluation Plan](#) stipulates the conditions under which Performance Improvement Plans are required (i.e., failure to meet academic responsibilities) as well as the general structure of such plans. The plans must conform to the Faculty Evaluation Policy, section #6: <http://policy.ku.edu/provost/faculty-evaluation-tenured-tenure-track>, which states: “the administrator and the faculty member shall develop a written plan of methods to improve the faculty member’s performance. The plan may include appropriate provisions for faculty development, such as campus opportunities for faculty continued renewal and development, reassignment of duties, or a change in teaching assignments. The unit administrator may call upon the University administration for assistance in constructing such a plan, including provision for additional resources, where needed. A faculty member may reject any plan recommended to aid performance levels, but the faculty member must understand that a sustained overall failure to meet academic responsibilities based on articulated performance criteria is a basis for dismissal.”

Plan Format. In general, the format of both types of plans are the same:

- (1) Clearly identify the area for growth. Borrow the phrasing from your evaluation instruments (e.g., improve command of subject matter, improve quality of teaching practices) so that the area for growth is tied to how the concern arose and so that improvements can be measured using the same instrument. Additional detail may need to be added for clarity and focus (e.g., improve quality of teaching practices with a focus on practices that enhance student engagement during class time).
- (2) List specific actions that the instructor (and possibly others) will undertake to enhance skills in the specified area. This can include learning resources and mentored activities. The Center for Teaching Excellence has a variety of resources and programs that may be relevant (see below) but also consider options for observing other instructors who have well-developed skills in the target area, co-teaching with more advanced instructors, and consultation/mentorship with more advanced instructors.
 - a. [Website Resources](#)
 - b. [2-Minute Mentor videos](#)
 - c. [Program Listing](#)
- (3) Construct a timeline for completing the specific actions listed in #2. Keep in mind that merely hearing about teaching methods is unlikely to produce change. Be sure to include

activities that require the instructor to apply newly learned methods to their own teaching.

- (4) Consider whether interim feedback and evaluation is needed prior to the end of the improvement plan and, if needed, specify how and when this will occur as well as by whom. Alternatively, specify any procedures for mentors to report interim progress.
- (5) Specify how improvements will be measured upon completion of the plan. Ideally, one of your existing evaluation instruments will have triggered the initial concern so you can continue to use this same instrument to document change.

Refer to the [sample plan](#) provided.

EEB Guiding the Peer Review of Teaching Process

The questions provided below can be used as a framework for the peer review of teaching. They are excerpted from sets of questions developed for peer review by Pat Hutchings for the first AAHE peer review project. They are congruent with a complete set of collaboration and review procedures described in Hutchings' (1996) *Making teaching community property: A menu for peer collaboration and peer review*. (Washington D.C.: American Association for Higher Education). This guide is not a recipe for conducting or reporting on peer review of teaching, rather it is designed to prompt reviewers to think about the possible components of the review process and the possible reporting of review outcomes.

1. Quality of intellectual content

The instructor can provide an annotated syllabus for a course, highlighting the decisions made in including material and choosing which particular aspects of the field will be included, excluded, or emphasized.

The reviewer should consider:

- Is the material in this course appropriate for the topic, appropriate for the curriculum and institution?
- Is the content related to current issues and developments in the field?
- Is there intellectual coherence to the course content?
- Are the intellectual goals for students well articulated and congruent with the course content and mission?

2. Quality of teaching practices

The instructor can provide a brief account of an overall plan for use of the students' time both in and outside of class. Some of this is found in the syllabus and some would be in the annotations to the syllabus.

The reviewer should consider:

- Is the contact time with students well organized and planned, and if so, are the plans carried out?
- How much of the time are students actively engaged in the material?
- Are there opportunities [in or out of class] for students to practice the skills embedded in the course goals?
- Are there any particularly creative or effective uses of contact time that could improve student understanding?
- Are there any course structures or procedures that contribute especially to the likely achievement of understanding by students?

