

# Draft Report of the Provost's Task Force on Assessing Teaching Effectiveness

Date: November 12, 2020

## Task Force Members

Amy W. Ando, **ACES** (*Chair*)

Timothy Bretl, **ENG**

Tara Earls Larrison, **SSW**

Karle Flanagan, **LAS**

Kim Graber, **AHS**

David Hays, **FAA**

Angela Kent, **ACES**

Denise Loyd, **BUS**

Shachar Meron, **COM**

Karla Moller, **EDUC**

Michelle Nelson, **COM**

Steve Petruzzello, **AHS**

Robert Rushing, **LAS**

Sharon Shavitt, **BUS**

Timothy Stelzer, **ENG**

Maryalice Wu, **CITL**

Ann Yeung, **FAA**

James Yoon, **LAS**

Direct questions and comments to Amy Ando at [amyando@illinois.edu](mailto:amyando@illinois.edu)

## Table of Contents

Executive summary of findings and recommendations	p. 3
1. Introduction	p. 4
1.1 <i>Task Force charge and goals</i>	p. 4
1.2 <i>Process</i>	p. 4
2. Relevant research and innovations at other universities	p. 6
2.1 <i>Research and calls for change</i>	p. 6
2.2 <i>Innovations at other universities</i>	p. 8
3. Recommendations for practice of evaluating teaching at UIUC	p. 9
3.1 <i>Establish criteria defining quality of teaching performance</i>	p. 11
3.2 <i>Gather inputs for teaching evaluation from three voices</i>	p. 11
3.2.1 Gather input from students	p. 12
3.2.2 Gather input from peers	p. 16
3.2.3 Gather input from instructors themselves	p. 18
3.3 <i>Holistic evaluation based on multiple inputs</i>	p. 20
4. Recommendations for functions that draw upon teaching evaluation	p. 22
5. Implementation process	p. 23
6. Topics for future work	p. 24
Appendix 1: Charge Letter from Provost Cangellaris	p. 25
Appendix 2: Annotated Bibliographies and Publications	p. 26
Appendix 3: Sample Innovations at Other Institutions	p. 29
Appendix 4: Draft Campus-Wide Definition of Teaching Excellence	p. 30
Appendix 5: Draft Student Feedback on Courses and Instruction	p. 32
Appendix 6: Draft Peer Review of Instruction	p. 35
Appendix 7: Draft Peer Review of Course Materials and Design	p. 37
Appendix 8: Sample Guidance for Peer Evaluation of Teaching	p. 39
Appendix 9: Draft Self-Reported Annual Inventory of Teaching Contributions	p. 42
Appendix 10: Draft Personal Self-Evaluation of Instruction	p. 43
Appendix 11: Draft Formal Self-Evaluation of Instruction	p. 45
Appendix 12: Draft Single-Year Holistic Evaluation of Instruction	p. 47

## Executive Summary of Findings and Recommendations

- Findings regarding current practice:
  - Bias is a chronic problem in traditional approaches to teaching evaluation, especially simple numeric student ratings of overall teaching and course quality.
  - Overall student evaluations of teaching or course quality on a simple numeric scale provides limited useful information. General ratings in response to vague prompts do not shed light on how well a course or instructor have met objective criteria for teaching quality, and do not help instructors to improve their work.
  - Bias in teaching evaluation inputs like student surveys makes comparing instructors to one another on the basis of those inputs problematic.
  - Current teaching evaluation practice at UIUC is inconsistent across units, and even across instructors within a unit.
  - The current system of teaching evaluation at UIUC does not take advantage of modern best practices in this activity.
  
- UIUC has the opportunity to develop a new system for teaching evaluation with a set of enhanced features. The new system can:
  - Use design to minimize bias.
  - Improve our capacity to enhance teaching quality, troubleshoot efficiently, celebrate multiple dimensions of teaching excellence, and minimize bias in personnel decisions for instructors.
  - Minimize time costs and administrative burdens to encourage high levels of engagement and compliance.
  
- Recommended features of the new system of teaching evaluation are:
  - Establish a set of campus-wide criteria that define criteria for high-quality instruction and course design.
  - Use routine and systematic input from multiple voices (students, peers, and self) in holistic reviews to evaluate an instructor's progress towards meeting those criteria.
  - Compare an instructor's performance to the articulated criteria defining teaching quality and chart their growth over time.
  - Permit units to tailor evaluation instruments to accommodate the diverse teaching needs and formats across campus.

## 1. Introduction

### 1.1 Task Force charge and goals

The University of Illinois Urbana-Champaign (UIUC) works to provide an excellent education to its students. Just as scholars review one another’s research to make the work stronger, teaching evaluation serves many important functions for advancing the quality of education at UIUC. A robust and unbiased system of teaching evaluation helps all instructors to hone their craft and keep their work up to date; guides hiring, promotion, and tenure decisions; and informs merit raises and awards that recognize excellence in teaching.

Recent innovations at other institutions and research both show opportunities to improve how we evaluate teaching at UIUC. Many different types of professionals make important contributions to teaching at UIUC, such as tenure-track faculty, non-tenure-track faculty, academic professionals, and graduate employees; our system of teaching evaluation can work to support all teaching professionals on campus. We can make teaching better, have more streamlined teaching evaluation processes, and support personnel functions that are more efficient and equitable.

The charge from Provost Cangelaris to the Provost’s Task Force on Assessing Teaching Effectiveness (“the Task Force”) on November 20, 2019 is provided in Appendix 1; the Task Force has worked to follow that charge.

### 1.2 Process

Table 1 shows a timeline of the process that produced this report. In Fall 2019, the Chair of the Task Force conducted foundational research about current practices at UIUC, recent research on teaching evaluation practices, and changes in practice at other universities. She also consulted stakeholders across campus to develop ideas about appropriate goals for a new system. The Provost constituted and charged the Task Force at the end of Fall 2019. The Task Force represents multi-dimensional diversity at the university, with members from nine colleges, the Center for Innovations in Teaching and Learning (CITL), the Teaching Advancement Board (TAB), and the Faculty Senate.

In Spring 2020, the Task Force agreed on the goals and general structure of a new teaching evaluation system at UIUC and created three subcommittees that drafted broad recommendations for how inputs should be gathered from students, peers, and instructors. The Task Force paused after the arrival of COVID-19. Members developed new ideas for how a teaching evaluation system should be developed and function in a time when many more classes are online, and all instructors and administrators have even less capacity for administrative functions (like evaluations) than before.

Over the summer of 2020, the Task Force developed actual drafts of the critical elements of a new teaching evaluation system: a definition of teaching excellence with concrete criteria against which performance can be compared; draft instruments for gathering inputs for evaluation from students, peers, and instructors themselves; and recommendations for how and when input should be gathered and used. These elements have been collected in this Draft Final Report for circulation and feedback from the broader campus community at UIUC.

