

Teaching Matters

News & Information from CTE

Spring 2018

Vol 21, No 2

Benchmarks seeks to broaden visibility of teaching work

Universities like KU have traditionally relied on methods for evaluating teaching that prioritize a narrow dimension of teaching activity (the behavior of the instructor in the classroom) and a limited source of evidence (student ratings). When other sources of information—such as peer observations or course material surveys—are available, reviewers often struggle to make sense of it all. Moreover, many new faculty members wish they had a better sense of how their teaching will be evaluated and had opportunities to learn from feedback. The Benchmarks for Teaching Effectiveness project increases visibility of all dimensions of teaching, clarifies faculty teaching expectations, enables quick identification of strengths and areas for improvement, and brings consistency across evaluations and over time.

CTE is taking part in Benchmarks, which is a five-year project supported by a grant from the National Science Foundation to foster improved methods of reviewing, documenting and evalu-



Sociology faculty members at Benchmarks project kick-off. Photo Doug Ward.

ating faculty teaching practices at KU and at the University of Colorado, Boulder, and the University of Massachusetts, Amherst.

Benchmarks is organized as a rubric and identifies seven dimensions of teaching practice that are designed to capture teaching in its totality. The rubric can be used to

guide the evaluation of teaching effectiveness (e.g., for promotion and tenure or progress toward tenure). It can also be used to foster teaching effectiveness, as part of a peer review or teaching mentoring system. This issue of *Teaching Matters* provides more information about the project.

In this issue

CTE View—Andrea Greenhoot discusses KU policies on faculty teaching evaluation and how the Benchmarks project intersects with them, pp. 2-3.

Snapshots—Examples of evaluating teaching at KU, pp. 4-5.

Innovations—Doug Ward describes the Benchmarks kickoff and ways that four departments and a school are piloting the rubric, pp. 6-7.

End Note—How COPUS documents use of class time, p. 8.

How have you spent your time on teaching this week?

Andrea Greenhoot
CTE/Psychology

How have you spent your time on teaching this week? When I ask KU faculty members this question, here are some of the activities they report:

- Identify what I most want students to learn
- Select readings
- Design in-class activities
- Develop and post materials to Blackboard
- Grade online discussion posts and integrate themes into class plans
- Figure out how to scaffold students' research papers
- Read papers and give feedback
- Create a rubric
- Plan and lead a field trip
- Write a reading quiz
- Write exam questions
- Lead classroom discussions
- Write letters of recommendation
- Explain lab protocols to new undergraduates
- Email students about how to prepare for an upcoming test

This is just a snapshot of the substantial intellectual work that goes into effective teaching and mentoring. Each of these activities requires thought, care, and most of all, time. Yet the most commonly used strategies for reviewing and evaluating teaching speak to only a narrow slice of these activities. Faculty at KU and other universities nationwide report that the evaluation of teaching relies heavily on student ratings. While the student perspective provides important feedback for instructors and review committees, and the ease of collection and their apparent applicability across disciplines make them appealing, the list of teaching activities provided above includes many that students have neither the opportunity nor the knowledge to address. Moreover, studies of student evaluations highlight several other limitations, including age-, gender- and race-related biases, as well as associations with

grading leniency and poorer student performance in follow-on courses (i.e., learning).

KU policy actually requires that multiple sources of information be used for faculty teaching evaluation, including students, peers, and the faculty member being evaluated. But the typical peer evaluation, consisting of an observation of a single class period followed by a letter describing the instructor's in-class performance, also prioritizes a narrow dimension of teaching activity. Review committees have few guidelines for how to make sense of and integrate the often voluminous information from the three sources. It is easy to understand why reviewers might gravitate toward the "clarity" provided by student evaluation numbers. The result is that most evaluations focus on a limited range of teaching practice and a limited source of evidence, and much work on teaching remains undocumented and invisible in the review process.