3. Quality of student understanding

The instructor provides samples of assignments the students use to demonstrate their understanding of some of the key goals articulated in the first part of the review. For each assignment chosen for review, there could also be graded student work, complete with the feedback provided (if any). A recommendation is to provide two A papers (or equivalent category of judgment), two B papers, and two C papers, along with the distribution of performance for the entire class.

The reviewer should consider:

- Is the performance asked of students appropriate for the course goals, for the level of course, and for the institution?
- Does the performance requested include challenging levels of conceptual understanding and critical evaluation of the material appropriate to the level of the course and of the students?
- Are students being asked to demonstrate competence in the stated course goals? If not, is it possible to identify why?
- Are there obvious changes in the course that could improve performance?
- Are the forms of evaluation and assessment appropriate to the stated goals of the course?
- Are they particularly creative or do they provide students with opportunities to demonstrate their understanding using intellectual skills typical of the field?
- Is the weighting of course assignments in grade calculation coordinated with the relative importance of the course goals?

4. Summarizing the evidence of reflective consideration and development

The reviewer should consider:

- Has this faculty member made a sincere effort to insure that students achieve the goals for the course?
- Has the faculty member identified any meaningful relationship between what (s)he teaches and how students perform?
- Is there evidence that the faculty member has changed teaching practices based on past teaching experiences?
- Is there evidence of insightful analysis of teaching practice that resulted from consideration of student performance?

Example letters from CTE that follow a similar prompt format – [Letter 1](#); [Letter 2](#)

Tips for Providing a Constructive Peer Review of Teaching

Peer review of teaching is a critical component of mentoring faculty and instructors in developing teaching skills. Peer reviewers are in a good position to provide suggestions for both content adjustments as well as changes to pedagogical practice. In contrast, student reviewers do not have this same perspective of what content needs to be learned or what options are available to support their learning. Although peer review is a critical component in mentoring faculty, many peer reviewers find the process awkward and are hesitant to provide constructive criticism. Yet, most of us appreciate the value of peer review to development of our research skills. In general, peer review improves grants and research articles. The following tips are offered to assist faculty in providing a constructive peer review of teaching.

1. Organize your review around key elements of teaching using a rubric or standard set of prompts/topics as you would for a journal article, manuscript, or grant review.
2. Objectively state what was observed that was relevant for a specific topic area. This allows others to know the basis for your evaluative comments. This is the evidence on which you are basing your evaluation.
3. State your evaluation of the evidence. Is this element of the course/instruction a strength of the course/instructional approach? Is this element of the course/instruction a mix of strengths and areas for improvement? Is this element of the course/instruction an area where attention is needed for improvement? Be specific.
4. If you identify areas for improvement, be specific about what needs to be improved and provide concrete examples of how it could be improved. This provides a clear direction to the instructor about how to improve the course/their approach to instruction. This particular element is crucial to providing a helpful review rather than a punitive review. This step allows you as a peer reviewer to coach the instructor on how to improve and casts your review in a supportive light.
5. Treat the instructor as a colleague. We have all struggled to teach difficult classes. We have all struggled to learn new approaches to teaching. We usually knew when we were struggling, and it would have been helpful to have a colleague point the way to a solution. Try to avoid a judgmental tone or an air of superiority. Likewise, strive to provide a balanced review that lauds the elements of the course that are successful, notes the areas for improvement, and provides a clear path towards improving the course/approach to instruction. If you know of resources that would assist the instructor in making improvements (e.g., specific CTE programs, specific sections of the CTE website, other faculty with expertise in that area), provide that information.

Examples of Constructive Peer Feedback:

Example 1:

Topic: Goal, content, and alignment – What are students expected to learn from the course? Are course goals appropriately challenging? Is content aligned with the curriculum?

Rubric Rating = 4 out of 5 (Exceeds Expectations)

The goals of this course are to provide an overview of language development from birth through adulthood as well as an introduction to research paradigms used in the study of language development. The course also addresses the distinction between a language difference and a disorder as well as the characteristics of different language disorders. The last two items receive less emphasis than the first two items, which is appropriate for the level of this course (i.e., soph-jr; mid-curriculum). Overall, these goals are in line with departmental expectations for the course. The course provides good coverage over the range of language topics and over the range of ages. The textbook selected is appropriate and up-to-date. The content of the observed lecture relied on foundational research findings as well as fundamental research paradigms that would be useful for students to know. Overall, the goals and content are a strong aspect of this course.