**Table 1: Timeline of Task Force Activity**

<b>Dates</b>	<b>Activity (CH=Chair of Task Force; TF=Task Force)</b>
Aug – Nov 2019	<ul style="list-style-type: none"> <li>- CH conducts foundational research on recent scholarship and practice (see Section 2.1)</li> <li>- CH consults with members of campus community: Provost’s Office; Center for Innovation in Teaching and Learning (CITL); the Teaching Advancement Board (TAB); the Graduate College; Associate Deans in the colleges of ENGR, LAS, AHS, ACES, BUS, FAA, and EDUC.</li> </ul>
Nov – Dec 2019	<ul style="list-style-type: none"> <li>- TF is recruited, has initial charge meeting with Provost</li> <li>- TF identifies strengths and weaknesses of current UIUC system</li> </ul>
Jan 2020	<ul style="list-style-type: none"> <li>- TF meets as a whole</li> <li>- TF agrees on goals and general structure of new system</li> <li>- TF creates three subcommittees to study inputs from self, students, peers</li> <li>- CH provides update on project to Council of Deans; gains feedback</li> </ul>
Feb – March 2020	<ul style="list-style-type: none"> <li>- TF subcommittees draft plans for three systems to gather inputs</li> <li>- CH consults with College of Media Executive Committee</li> <li>- CH planned consultation with Provost Undergrad Student Advisory Board; was cancelled due to pandemic</li> </ul>
April 2020	<ul style="list-style-type: none"> <li>- Regrouping after COVID-19</li> </ul>
May – July 2020	<ul style="list-style-type: none"> <li>- TF develops core definition of teaching excellence at UIUC</li> <li>- CH gains feedback on definition of teaching excellence from Provost’s office, CITL, Associate Deans, TAB, students</li> <li>- TF drafts new instruments for gathering inputs from students, peers, instructors themselves</li> </ul>
August-Sept 2020	<ul style="list-style-type: none"> <li>- Draft Final Report written</li> <li>- Feedback solicitation begins</li> </ul>
Oct – Nov 2020	<ul style="list-style-type: none"> <li>- Circulated for feedback to: Office of the Provost; CITL; administrative offices of all UIUC Colleges, Chair of Campus Committee on Promotion and Tenure; Vice Chancellor for Diversity, Equity, and Inclusion; Senate Executive Committee; General University Policy Committee; Educational Policy Committee; Teaching Advancement Board; DRIVE Committee; Illinois Student Government; Provost’s Undergraduate Student Advisory Board.</li> </ul>

## 2. Relevant research and innovations at other universities

### 2.1 Research and calls for change

Many scholars study how best to evaluate teaching in institutions of higher education. The Task Force reviewed several recent strains of peer-reviewed research that have raised concerns about traditional models of teaching evaluation that rely almost exclusively on simple student surveys. **Those concerns have driven calls for change.** For example:

- A successful lawsuit at Ryerson University disallowed student evaluations of teaching in personnel decisions (Flaherty 2018).
- Faculty union negotiations at the University of Oregon led to a signed memorandum of understanding<sup>1</sup> to establish new policies for teaching evaluation that replace traditional simple student surveys.
- Carl Wieman, a Nobel-Prize winning physicist, uses critical reviews of the research on traditional teaching evaluation methods to inform his new approach to teaching evaluations in STEM fields (Wieman 2015).
- In 2019, the American Sociological Association and over a dozen other associations issued a statement (ASA 2019) calling for broad reform in the practice of teaching evaluation in colleges and universities. This statement provides a clear summary of the research on limitations of traditional student evaluations and provides a set of recommendations for change that are also based on research.

This Task Force report will not include a full review of the literature on teaching evaluations; readers can view the annotated bibliographies and articles listed in Appendix 2. Rather, this section highlights just two of the widely recognized main **concerns that are driving change in teaching evaluation practice.**

- Bias is a problem in traditional simple student evaluations of teaching (SETs). Many studies, including experiments designed to control for confounding factors, find that SETs with simple numeric ratings of overall teaching quality are biased against women in general, and against Black, Asian, and foreign instructors of all genders. As Wieman (2015) summarizes, “the data indicate that it would be nearly impossible for a physically unattractive female instructor teaching a large required introductory physics course to receive as high an evaluation as that of an attractive male instructor teaching a small fourth-year elective course for physics majors, regardless of how well either teaches.” The ASA (2019) report states:

---

<sup>1</sup> <https://provost.uoregon.edu/revising-uos-teaching-evaluations>

“In addition, in both observational studies and experiments, SETs have been found to be biased against women and people of color (for recent reviews of the literature, see Basow and Martin 2012 and Spooren, Brockx, and Mortelmans 2013). For example, students rate women instructors lower than they rate men, even when they exhibit the same teaching behaviors (Boring, Ottoboni, and Stark 2016; MacNell, Driscoll, and Hunt 2015), and students use stereotypically gendered language in how they evaluate their instructors (Mitchell and Martin 2018).... Further, Black and Asian faculty members are evaluated less positively than White faculty (Bavishi, Madera, and Hebl 2010; Reid 2010; Smith and Hawkins 2011), especially by students who are White men. Faculty ethnicity and gender also mediate how students rate instructor characteristics like leniency and warmth (Anderson and Smith 2005).”

In an example of one such study, Boring (2017) studies teaching evaluations in a French university setting with several features that provide an excellent opportunity to explore the possibility of bias in SETs. First, the university functionally randomizes the gender of the professors to whom students are assigned, and there are professors of different genders teaching the six mandatory courses taken by all first-year undergraduate students. This eliminates concern that students might self-select into certain classes or to learn from certain professors. Second, all students are required to complete SETs of their courses, so there is no sample selection in who responds. Third, all students in the same course take the same final exam, so one can compare a measure of learning that is common across instructors of different genders. Finally, the dataset includes the universe of students, faculty, and SETs over five years, with 20,197 individual SET Scores; this is not a small idiosyncratic sample. The study finds that satisfaction scores are not correlated with the exams that proxy for actual learning, and male professors receive statistically significantly higher overall average satisfaction scores (driven heavily by male students being more likely to rate them as “excellent.”)

Not all research on SETs finds evidence of bias. However, such findings are common enough to raise concern.

Simple SETs do not provide useful measures of teaching quality. Even if SETs were unbiased, other research casts doubt on how useful traditional SETs can be. If an instructor scores poorly on a five-point scale of general teaching effectiveness, that information does not help the instructor to improve. Furthermore, as stated in ASA (2019):

“SETs are weakly related to other measures of teaching effectiveness and student learning (Boring, Ottoboni, and Stark 2016; Uttl, White, and Gonzalez 2017); they are used in statistically problematic ways (e.g., categorical measures are treated as interval, response rates are ignored, small differences are given undue weight, and distributions are not reported) (Boysen 2015; Stark and Freishtat 2014); and they can be influenced by course characteristics like time of day, subject, class size, and whether the course is required, all of which are unrelated to teaching effectiveness.”

For example, Uttl, White, and Gonzalez (2017) re-analyzed data from previously-published studies comparing SET scores and measures of student learning in courses with multiple sections and the same final exam; they found that correcting for small-

sample bias and publication bias eliminated any correlation between SET scores and student learning.

Other research has informed innovations in modern systems of teaching evaluation. That work is reflected in the innovations at other universities and the recommendations for a new system at UIUC, below.