Interestingly, whenever the topic of teaching evaluation arises, most faculty express deep dissatisfaction with these ways of doing things. These sentiments are corroborated by the results of a recent survey of over 1000 instructors at eight AAU Universi-

... most evaluations focus on a limited range of teaching practice and a limited source of evidence, and much work on teaching remains undocumented and invisible in the review process.

ties: Less than 10% of those surveyed reported that the evidence used for teaching evaluation on their campus was of high quality.

In response to this state of affairs, my CTE colleagues and I have been working on a framework for the representation, review, and evaluation of teaching that we call Benchmarks for Teaching Effectiveness. A major goal is to yield a richer, more comprehensive view of faculty teaching contributions by broadening both the types of activities that are reviewed and the sources of information that speak to the quality and efficacy of those activities.

Drawing on the work of former CTE director Dan Bernstein, and two decades of scholarship on scholarly teaching and its evaluation, Benchmarks identifies seven dimensions of teaching practice that give visibility to activities such as designing and refining a course and evaluating and improving student learning, in addition to the activities that take place in the classroom. Benchmarks also identifies three critical sources of information to be triangulated: materials assembled by the instructor, peers, and students.

We have articulated the Benchmarks for Teaching Effectiveness framework in the form of a rubric, with guiding questions and defined expectations for each dimension. As we developed the rubric, we gathered feedback from CTE Ambassadors and department chairs to refine the categories and the language.

First and foremost, we would like the rubric to serve as a starting point for departments to have conversations about what teaching effectiveness “looks like” in their discipline(s). Most disciplines have some shared vision of what constitutes high-quality research in their field, but there seems to be far less consensus about what constitutes effective teaching. We encourage departments to adapt the rubric to their own disciplinary identities and build consensus on the categories and criteria. In this way, the rubric can help to clarify faculty teaching expectations and bring greater transparency and consistency to the evaluation process.

The Benchmarks rubric is designed to help department committees integrate information from multiple sources and improve teaching evaluation, but we are just as enthused about its potential for guiding teacher growth and development. For instance, the rubric can facilitate the mentoring of junior faculty and graduate student instructor by providing clear expectations for teaching effectiveness and enabling quick identification of strengths and areas for future growth. And some KU departments are already using adaptations of the rubric to guide regular peer reviews of teaching and to prompt reflection and iterative improvement.

We are excited that these KU units will be exploring innovative and effective applications of the

Benchmarks framework as part of the project funded by the National Science Foundation:

- Biology
- Chemical and Petroleum Engineering
- Philosophy
- Public Affairs and Administration
- Sociology

We have just launched a five-year collaborative project with institutional partners through the Bay View Alliance, in which three cohorts of KU departments will adapt and use the rubric for peer review, mentoring, and/or evaluation. The goal is to develop models of approaches to teaching evaluation that accurately reflect teaching effectiveness, minimize bias, and can be practically applied in university contexts.

For more information see:

Bernstein, D. (2008). Peer Review and Evaluation of the Intellectual Work of Teaching. *Change: The Magazine of Higher Learning*, 40(2), 48–51.

Dennin, M., Feig, A., Finklestein, N., Greenhoot, A. F., Hildreth, M., Leibovich, A., Martin, J., Miller, E., Moldwin, M., O'Dowd, D., Posey, L., Schultz, Z., & Smith, T. L. (2017). *Aligning Practices to Policies: Changing the Culture to Recognize and Reward Teaching at Research Universities*. *CBE-Life Sciences Education*, 16(4), es5. cbe.17-02-0032 doi:10.1187

Hutchings, P., Huber, M. T., & Ciccone, A. (2011). *The Scholarship of Teaching and Learning Reconsidered: Institutional Integration and Impact*. San Francisco, CA: Jossey-Bass.

Three examples of evaluating teaching

Implementing in-class peer reviews

Jordan Bass, HSES

In Spring 2016, the Department of Health, Sport and Exercise Science received support from CTE to design and implement a standardized peer review system for teaching. Having a consistent and agreed upon peer evaluation form was something both faculty members and the department's leadership were committed to developing.