The only place where minor change may be needed is that the syllabus is a bit brief and cursory making it difficult to get an in-depth view of what is being covered across classes. There is an online calendar that provides more detail but this is still at a pretty global level (phonology, lexicon, syntax). It would be nice if the core issues being addressed were better highlighted in the syllabus (e.g., dividing phonology into the actual topics covered, such as babbling, producing sounds to convey meaning, speech perception). This would be helpful for students in recognizing the content or core issue of different classes and in communicating course coverage to instructors of other courses at KU or instructors at other institutions where students may attend graduate school (and their undergraduate courses may be scrutinized to ensure appropriate coverage of pre-requisite content).

Comment on the example – Note that the coverage of the courses is described and then both strengths and areas for improvement are noted with a specific example of how to improve the syllabus (i.e., dividing large topics into subtopics).

Example 2:

Topic: Teaching practices – How is in-class and out-of-class time used? What assignments, assessments, and learning activities are implemented to help students learn?

Rubric Rating = 3.5 out of 5 (Meets/Exceeds Expectations)

In-class activities are lecture and discussion based. The first part of class (~15 minutes) was a re-cap of material from the prior class. Students actively answered questions and were usually correct or at least on the right track. This was a good way to check understanding. The remainder of class (~60 minutes) was more lecture based with a description of several classic psycholinguistic studies with infants that illustrate some core findings about what infants know and don't know about language. No PowerPoint was used but key points were written on the board. The lack of PowerPoint slides seemed to focus the students on the lecture. Specifically, no overt off-task behavior was observed (e.g., online shopping, Facebook). The pace of the lecture seemed appropriate for note-taking. Students were asked to volunteer information at different points but there was much less student involvement than what occurred during the first 15

minutes of class. INSTRUCTOR was a dynamic speaker and she appeared to be holding student attention during this longer lecture period. Overall, the structure of class activities was strong and the students appeared to be attentive. However, I would encourage some thinking about incorporating short interactive activities within the longer lecture segment. 60 minutes is a long time to listen to someone else talk. Also, passive listening may not encourage deep or critical thinking. A very brief activity could be effective in increasing engagement and deepening thinking. For example, INSTRUCTOR asked students to volunteer what their first word was as a baby. This activity could have been expanded to have students note the phonological structure of their first word, guess if their first word was a common first word spoken by other children, did they produce the first word in an adult-like way or in a “baby talk” version, etc. This brief activity would then support the lecture that followed on characteristics of first words and phonological idioms.

Out-of-class activities include objective quizzes and discussion boards. The quizzes are primarily multiple choice and thus automatically graded. Quizzes also include some short answer or more open-ended questions. Quizzes also are timed presumably to limit student’s ability to look-up every single answer. Generally, quizzes assessed basic understanding of course content and students appeared to perform well (majority of students earning grades of A or B), indicating a good understanding of that content. Discussion boards ask students to pose open-ended questions about a section of content and then to answer these questions. This is intended as preparation for exams which use open-ended questions. This was a great activity! The students posed some really good questions! Likewise, some of the answers also were really good! This was a nice way to encourage deep thinking. My only concern was that it didn’t seem like students got any feedback during this activity. So, I’m not sure how it would help them figure out what the “right” answer was to a particular question where “right” probably means more about having a complete and well thought out answer as opposed to a correct answer. Greater feedback might enhance learning in this activity.