## 2.2 Innovations at other universities

In response to research calling traditional approaches to teaching evaluation into question, many colleges and universities in the U.S. are transforming their systems of teaching evaluation. For example:

- The University of Southern California was an early pioneer in overhauling its system of teaching evaluation.
- The University of Oregon created a new multi-voice teaching evaluation system oriented around a set of teaching expectations.
- The National Science Foundation has funded the “Transforming Higher Education - Multidimensional Evaluation of Teaching” project (TEval) in which the University of Kansas, the University of Massachusetts at Amherst, and the University of Colorado Boulder are co-creating a new holistic system for teaching evaluation, and Michigan State University is analyzing the change process at those three institutions.

Figure 1 shows the research universities the Task Force studied most comprehensively and drew on in formulating recommendations for UIUC. We note that other institutions of higher education are also engaging in similar innovation.

**Figure 1: Sample Research Universities Changing Teaching Evaluation**



Appendix 3 provides websites for our exemplars with detailed information about their work. Here we identify a set of common themes in transformed teaching evaluation system at research universities.

**Innovations in teaching evaluation systems include:**

- Gather regular and systematic input on teaching quality from multiple sources: students, peers, instructors themselves.
- Work to reduce bias in the inputs, such as student surveys, used for teaching evaluation.
- Drop or de-emphasize student evaluations of overall instruction or course quality on a simple numeric scale.
- Use instruments that elicit feedback on teaching quality from all three sources by asking about specific features of teaching design and implementation that are under instructors' control, and that are aligned with elements of instruction that research indicates are effective and important.
- Use simple metrics and rubrics to increase administrative efficiency.
- Gather small amounts of focused qualitative input to capture features of teaching beyond metrics.
- Allow different units of an institution to tailor teaching evaluation instruments to accommodate diverse teaching formats.
- Design administrative functions that use teaching evaluation inputs from students, peers, and instructors themselves to guard against biased outcomes because there is no way to eliminate bias in the inputs entirely.
- Do not compare instructors to one another.
- Do evaluate whether an instructor has achieved objective benchmarks in performance and how they are improving over time.

### **3. Recommendations for practice of evaluating teaching at UIUC**

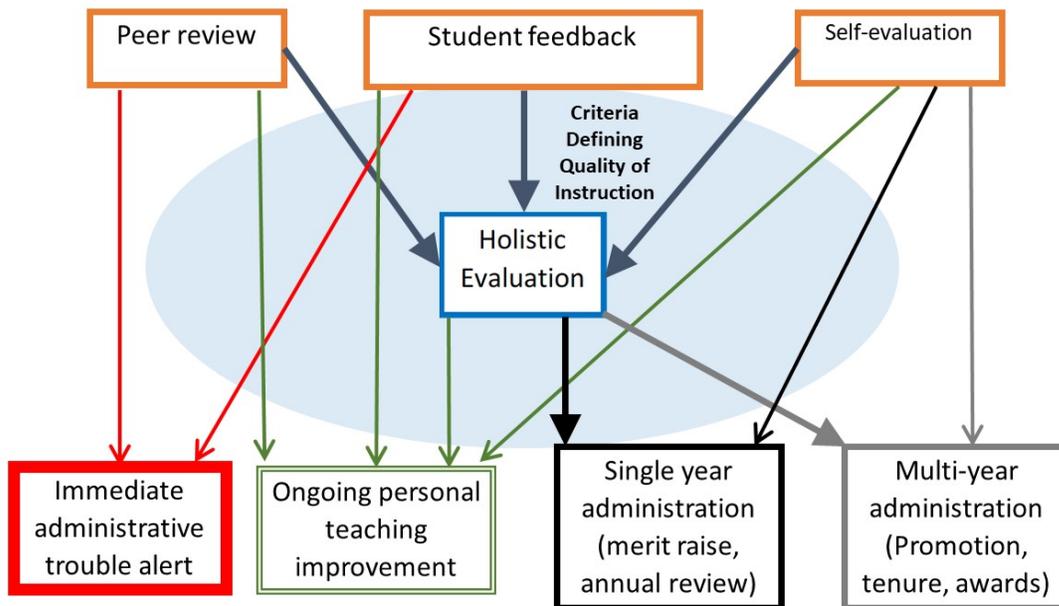
UIUC already draws upon input from students, peers, and instructors themselves to gather information for teaching evaluation. However, only student ratings are carried out campus wide in a systematic fashion, and the university has no definition of teaching excellence against which a course or instructor can be objectively compared. The Task Force recommends a set of changes in the practice of teaching evaluation at UIUC, each of which is further detailed in a section below.

In general, we recommend that the university should establish a set of criteria that define excellent instruction at UIUC, and then gather concrete and routinized feedback and evaluation from students, peers, and instructors themselves to regularly chart the progress made by course designers and instructors toward accomplishing those standards of excellence. Those inputs will be used in holistic reviews that use input from multiple sources to chart and

celebrate an instructor’s progress in achieving excellence in multiple dimensions of teaching quality.

The teaching evaluation system informs a range of functions: self-improvement; trouble alerts; annual personnel functions like merit raises and contract renewals; and multi-year personnel functions like tenure and promotion decisions and awards. Trouble alerts can be triggered quickly by student feedback or peer reviews (though administrators should gain more information before taking action on the potential issue that was flagged by the alert). Personal teaching improvement can be informed by all individual inputs as well as holistic evaluations. Formal administrative actions should only be informed by holistic evaluations and formal instructor self-evaluation. A schematic of this system is shown in Figure 2.

**Figure 2: Schematic of New Teaching Evaluation System**



This new system will address the limitations of our current system in two general ways. First, the system will provide more consistent, actionable, and useful evaluations of teaching design, instruction, and activity. Student feedback and peer evaluations will be more consistent across campus, and will provide clear assessments of progress towards campus expectations on concrete measures of teaching quality that are in the control of instructors. Second, the system will reduce bias in the teaching evaluation system and decisions that draw upon those evaluations. Student and peer assessments of teaching will be focused on specific performance criteria and thus less influenced by implicit bias than current approaches. The self-evaluation process will provide a structured format to document and reward previously under-appreciated labor such as informal mentoring. Personnel decisions will be made on the basis of holistic

evaluations that draw on three voices rather than one, mitigating the influence of bias in any one measure and giving instructors themselves regular and systematic opportunity to provide context for potentially biased assessments. Finally, the system of teaching evaluation will not compare instructors directly to each other on metrics that are likely biased against some instructors, but rather evaluate each instructor in terms of their growth towards accomplishing campus standards of excellence.

### 3.1 Establish criteria defining quality of teaching performance

- Current practice: UIUC currently has no campus-level definition of the features that characterize high quality teaching performance.
- Recommended new system: The Task Force has drawn on research and models from several other universities to develop a **Campus-Wide Definition of Teaching Excellence at the University of Illinois Urbana-Champaign**; see Appendix 4. This definition is broad enough to guide assessment of teaching across our diverse campus. At the same time, the definition provides concrete criteria for what determines the quality of instruction: contributions outside course-based teaching should be valued, and course-based teaching should be well-designed, well-delivered, inclusive and ethical, reflective and evolving. Thus, assessment of teaching can be aligned to specific performance criteria and thus be less influenced by implicit bias.

The Definition stipulates that excellent course teaching is well-designed, well delivered, inclusive and ethical and reflective and evolving. The Definition is flexible enough to be relevant for the diverse range of teaching activities at the University. However, it is critical that the term “learning objectives” be considered broadly. Well-designed teaching creates clear learning objectives and well-delivered teaching accomplishes those objectives; however, objectives need not be factual content to be conveyed. Objectives could include providing students with opportunities to engage with the public, exposing them to multiple disciplinary perspectives on a problem, or any number of other goals that may be hard to quantify.