To create the form, faculty members came together and had in-depth discussions about what they really wanted to get out of the peer review system. What followed was a comprehensive document that uses both qualitative and quantitative measure to help guide faculty as we evaluate each other in the classroom. Faculty members have been using the form successfully since Spring 2016.

We believe the form is not only helping us in the short term, but also improving the overall level of teaching as we look at the ways in which we communicate knowledge and information to our growing student body in HSES.

Sample topics from HSES Peer Evaluation of Teaching Observation Form

1. Organization/clarity:
 - Presented overview of lesson
 - Related lesson to past/present lessons
2. Presentation:
 - Used good examples to supplement
 - Emphasized important components
3. Instructor-group interaction:
 - Actively encouraged student questions
 - Asked questions to monitor learning
 - Provided adequate feedback
4. Content knowledge:
 - Presented material at appropriate level
 - Advanced knowledge level in the class

Transforming observations

Kellie Smith Herrod, Applied English Center

When being tasked as a classroom observer, we all quickly realize that traditional observations can take many forms and have many purposes. Over the years I have encountered the full range; for example, some observations are evaluative in nature, some are in place to support less-experienced instructors, while others are simply mandatory, and we are not really sure what the outcome is or how to lead conversations on instructor growth and teaching. For many observers, the most difficult piece is to identify the purpose of the observation.

In Ruth Wajnryb's book, *Classroom Observation Tasks: A Resource for Language Teachers and Trainers*, the author outlines specific observation tasks as a learning tool for instructors and observers. The categories of observation range from observations of the learner in the classroom, instructor use of language, learning environment, lesson, teaching skills and strategies, classroom management, and materials and resources. There is no evaluation piece. The focus is truly on professional growth through objective observation. Because the observations are focused on teaching, they can be done outside of one's discipline.

In my experience as an instructor observer, some of the most unexpected moments of growth have come in post-observation conversations. The conversation and reflection questions outlined by Wajnryb are focused and lead to surprising insights. They even worked well when tasked to observe our unit's most seasoned, awarded teachers. When the observation is transformed into an objective collection of data, followed by reflective writing and collegial discourse, teaching growth happens.

Evaluating GTAs in the Spanish Language Program

Amy Rossomondo, Spanish & Portuguese

All instructional staff in the Spanish Language Program participate in a rigorous and iterative class observation program. New instructors observe experienced instructors and their own peer group throughout their tenure in the program. The program director observes new instructors mid-way through their first semester and again at the end of the first semester; in subsequent semesters the course coordinators observe GTAs and lecturers at least annually and often each semester with supervision from the program director. All observations are followed up with structured conversations that promote self-

reflection to identify areas for improvement and to develop a plan of action toward that end.

In addition to the class observation program, GTAs complete an annual self-evaluation form (excerpts below) that serves as a guided reflection on their performance and goals. It is modeled on the procedure that is in place for faculty. The GTAs reflect on their successes and challenges, acknowledging specific areas for improvement based on their student course evaluations and the feedback they have received from their coordinators both on in-class performance and the degree to which they have been helpful

members of their instructional team each semester (meeting grading deadlines, professional interactions with their students, timely completion of assigned tasks, general collegiality, etc.). GTAs also have the opportunity to request support from the director and coordinators and make suggestions to improve the course or the program in general.

We try to create an environment in which teamwork is valued and individual growth is supported in service to undergraduate students. We also seek to instill a sense of purpose and pride in their contributions as GTAs become professionals.

Graduate Teaching Assistant (GTA) Self-Evaluation

1. Teaching expectations: prepares well for all class sessions; provides instruction with accurate information and teaches course content with enthusiasm; monitors and evaluates student progress and assignments; develops and organizes relevant teaching materials; utilizes approved texts or other instructional materials; administers examinations; provides opportunities for student engagement; maintains effective working relationships with students; grades and returns students' work in a timely manner
2. Quality of work expectations: completes work thoroughly, accurately and on time; pays attention to details; is well-organized; maintains a consistent and high level of performance; maintains standards of excellence in evaluation of student performance; creates instructional materials consistent with learning outcomes and course goals; follows federal law concerning confidentiality and academic and privacy rights of students (e.g., FERPA).
3. Accountability and self-management expectations: demonstrates organizational and time management skills; holds office hours at specified time and location; displays a team-oriented attitude; regularly attends course staff meetings and other related events as specified by supervisors; maintains communication with supervisor for support and guidance; responds promptly to requests of supervisors (e.g. submission of grades, meetings to review work); takes responsibility for all aspects of work; willing to ask for help and to help others; adheres to course outline and schedule on the syllabus; keeps clear records of all evaluations of student work.