Update to Comments after Discussion: This issue regarding discussion board feedback was clarified in a face-to-face meeting with INSTRUCTOR and PEER EVALUATOR. Students do receive feedback on the discussion board posts from the GTA but that feedback occurs privately (one-on-one communication between GTA and student with input and oversight from INSTRUCTOR). It would be ideal to have the GTA comment within the discussion board but this type of public feedback needs to be constructive and supportive and not all GTAs are successful in striking the right tone in their comments (as noted by the instructor). Thus, in those situations, private feedback is used initially (with feedback and coaching by INSTRUCTOR) until the GTA develops the right skill set to do this publicly on the discussion board. This is a nice way to balance student needs with GTA “learning curve” and demonstrates how seriously INSTRUCTOR approaches GTA training and how carefully she manages the GTA’s involvement in the course to ensure a quality experience for the students (and the GTA). Other instructors could learn from this approach. This could be a nice topic for discussion at a faculty retreat or at the KU Teaching Summit.

Comment on the example – Again, description of what was observed along with evaluation of strengths and areas for improvement are noted. Specific suggestions/examples are used to illustrate how to improve the approach to instruction. Also, note the importance of having a follow-up discussion between the instructor and peer evaluator. This discussion clarified an apparent weakness, revealing a strength in GTA training. The peer evaluator makes a suggestion for the instructor to become more involved in the scholarship of teaching and learning through a departmental or university venue.

Sample Teaching Enhancement/Improvement Plan

Goal: Improve quality of teaching practices in undergraduate courses with a focus on practices that enhance student engagement during class time.

Action	Timeline
(1) Attend sessions at the Teaching Symposium that focus on engaged learning	August, 20xx
(2) Attend monthly C21 and TRESTLE activities at CTE	Academic Year 20xx-20yy
(3) Observe undergraduate classes in and outside the department that have adopted engaged learning activities: CLASS 1 – INSTRUCTOR 1; CLASS 2 – INSTRUCTOR 2	October - November, 20xx
(4) Outline a plan for incorporating engaged learning activities to the first 2 units in CLASS XXX, which will be taught in Spring 20yy. Meet with INSTRUCTOR 1 and INSTRUCTOR 2 to review plan and receive feedback.	November – December, 20xx
(5) Revise outline for engaged learning for first 2 units in CLASS YYY. Outline plan for incorporating engaged learning activities for last 2 units in CLASS XXX. Share with INSTRUCTOR 1 and 2 for feedback.	December, 20xx – February, 20yy
(6) Implement plan for incorporating engaged learning in CLASS YYY with INSTRUCTOR 1 and INSTRUCTOR 2 observing class once during the first 4-weeks of class and once during the second 4-weeks of class. Meet to discuss feedback after each observation.	January-March, 20yy
(7) Peer evaluation of teaching conducted by INSTRUCTOR 3 for CLASS YYY post-Spring Break	March-May, 20yy
(8) Present course transformation activities at a C21 meeting or at the CTE Annual Poster Session and Celebration of Teaching on Stop Day of Spring 20yy	March-May, 20yy

Interim Benchmarks/Feedback: XXX is responsible for initiating all actions noted above. As noted above, XXX will be working closely with INSTRUCTOR 1 and INSTRUCTOR 2 for interim feedback and mentoring. The department chair will consult with INSTRUCTOR 1 and INSTRUCTOR 2 in October, December, and February to determine if XXX is following the plan outlined above.

Measure Change:

- (1) Post-Spring Break peer evaluation of teaching for CLASS YYY will be used to document whether more active learning activities have been incorporated into this specific class. The peer evaluator will not be one of the mentors (INSTRUCTOR 1 or INSTRUCTOR 2) for this goal.
- (2) Student evaluations of CLASS YYY from Spring 20yy will be examined. There is no rating item that specifically taps active learning, so comments will be examined. Prior comments noted: the “instructor lectured a lot,” “75-minutes is a long time to sit and listen,” and that “class was boring.” It is expected that these types of comments will be reduced with successful incorporation of active learning activities.
- (3) The department chair will check with CTE during Fall 20xx about the possibility of conducting a Classroom Observation Protocol for Undergraduate STEM (COPUS) on CLASS YYY in Spring 20yy to better document how class time is being spent.