### 3.2 Gather inputs for teaching evaluation from three voices

#### **Best practice in teaching evaluation gathers input from students, peers, and the instructor.**

Gathering input from three sources mitigates the impact of bias from any one source. This strategy also gathers information from sources that have different areas of expertise and experiences with the course design and instruction being evaluated. For example, students are uniquely positioned to know how an instructor interacts with them over the course of the semester. Peer instructors are experts in the quality and currency of the course material being taught. And only the instructor and course designer can explain what challenges they had to

overcome with students or instructional technology, and how they are using feedback and experimentation to improve course design and/or instruction over time.

**The Taskforce has drafted instruments for all three types of input collection.** These draft instruments are aligned to measure teaching effectiveness according to the criteria set out in the draft “Campus-Wide Definition of Teaching Excellence at the University of Illinois Urbana-Champaign.” Many details of the actual instruments used will be refined; for example, the systems for student and peer evaluation should eventually have the capacity for departments and instructors themselves to add a few questions at the end. However, these drafts provide concrete examples of the new processes.

**Input collection instruments will be piloted and refined before full implementation.** These instruments will be refined in consultation with faculty, students, administrators, and pedagogical experts at CITL. All three instruments should be programmed to be completed in online surveys or webforms to facilitate efficient and uniform collection, storage, and communication of information. Design experts at UIUC should be consulted to produce forms that are clear and easy to fill out.

### ***3.2.1 Gather input from students***

Student input plays a vital role in teaching evaluation systems. UIUC has a tradition of sharing governance between faculty, students, and administrators; giving students voice in teaching evaluations is consistent with that tradition. Students are uniquely positioned to provide information about some elements of an instructor’s teaching performance. More practically, it is relatively straightforward and cost-effective to collect information from students for every course offered. Below we list features of how student input is currently collected and used at UIUC, and offer recommended changes in those practices with explanations for each one.

#### Current practice at UIUC:

- Instructors administer student rating surveys through “Instructor and Course Evaluation System Questionnaires” (ICES) forms.
- Student ratings are not mandatory for all courses. In some units, student ratings are carried out only if the instructor chooses to do so.
- Surveys are carried out in a mixture of paper and online formats.
- Student rating results are not released to anyone but the instructor without instructor permission.
- Faculty may choose from over 400 different questions to include on the surveys for their classes. Many units have their own sets of questions, though not all sets are actively curated.

- Only two summary (or “global”) ratings of quality (on a scale of one to five) for the instructor and the class are used in formal tenure and promotion proceedings, with indicators of how an instructor’s ratings compare to others on campus.
- Instructors who have average global ratings above a given bar are listed on the “List of Instructors Rated by their Students as Excellent.”

Suggested changes and practices for gathering student input for teaching evaluation:

(See Appendix 5 for draft new student feedback survey)

- Campus will refer to “student feedback on courses and instruction” rather than “student evaluations.”
  - This change will better reflect the fact that students alone are not evaluating instruction.
- Student feedback will be solicited automatically for all courses, instructors may not choose to forego student feedback.
  - This change will ensure that student feedback is available for all courses, to help unit EOs identify areas of excellence and problems that might need to be addressed.
- Student feedback surveys will be conducted online for all courses.
  - This change will improve consistency in feedback practice across campus, be feasible for all courses regardless of instructional mode, reduce administrative costs, and provide results to instructors in a timely manner.
- Student feedback results will be immediately and automatically available to the instructor, their unit EO, and other university program administrators with a need for the information. This information should otherwise be a private part of instructors’ personnel files.
  - This change will help EOs to carry out more well-informed performance reviews of instructional personnel and facilitate instructional program review and improvement.
- A core set of student feedback questions will be administered uniformly to all courses. In final form, each department may add a small number of additional questions to the feedback instrument, and instructors should be able to add a few questions themselves.
  - This set of practices will provide important consistency in evaluation practice across campus, while still providing opportunities for units and instructors to gain information tailored to their particular needs.
- Student feedback surveys will not include summary 1-5 ratings to “Rate the instructor’s overall teaching effectiveness” or “Rate the overall quality of this course.”
  - This change will improve the validity of our teaching evaluations because those ratings are particularly subject to bias and do not provide meaningful input for teaching assessment on the basis of agreed-upon criteria.
- Campus policy should be that an instructor’s feedback results will be viewed relative to articulated standards or teaching quality and in the context of their own past results, not in comparison to other instructors.

- This change is consistent with recommendations from research on teaching evaluation practices, and will reduce the harm done to instructors by biased interpersonal comparisons.
- The report of results will include information about the number and full distributions of responses to individual categorical feedback questions in addition to the means.
  - This practice will allow instructors and EOs to identify if there is a problem that only a few students experience. It will also provide valuable context for results, such as low response rates that render some averages unreliable, or bi-modal responses that tell a very different story from a moderate mean.
- Appendix 5 has the draft core student feedback survey being proposed. This survey has the following features:
  - Design that minimizes bias. This features questions about specific features of instruction that are under the instructor’s control and may include an anti-bias statement at the beginning (Peterson et al. (2019) find evidence that such a statement may help mitigate bias.)
  - Questions that elicit student feedback on features of instruction that align directly with UIUC’s new definition of teaching excellence.
  - Questions, which if answered negatively, should trigger the EO to take further action to identify whether there is a serious problem with instruction in the class.
  - Questions that identify how much engagement the students themselves had with the class.
  - A few spaces for open-ended comments from the students in response to structured prompts. This includes several universal prompts and could allow space for several added by the instructor.

Notes on implementation:

- Immediate critical needs:
  - Refine campus core student feedback survey: The taskforce will work with others to refine the core survey.
  - Refine system of student feedback survey for lab or discussion sections of larger classes. Additional questions might probe whether these sections align well with the class as a whole. The questions should work well even if the auxiliary instructors do not control the content of the sections they teach.
  - Develop optimized mechanism for electronic administration: Best practices will be used to maximize response rates and representativeness. UIUC should draw on research (e.g. Spooren and Van Loon 2012; Goodman et al. 2014) and consider a range of policies (e.g. extra credit) or processes (e.g. dedicated day for student feedback surveys) to accomplish this goal. Care will need to be taken with special situations such as courses in hospital settings and studio teaching where many students are taught in small separate sections.
  - Develop system for secure storage and communication of results.

- Next steps:
  - Recognizing accomplishment: Develop a replacement for the old “List of Faculty Rated by their Students as Excellent” that can be reported for individual instructors and classes each term. This recognition each term could, for example, draw on responses to the question that asks “Do you think your instructor’s teaching was excellent?” This system of recognition could also highlight instructors that had high levels of positive feedback in questions related to one of the elements of teaching excellence (“Inclusive and ethical” or “Well-delivered”) and courses could be recognized for good feedback related to being “Well-designed.”
  - Unit additions: During the first semester of implementation, units may identify up to five additional questions they wish to add to the survey instrument for classes in their departments. Those will be added in future semesters.
  - Develop longitudinal visualized profile: Develop a data visualization approach to represent the longitudinal profile of an instructor’s multi-dimensional student feedback results over time. This could take a form such as that used by the OECD for its cross-country quality of life ratings (see Figure 3).
  - Reporting: Develop system for aggregate campus reports about student feedback.
  - Individual questions: Build capacity for individual instructors to add a few idiosyncratic questions to the end of the survey intended to help them to improve their teaching.
  - Processing qualitative answers: Consider using machine learning to synthesize responses to open ended questions.