Teaching rubric project kicks off

Doug Ward

CTE/Journalism & Mass Communications

A peer review of teaching generally goes something like this:

An instructor nears third-year review or promotion. At the request of the P&T committee, colleagues who have never visited the instructor's class hurriedly sign up for a single visit. Sometimes individually, sometimes en masse, they sit uncomfortably among wary students for 50 or 75 minutes. Some take notes. Others don't. Soon after, they submit laudatory remarks about the instructor's teaching, relieved that they won't have to visit again for a few years.

If your department or school has a better system, consider yourself lucky. Most peer evaluations lack guidelines that might offer meaningful feedback for a candidate, and they focus almost exclusively on classroom performance. They provide a snapshot at best, lacking context about the class, the students or the work that has gone into creating engagement, assignments, evaluations and, above all, learning.

Those peer evaluations don't have to be a clumsy and awkward free-for-all, though. Through the Benchmarks for Teaching Effectiveness Project, we have begun a process intended to make the evaluation of teaching much richer and more meaningful. The project started with the selection

of four departments and one school that will pilot the use of a rubric we have developed that delineates and weights several aspects of teaching. (My colleague Andrea Greenhoot writes more about the project starting on p. 2 of this newsletter.) Faculty members in those units will work with colleagues to define and identify the elements of good teaching in their discipline, decide on appropriate evidence, adapt the rubric,

By using the rubric to foster discussion about the central elements of teaching and then testing it in a variety of circumstances, instructors will learn valuable information about the teaching process.

apply it in some way, and share experiences with colleagues inside and outside the department and the university.

Not surprisingly, the instructors involved in the project had many questions about how the process might play out as they gathered for the first time in February: What types of evidence are most reliable? How do we reduce conscious or unconscious bias in the evaluation process? How do we gain consensus among colleagues for an expanded evalua-

tion process and for application of a new system of evaluation? How do we create a meaningful process that doesn't eat up lots of time?

Those are important questions without simple answers, but the departments that have signed on in this initial stage of the project have already identified many worthy goals. For instance, Sociology, Philosophy and Biology hope to reduce bias and improve consistency in the evaluation process.

Chemical and Petroleum Engineering plans to create triads of faculty members to provide feedback on teaching to each other. Public Affairs and Administration sees opportunities for enriching teaching and inspiring instructors to take risks to innovate teaching.

All the units will use the rubric to foster discussion among their colleagues, identify trustworthy standards of evidence, and, ultimately, evaluate peers. Philosophy sees opportunities for better evaluating graduate teaching as-

sistants. Chemical and Petroleum Engineering hopes to use the rubric to guide and evaluate 10 faculty members on tenure track. Sociology plans to use it to guide peer evaluation of teaching. Public Affairs and Administration plans to have a group alternate between evaluator and evaluatee as they hone aspects of the rubric. Biology plans to explore the best ways to interpret the results.

That range of activities is important. By using the rubric to foster discussion about central

elements of teaching and then testing it in a variety of circumstances, instructors will learn valuable information about the teaching process. That feedback will allow us to revise the rubric, create better guidelines for its use, and ultimately help as many departments as possible adopt it for the promotion and tenure process.

All the faculty members working in the initial phase of the Benchmarks project recognize the complexity and challenge of high-quality teaching. They also recog-

nize the challenges in creating a better system of evaluation. Ultimately, though, their work has the potential to make good teaching more transparent, to make the evaluation of teaching more nuanced, and to make teaching itself a more important part of the faculty evaluation process.