Appendix III

Course Survey Questions from the School of Health Professions

Updated and approved, February 2019

Course Related Questions: *[NA, Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, or Strongly Agree]*

1. Given the importance of this course to my field, I felt the amount of time provided for the content was sufficient.
2. My knowledge of topics in this course improved.
3. The readings in this course helped me to prepare for and integrate course-related activities.
4. The course assignments were useful to my learning.
5. I understood the relevance of the topics in this course for my profession.
6. The instructional technology used in this course supported my learning.
7. Course design offered opportunities for collaboration with peers.

Course Open-Ended Questions:

1. Please provide an explanation for any items on which you marked "strongly disagree" or "disagree".
2. With regards to course structure, content and design, what did you find most beneficial to your learning?
3. If there were aspects of this course that were not beneficial to your learning, what changes would you suggest?

Instructor Related Questions: *[NA, Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, or Strongly Agree]*

1. The instructor was available and responsive to student needs and concerns.
2. The instructor was organized and prepared for each class in a way that was helpful to my learning.
3. The instructor provided quality feedback.
4. The instructor provided feedback in a timely manner. (e.g. responded to emails, returned assignments and exams within a reasonable timeframe).
5. The instructor conveyed content and materials in a manner that was helpful to my learning.
6. The instructor set and met clear goals and objectives for the course as described in the course syllabus.
7. The instructor encouraged student participation.
8. The instructor's assessment of me was fair and impartial.
9. The instructor demonstrated respect for students and their points of view.
10. The instructor was knowledgeable about the subject.

Instructor Open-Ended Questions:

1. Please provide an explanation for any items on which you marked "strongly disagree" or "disagree".
2. If you'd like to make comments about any of the above items, please type them below.

GTA Related Questions: *[NA, Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, or Strongly Agree]*

1. The GTA conveyed content and materials in a manner that was helpful to my learning.
2. The GTA was available and responsive to student needs and concerns.
3. The GTA provided quality feedback.
4. The GTA provided feedback in a timely manner. (e.g. responded to emails, returned assignments and exams within a reasonable timeframe).
5. The GTA contributed to meeting the goals and objectives for this course as described in the syllabus.
6. The GTA encouraged student participation.
7. The GTA's assessment of me was fair and impartial.
8. The GTA demonstrated respect for students and their points of view.
9. The GTA was knowledgeable about the subject.

GTA Open-Ended Question:

1. Please provide an explanation for any items on which you marked "strongly disagree" or "disagree".
2. If you'd like to make comments about any of the above items, please write them below.

Lab Questions: *All lab courses will have the same questions presented on page 1 of this document as well as the questions listed below.*

Lab Related Questions: *[NA, Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, or Strongly Agree]*

1. I feel prepared to perform the procedures/techniques taught in this lab.
2. I feel prepared to operate equipment/systems covered in this lab.
3. Adequate time was provided to practice and apply skills.
4. The lab activities helped me build upon other content in the program.
5. I feel able to relate the lab activities to clinical, technical, or policy relevance.
6. The instructor encouraged critical thinking (e.g., problem solving, trouble shooting, applying information) during lab activities.

Lab Open-Ended Questions:

1. Please provide an explanation for any items on which you marked "strongly disagree" or "disagree".
2. With regards to lab structure, content and design, what did you find most beneficial to your learning?
3. If there were aspects of this lab that were not beneficial to your learning, what changes would you suggest?

Online Course/Hybrid Questions: *All online courses will have the same questions presented on page 1 of this document as well as the questions listed below.*

Online Course/Hybrid Related Questions: *[NA, Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, or Strongly Agree]*

1. The online classroom was well-organized and accessible.
2. Technology used in the course enhanced peer interactions.
3. Technology used in the course enhanced student-faculty interactions.
4. The sound quality of audio or audiovisual resources was excellent.
5. The visual quality of visual resources was excellent.

Online Course/Hybrid Open-Ended Questions:

1. Please provide an explanation for any items on which you marked "strongly disagree" or "disagree".
2. With regards to online and technical content structure and design, what did you find most beneficial to your learning?
3. If there were aspects of this online course or its use of technology that were not beneficial to your learning, what changes would you suggest?