**Figure 3: Sample visualization of multi-dimensional ratings**



Note: The OECD Better Life Index is a good example of how multi-criteria data could be visualized. This website compares countries to each other; UIUC would track an index for a single instructor over time.

<http://www.oecdbetterlifeindex.org/#/1111111111>

### **3.2.2 Gather input from peers**

Peer evaluation plays an important role in teaching evaluation. Other professionals bring expertise in content and pedagogy to evaluation that is different from the expertise of students and important for formal processes like tenure and promotion. Continued mutual observation and assessment of each other's instructional activities also promotes a culture of ongoing improvement of instruction.

#### Current practice at UIUC:

- Most units have occasional peer review of course-based teaching for junior faculty, but both frequency and practice vary widely. Some units do no meaningful evaluation of teaching, others have systematic evaluation of materials and visits to the class carried out regularly with systematic rubrics.
- P&T and promotion dossiers: Units designate someone to write an evaluation of the candidate's teaching. The thoroughness of those evaluations is extremely variable.

#### Suggested changes and practices for formal peer evaluation practice:

(See Appendices 6, 7, and 8 for draft rubrics and guidance for peer review of instruction and course design)

- Every instructor with a regularly recurring contract should be regularly reviewed by a peer, or by two peers who work as a team.  
-Peer evaluation provides an important lens to complement feedback from students.
- "Peers" who conduct teaching evaluations could be chosen by units to be another faculty member in that department, an instructor's teaching mentor, a campus pedagogical expert, or an instructor in a different department with appropriate expertise.  
-Allowing flexibility in this role reduces the time cost of peer evaluation on faculty, especially those in small units.
- Practices for peer evaluations of individual courses:
  - The peer and instructor (and course designer if not the same as instructor) complete a brief pre-meeting prior to visiting the class. They discuss the observation and review the rubrics to be used for evaluation.  
-Note: Allowing separate evaluations of course design and instruction allows credit to flow to course designers who do not teach the courses, and makes more sense when evaluating instructors (often graduate students) who did not design the course they teach.
  - The peer formally reviews course materials with a simple rubric designed to capture a large amount of feedback in little time.

- The peer observes at least two classes or equivalent online course modules to formally review instruction with a rubric designed to capture a large amount of feedback in little time.
- Each rubric has a small section for comments from the peer to the instructor that are confidential.
- The peer meets with instructor and course designer to review observations
- The peer completes a formal report (two rubrics); all but the confidential comments are submitted to the unit EO and part of personnel files.
- Recommended frequency: More frequent review could be conducted as warranted by the unit. For example, specialized faculty may benefit from more frequent peer review to support merit raises, or concern about a tenured faculty member’s teaching may warrant a period of more frequent review until the issue is addressed. Below, however, are baseline recommendations.
  - Faculty at the Assistant rank, lecturers, and instructors: one time every year teaching.
    - This regular evaluation captures growth in course design and instruction during the pre-promotional period.
  - Faculty at the Associate rank: once every other year.
    - Continued semi-regular review keeps building a record of growth to support the case for promotion to Full.
  - Full faculty: once every five years.
    - Ongoing peer review of full faculty can support merit raises and awards, and encourages continued teaching improvement throughout faculty careers.
- Practices for third-year reviews, promotion and P&T dossiers:
  - Multi-year data visualizations will show how peer ratings for elements of their instruction and their course design have evolved over time.
  - Multiple years peer evaluations are included in multi-year holistic evaluations. See Section 3.3 for details.

Notes on implementation:

- Immediate critical needs:
  - Refine baseline rubric and guidance sheets: The baseline rubric will be general enough to be relevant across campus and simple enough to use well without extensive training.
    - Evaluation of single course materials
    - Evaluation of single course instruction
  - Develop guidance sheets to give detailed guidance for evaluating the criteria in the rubric for different kinds of courses, e.g. face-to-face vs. online, lecture vs. discussion, lab science vs. studio art, etc.

- Develop peer review training documentation: EOs and initial participating peer evaluators must have guidance to evaluate the teaching effectiveness of others. Initial guidance for carrying out peer reviews will be through written materials.
- Provide campus support to departments for peer review: Some units may have trouble finding enough faculty well-positioned to carry out peer reviews. Campus administration should facilitate cross-department peer matching and hire additional instructional review experts who can serve as peers for this purpose.
- Next steps:
  - Units refine the guidance documents for the types of courses their instructors teach.
  - Units set forth policies for peer-review labor: Units should establish clear guidelines for who can and should serve as peer reviewers for instructors in the unit and how that work is recognized and rewarded. Practice will necessarily vary depending on factors such as how specialized teaching is and how many faculty there are in the unit. Guidelines should avoid having instructors reviewing other instructors who could be in a position to evaluate them in the future for promotion, tenure, reappointment, or other personnel matters.
  - CITL develops peer-review training modules for various assessment criteria. Modules can be customized for the various instructional approaches imperative to each college or program
  - Widespread training: Campus should provide streamlined opportunities for EOs and faculty (tenure-stream and specialized) at the associate and full levels to learn to review the teaching effectiveness of other instructors.
  - Develop multi-year visualized peer-review profile: UIUC should have a system like that used for student survey results that visualizes the pattern of results for a given instructor over time. Ratings for elements of instructional quality can be in one visualization, and ratings for elements of course design can be in another.
  - Graduate and adjunct instructors: Units and the campus should have further discussion about professional evaluation of graduate students and other instructors who do not have regularly recurring appointments.
  - Connect design with instruction: When a course is designed and taught by different people, the instructor can provide useful feedback to the designer. Campus should consider developing a simple mechanism for accomplishing that.

### ***3.2.3 Gather input from instructors themselves***

In many modern teaching evaluation systems, self-evaluation plays an important role. Self reflections on teaching provide important context for assessment data from students and peers. A record of real-time self-evaluation can provide evidence of the evolution of an instructor's performance. Furthermore, the process of regularly reflecting on one's own teaching promotes a culture of continued improvement in instruction.

Current practice at UIUC:

- Annual reviews: When information on teaching is collected, often the instructor is just asked to list any teaching awards, students advised, instructional units taught, ICES scores, and so forth. There is no consistency across units in what information is collected or how it is used.
- Third-year reviews: Many units collect the full candidate-authored components of the promotion and tenure dossier, but practice is inconsistent.
- Promotion and tenure dossiers include:
  - A candidate essay on “statement of teaching.” This is long, but has no prescribed structure.
  - A self-reported list of activities related to teaching that the candidate did other than classroom teaching. There is a category for “Other Contributions to Instructional Programs” but little guidance for what might go there, and little opportunity for the candidate to reflect on the impact and quality of those activities.