A summary of the rubric is below. The complete rubric can be found at https://cte.ku.edu/sites/cte.ku.edu/files/docs/KU_Rubric_for_Evaluating_Teaching_May_2017.pdf

Benchmarks Project: Categories and Questions for Evaluation of Teaching



GOALS, CONTENT, & ALIGNMENT

What are students expected to learn from the courses taught?
Are course goals appropriately challenging?
Is content aligned with the curriculum?



TEACHING PRACTICES

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



ACHIEVEMENT OF LEARNING OUTCOMES

What impact do these courses have on learners?
What evidence shows the level of student understanding?



CLASSROOM CLIMATE & STUDENT PERCEPTIONS

What are the students' views of their learning experience?
How has student feedback informed the faculty member's teaching?



REFLECTION & ITERATIVE GROWTH

How has the faculty member's teaching changed over time?
How has this been informed by evidence of student learning?



MENTORING & ADVISING

How effectively has the faculty member worked individually with undergraduates or graduate students?



INVOLVEMENT IN TEACHING COMMUNITY

In what ways has the instructor contributed to the broader teaching community, both on campus and off campus?

Graphic designed by Kali Jo Wolkow

END NOTE

COPUS documents how instructors use class time

The Classroom Observation Protocol for Undergraduate STEM (COPUS) is an adaptation of the Teaching Dimensions Observation Protocol that was developed at the University of British Columbia (Smith et al., 2013). This protocol documents classroom behaviors in two-minute intervals throughout a class period. These behaviors encompass activities by both instructors and students, with 25 categories captured (see below).

At KU, we began using the COPUS protocol in Spring 2016 through the TRESTLE project as a way to measure change in instructive practices by faculty in participating STEM departments. Observations are completed over three class periods a semester, and a report is generated that is provided to the instructor. Reports include a breakdown of activities across time in a course, in addition to the percent of time each activity is used throughout the class period.

Since beginning this project, the number of COPUS observations across campus has grown each semester (22 courses in Spring 2016 to over 40 in Spring 2018). Feedback from instructors has been positive; for example, many state that it has been helpful to see how class time is being used.

COPUS observations are available for STEM courses at all levels. For more information, contact Blair Schneider at blair@ku.edu, or check trestlenetwork.org/copus-observation-resources/

Students are doing

L	Listening to instructor / taking notes, etc.
AnQ	Student answering a question posed by instructor with rest of class listening
SQ	Student asks a question
WC	Engaged in whole class discussion
SP	Presentation by student(s)
Ind	Individual thinking / problem solving
CG	Discuss clicker questions in groups of 2 or more students
WG	Working in groups on worksheet activity
OG	Other assigned group activity
Prd	Making a prediction about the outcome of a demo or experiment
TQ	Test or quiz
W	Waiting (instructor late, working on fixing AV problems, etc.)

Instructor is doing

Lec	Lecturing
RtW	Real-time writing on board, document projector, etc.
D/V	Showing or conducting a demo, experiment, simulation, video or animation
Fup	Follow-up / feedback on clicker question or activity to entire class
PQ	Posing non-clicker question to students
CQ	Asking a clicker question
AnQ	Listening to and answering student questions with entire class listening
MG	Moving through class guiding ongoing student work during active learning task
1o1	One-on-one extended discussion with 1-3 individuals
Adm	Administration (assigning homework, return tests, etc.)
W	Waiting when there is an opportunity for an instructor to be interacting with or observing / listening to student or group activities and the instructor is not doing so.)

Smith, M.K., Jones, F.H.M, Gilbert, S.L, and Wieman, C.E., 2013, The Classroom Observation Protocol for Undergraduate STEM (COPUS): A New Instrument to Characterize University STEM Classroom Practices: *CBE Life Sciences Education*, v. 12, p. 61

Teaching Matters is published by CTE and edited by Judy Eddy. We welcome your comments and suggestions.

KU Center for Teaching Excellence
1455 Jayhawk Blvd., Room 135
Lawrence, KS 66045-7573
785-864-4199
cte@ku.edu
www.cte.ku.edu