Suggested changes and new practices for self-evaluation practice:

(See Appendices 9, 10, and 11 for draft instruments for self-evaluation)

- Annotated annual inventory of teaching contributions:
  - For annual reviews, each instructor will complete a rubric for an annotated inventory of other contributions to instruction that year (see Appendix 8).
  - The inventory will be simple so as not to be burdensome like a traditional teaching dossier, but have space for comments required to explain the scope, quality, and impact of each of the items in the list.
  - The inventory will have an extensive list of categories of activities that can be included, such as: advising, mentoring, course and curriculum development, committee service and leadership, peer review, engagement in instructional improvement on campus, advising student clubs, and so forth.
- Course self reflections:
  - At the end of every course, each instructor will complete an automatic short online self-reflection on their own teaching.
  - The first part of the self-reflection is a short rubric; see Appendix 10. This formative assessment is only for the instructor’s personal records and use, and not part of their formal record.
  - The second part of the self-reflection is a formal set of short open ended questions to be answered after completing the rubric; see Appendix 11. This formal part of the self-reflection goes into the instructor’s record of teaching assessment along with the student ratings and peer evaluations (if any) for that course. The self-reflection

provides space for the instructor to describe their goals and accomplishments. It also makes space for the instructor to describe challenges they faced (which could include bias and harassment from students) and how they worked to address them.

- Self reflections will be prompted for all instructors, not just Assistant and Associate faculty.
- Annual reviews can draw upon the formal course self reflections and annual inventory on teaching contributions for that year.
- Promotion and tenure decisions and award applications:
  - The candidate must complete a guided, structured essay based on the self-reflection questions that captures the entire arc of their course-based teaching during the relevant period of time. What is their teaching philosophy – what are their goals in course design and delivery, how do they choose to accomplish those goals, and why? In what areas have they been consistently excellent? In which areas have they improved? What were their challenges, improvements, and accomplishments? What are their goals for teaching in the future? This will replace the previous “teaching statement.”
  - The candidate uses the annual inventories to complete a cumulative inventory of contributions to teaching with a final one paragraph summary about the holistic body of their total contributions to teaching.

#### Notes on implementation:

- Immediate critical needs
  - Refine personal and formal instruments for self reflection on a single course.
  - Refine instrument for annotated annual inventory of contributions to teaching.
- Next steps:
  - Develop instrument for self reflection on courses over several years for promotion, tenure, and award decisions.
  - Develop instrument for annotated inventory of contributions to teaching over the course of several years.

### 3.3 Holistic evaluation based on multiple inputs

- Single-year holistic evaluation
  - Official evaluation for actions such as annual reviews and merit raises draw upon a holistic assessment of teaching contributions for that year.

- When evaluative inputs for a year have been received, the unit EO or their designee gathers together the inputs to complete a Single-Year Holistic Evaluation of Instruction (see Appendix 12).
  - This rubric solicits broad evaluations (Area for Improvement, Acceptable, or Area of Strength) for the four elements of excellence in course-based teaching and for the overall inventory of teaching contributions.
  - Instructor course self-evaluations should be used to provide context for negative responses in student ratings or peer reviews.
  - Comments should provide reasons for findings and suggestions for improvement, and are especially important where input responses are negative or are mixed.
  - An instructor need not have excellence in all sub-parts of an element for that to be an area of strength.
  - Instructors need not exceed “acceptable” in the elements of teaching excellence in order to earn tenure (for tenure-track faculty) or contract renewal (for other instructors).
- Multi-year holistic evaluation:
    - A multi-year holistic evaluation will be carried out for the relevant period of time to inform actions such as promotion and/or tenure decisions and teaching awards.
    - The multi-year evaluation will draw upon all the single year evaluations for the relevant time period.
    - Individual instruments may be used to inform the comment sections.
    - The evaluation will include a multi-year multi-criteria visualization of the summary ratings from the single year evaluations.
    - Qualitative comments will discuss the arc of the instructor’s accomplishments. For example: Have some of the criteria improved over time? What challenges have they faced? What are areas of sustained and/or current excellence? What are areas of sustained and/or current potential for improvement? Is the instructor meeting expectations for their position and rank?

Notes on implementation:

- Immediate critical needs:
  - Refine Single-Year Holistic Evaluation of Instruction: The taskforce will work with others to refine this instrument.
- Next steps:
  - Develop Multi-Year Holistic Evaluation of Instruction: This will include a visualization of annual holistic ratings over the time period.

## 4. Recommendations for functions that draw upon teaching evaluation

UIUC's system of teaching evaluation will be structured to support actions taken at multiple levels of the university, as follows.

### 4.1 Instructor: own efforts to improve teaching, gain advancement

- Instructors should take advantage of opportunities to add questions to student surveys to gain more specific feedback on their teaching.
- Instructors should promptly review feedback from students and peers and use that feedback to improve future course design and instruction.
- Instructors should take advantage of the brief structured self reflections after courses to consider what they have learned from the semester and what changes they want to make in the future.
- Instructors can use the annual inventory of contributions to teaching to build a record over time of their work in this area. That record can support promotion cases, merit raises, and teaching awards.

### 4.2 Administrators

- Troubleshooting: EOs should promptly review the results of student ratings and peer reviews to identify signs of significant problems with a course or its instruction.
  - Note that a problem such as disrespectful behavior may be present even if only flagged by a small number of students in a class.
  - The EO should determine whether a problem really exists that needs to be corrected and the complaints were not made capriciously or in bad faith, and take appropriate action if a problem is found.
- Annual personnel functions:
  - These functions include annual reviews, determinations of regular merit raises, decisions about contract renewals for specialized faculty
  - EOs or their designated agents (the Director of Undergraduate Studies, Associate Head, Executive Committee, etc.) should construct a Single-year Holistic Evaluation of Teaching according to Section 3.3.
  - EOs should use that holistic evaluation in conjunction with the year's individual self reflections, inventory of contributions, peer reviews, and student surveys to inform these functions.
- Multi-year personnel functions:

- These decisions, including promotion and tenure processes and selection of teaching awards, are made on the basis of multi-year arc assessments of teaching activity.
- EOs or their designated agents (the Director of Undergraduate Studies, Associate Head, Executive Committee, etc.) should construct a Multi-year Holistic Assessment of Teaching according to Section 3.3.
- Administrators and decision making committees should draw upon that multi-year holistic assessment in conjunction with: the longitudinal student survey report, the longitudinal peer review report, and the instructor’s multi-year self-assessment of course teaching and cumulative self-reported inventory of contributions to inform these functions.

## 5. Implementation process

One does not change a system like this overnight. Implementation of a new teaching evaluation system requires extensive consultation, piloting, and feedback to be successful. Below is an outline of the stages this Task Force recommends for implementation.

- Consultation and final report: Large scale change in a process as fundamental to the university as teaching evaluation must be consultative. In addition to consultation listed in Table 1, the Chair of the Task Force will solicit feedback from the Council of Deans and the Faculty Senate and give an open webinar on the report via Zoom for anyone to watch. That webinar will be recorded for asynchronous viewing, and there will be a survey accompanying the webinar in which any person on campus can submit comments. The Task Force will make revisions and issue a final report.
- New Campus Guidance: A version of these recommendations will be codified as a “Guide to Evaluation of Contributions to Teaching at the University of Illinois at Urbana-Champaign” and maintained on a public website. This guide can be referenced by other policy documents on campus, and details of the guide can be updated as the system is implemented, evaluated, and improved over time.
- New Task Force: A new Task Force will be created to work with Provost Fellow Ando on implementation. It will include some members of the original Task Force, but include others whose expertise is particularly relevant to the implementation stage of this effort.
- Instrument refinement including focus groups: This system for teaching evaluation has many instruments for gathering input and producing reports. The new Task Force will work with others on campus to refine the draft instruments and write those that have not been drafted. This process will include focus group exercises with members of the campus

community (students, peer reviewers, instructors, EOs) who would actually be using these instruments in the evaluation system.

- Develop administration infrastructure: This system will work best if all elements are implemented in an efficient online system that conducts student surveys, prompts self reflections and annual self-evaluation.
- Pilot implementation: Units from several different colleges around campus will be selected for pilot implementation of the new system. Experience with all elements of the system will be evaluated so the instruments and systems can be improved for campus-wide implementation. At the same time, all departments will be encouraged to do focused tailoring of the instruments to meet their needs.
- Complete implementation and ongoing evaluation: The entire system will be implemented campus wide. However, the Task Force recommends that data be collected on system performance, and that the system be evaluated for possible revision within three to five years of full implementation. Implementation should be planned with provisions for evaluating several important features of success, such as reduced bias, increased compliance, and instructor improvement. Campus should continue for several years to collect a sample of traditional ICES scores in parallel with revised student feedback surveys to facilitate “translation” of results from one system into the other.

## 6. Topics for future work

This task force focused on recommending changes that can be implemented in the immediate future. There are important issues where best practices are still to be developed. The task force recommends that UIUC consider working on these elements of teaching evaluation for possible future incorporation into policy and procedure.

- Research to learn more about best practices in reducing bias in student rating surveys, peer evaluations, and even self-evaluations. This research could begin during the pilot studies for the student surveys and peer review systems.
- Develop practices to evaluate the quality of mentoring and advising of graduate students and undergraduate researchers.
- Developing scalable practices to quantify student accomplishment of learning objectives.

## Appendix 1: Charge Letter from Provost Cangelaris

November 20, 2019

Dear Colleagues,

Well-designed teaching evaluation systems help instructors to be more effective. Teaching evaluations also play important roles in decisions regarding promotion and tenure, specialized faculty contracts, and award selection. Thus, it is vital that evaluation systems produce valid and unbiased measures of performance. Modern best practices in teaching evaluation draw on evidence from sources such as student feedback, self-evaluation, and peer assessment to show dimensions of accomplishment and highlight areas for improvement. The task force will develop a teaching evaluation system for UIUC that is constructive, unbiased, administratively efficient, and mindful of diverse teaching needs and styles across campus.

I ask that you develop a system to evaluate teaching effectiveness. In addition, please provide recommendations on how Communication #9 on Promotion & Tenure and Communication #26 on promotion of specialized faculty should be modified to incorporate the new approach to evaluating teaching effectiveness. If appropriate, the Provost would welcome recommendations to incorporate the proposed new approach into the Annual Review process (Communication #21) and the Third Year Review Process (Communication #13).

During your work, I ask that you consult widely with the appropriate stakeholders across campus, including tenured faculty, pre-tenure faculty, campus leaders, members of the Faculty Senate, and the Center for Innovation in Teaching and Learning (CITL). Also, please engage with other campus committees who have an interest in teaching evaluations, particularly the DRIVE committee; the Teaching Advancement Board; and the Provost's Task Force on the Revision of Promotion and Tenure Guidelines.

Professor Amy Ando (ACES), in her role as Provost Fellow, will lead this task force. I anticipate that the committee will meet throughout the 2019-2020 academic year and, possibly, into the 2020-21 academic year.

I look forward to your recommendations, and I thank you for your willingness to participate in this important effort.

Sincerely,

Andreas C. Cangelaris

Vice Chancellor for Academic Affairs and Provost

M. E. Van Valkenburg Professor of Electrical and Computer Engineering

## Appendix 2: Annotated Bibliographies and Publications

Holman, Mirya, Ellen Key, and Rebecca Kreitzer. 2019. "Evidence of Bias in Standard Evaluations of Teaching." <http://www.rebeccakreitzer.com/bias/>.

**Annotated running bibliography with link to document of summaries of articles.**

Savonick, Danica and Cathy N. Davidson. 2018. "Gender Bias in Academe: An Annotated Bibliography of Important Recent Studies." HASTAC (Humanities, Arts, Science, and Technology Alliance and Collaboratory) Blog.

<https://www.hastac.org/blogs/superadmin/2015/01/26/gender-bias-academe-annotated-bibliography-important-recent-studies> .

**Annotated running bibliography of research related to possible discrimination in academia, including in teaching evaluations. Also has a link to a public document with more papers listed and described:**

<https://docs.google.com/document/d/1QRcQU4RSizlu-HxDY2uZxYp4EmYslmvm9BMtcd-RUis/edit>

\*\*\*

American Sociological Association (ASA). 2019. *Statement on Student Evaluations of Teaching*. [https://www.asanet.org/sites/default/files/asa\\_statement\\_on\\_student\\_evaluations\\_of\\_teaching\\_sept52019.pdf](https://www.asanet.org/sites/default/files/asa_statement_on_student_evaluations_of_teaching_sept52019.pdf).

Anderson, Kristin J., and Gabriel Smith. 2005. "Students' Preconceptions of Professors: Benefits and Barriers According to Ethnicity and Gender." *Hispanic Journal of Behavioral Sciences* 27(2): 184-201.

Basow, Susan A., and Julie L. Martin. 2012. "Bias in Student Evaluations." Pp. 40-49 in *Effective Evaluation of Teaching: A Guide for Faculty and Administrators*, edited by Mary E. Kite. Washington, DC: Society for the Teaching of Psychology. Retrieved from <http://teachpsych.org/ebooks/evals2012/index.php>.

Bavishi, Anish, Juan M. Madera, and Michelle R. Hebl. 2010. "The Effect of Professor Ethnicity and Gender on Student Evaluations: Judged Before Met." *Journal of Diversity in Higher Education* 3(4): 245-256.

Boring, Anne, Kellie Ottoboni, and Philip B. Stark. 2016. "Student Evaluations of Teaching (Mostly) Do Not Measure Teaching Effectiveness." *ScienceOpen Research* (DOI: 10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1).

Boring, Anne. 2017. "Gender Biases in Student Evaluations of Teaching." *Journal of Public Economics* 145: 27-41. <http://dx.doi.org/10.1016/j.jpubeco.2016.11.006>

- Boysen, Guy A. 2015. "Significant Interpretation of Small Mean Differences in Student Evaluations of Teaching Despite Explicit Warning to Avoid Overinterpretation." *Scholarship of Teaching and Learning in Psychology* 1(2):150-162.
- Flaherty, Colleen. 2015. "Flawed Evaluations." *Inside Higher Ed*.  
<https://www.insidehighered.com/news/2015/06/10/aaup-committee-survey-data-raise-questions-effectiveness-student-teaching#.XephseboY5k.link>.
- Flaherty, Colleen. 2018. "Arbitrating the use of student evaluations of teaching." *Inside Higher Ed*. <https://www.insidehighered.com/quicktakes/2018/08/31/arbitrating-use-student-evaluations-teaching>
- Goodman, James, Robert Anson, and Marcia Belcheir. 2015. "The Effect of Incentives and Other Instructor-Driven Strategies to Increase Online Student Evaluation Response Rates." *Assessment and Evaluation in Higher Education* 40(7): 958-970.  
<https://doi.org/10.1080/02602938.2014.960364>
- MacNell, Lillian, Adam Driscoll, and Andrea N. Hunt. 2015. "What's in a Name: Exposing Gender Bias in Student Ratings of Teaching." *Innovative Higher Education* 40(4):291-303.
- Mitchell, Kristina M. W., and Jonathan Martin. 2018. "Gender Bias in Student Evaluations." *PS: Political Science and Politics* 51(3):648-652.
- Peterson, David A. M., Lori A. Biederman, David Andersen, Tessa M. Ditonto, and Kevin Roe. 2019. "Mitigating Gender Bias in Student Evaluations of Teaching." *PLoS ONE* 14(5): e0216241. <https://doi.org/10.1371/journal.pone.0216241>
- Reid, Landon D. 2010. "The Role of Perceived Race and Gender in the Evaluation of College Teaching on RateMyProfessors.com." *Journal of Diversity in Higher Education* 3(3):137-152.
- Smith, Bettye P., and Billy Hawkins. 2011. "Examining Student Evaluations of Black College Faculty: Does Race Matter?" *The Journal of Negro Education* 80(2):149-162.
- Spooren, Pieter, Bert Brockx, and Dimitri Mortelmans. 2013. "On the Validity of Student Evaluation of Teaching: The State of the Art." *Review of Educational Research* 83(4):598-642.
- Spooren, Pieter, and Francis Van Loon. 2012. "Who Participates (Not)? A Non-Response Analysis on Students' Evaluations of Teaching." *Procedia-Social and Behavioral Sciences* 69: 990-996. <https://doi.org/10.1016/j.sbspro.2012.12.025>.
- Stark, Philip B., and Richard Freishtat. 2014. "An Evaluation of Course Evaluations." *ScienceOpen Research* (DOI: 10.14293/S2199-1006.1.SOR-EDU.AOFRQA.v1).

Uttl, 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(1):22-42.

Uttl, Bob, Kelsey Cnudde, and Carmela A. White. 2019. "Conflict of Interest Explains the Size of Student Evaluation of Teaching and Learning Correlations in Multisection Studies: a Meta-Analysis." *PeerJ* 7: e7225. <https://doi.org/10.7717/peerj.7225>

Weaver, Gabriela C., Ann E. Austin, Andrea Follmer Greenhoot, and Noah D. Finkelstein. 2020. "Establishing a Better Approach for Evaluating Teaching: The TEval Project." *Change: The Magazine of Higher Learning*, 52:3, 25-31. <https://doi.org/10.1080/00091383.2020.1745575>

Wieman, Carl. 2015. "A Better Way to Evaluate Undergraduate Teaching." *Change: The Magazine of Higher Learning* 47(1): 6-15. <https://doi.org/10.1080/00091383.2015.996077>

### Appendix 3: Sample Innovations at Other Institutions

Institution	Innovations; Website where available
University of Southern California	Complete new system: definition of teaching quality, peer review rubrics, templates for annual instructor self-evaluations, use of student surveys minimized <a href="http://cet.usc.edu/resources/instructor-course-evaluation/">http://cet.usc.edu/resources/instructor-course-evaluation/</a>
University of Oregon	Complete new system: definition of teaching quality, automatic self-evaluation after class; student ratings changed to align with definition. <a href="https://provost.uoregon.edu/revising-uos-teaching-evaluations">https://provost.uoregon.edu/revising-uos-teaching-evaluations</a>
University of Kansas	NSF funded TEval participant. Excellent holistic rubric of accomplishment. <a href="https://cte.ku.edu/evaluating-teaching-ku">https://cte.ku.edu/evaluating-teaching-ku</a>
University of Colorado – Boulder	NSF funded TEval participant. Implementing three-voices holistic teaching evaluation. <a href="https://www.colorado.edu/teaching-quality-framework/">https://www.colorado.edu/teaching-quality-framework/</a>
University of Massachusetts – Amherst	NSF funded TEval participant. Changing student surveys. <a href="https://www.umass.edu/oapa/program-assessment/instructional-innovation-assessment/evaluation-teaching-new-approach">https://www.umass.edu/oapa/program-assessment/instructional-innovation-assessment/evaluation-teaching-new-approach</a>
University of Minnesota	Expanding peer review. <a href="https://faculty.umn.edu/faculty-support-and-resources/peer-review-teaching">https://faculty.umn.edu/faculty-support-and-resources/peer-review-teaching</a>
Colorado State University	New student survey. <a href="https://coursurvey.colostate.edu/instructions">https://coursurvey.colostate.edu/instructions</a>
University of Texas – Austin	Expanded peer observation. <a href="https://facultyinnovate.utexas.edu/peer-observation">https://facultyinnovate.utexas.edu/peer-observation</a>
Pennsylvania State University	Definition of teaching quality. <a href="http://www.schreyerstitute.psu.edu/Definition/">http://www.schreyerstitute.psu.edu/Definition/</a>
University of California – Irvine; University of Michigan	Resolution that more than student evaluations should be used in teaching evaluation
University of Nebraska – Lincoln	New student experience survey awaiting approval
Yale, Amherst, Stanford	Including Student Assessment of Learning Gains <a href="http://www.salgsite.org/">http://www.salgsite.org/</a>

## **Appendix 4: Draft Campus-Wide Definition of Teaching Excellence at the University of Illinois Urbana-Champaign**

Excellent course teaching at the University of Illinois Urbana-Champaign (UIUC) is:

- Well-designed
- Well-delivered
- Inclusive and ethical
- Reflective and evolving

These four elements of teaching excellence are further defined in the context of course-based teaching on the next page. Detailed instruments for evaluation by units will refine and tailor the criteria even further, adding details that show how the general principle is manifest in that discipline.

An excellent teacher might make a range of important contributions<sup>2</sup> to teaching at UIUC including (but not limited to):

- contributing to student learning outside course environments (e.g. clubs, study abroad)
- developing new curricula
- engaging in pedagogical research and innovation
- mentoring or supervising undergraduate projects
- peer review of other instructors
- serving as a formal or informal mentor to students
- serving on and contributing to graduate student thesis/dissertation committees
- serving on and/or providing leadership for administrative committees that shape and oversee teaching and program coordination
- teaching to audiences outside the traditional Illinois student community
- training and supervising students in research, teaching, and/or internships

Holistic review of teaching excellence should document and evaluate these contributions.

---

<sup>2</sup> Items are listed alphabetically.

Excellent course-based teaching at the University of Illinois is:<sup>3</sup>

DES

• **Well-designed**

- Creates clear learning objectives<sup>4</sup> that are both challenging and attainable, and intersect with the overall course goals.
- Designs student-centered learning and assessment experiences that support the learning objectives.
- Provides readily available, coherently organized, and high quality course materials (including syllabi) that establish student workload, learning objectives, performance criteria, grading procedures, and class policy expectations.

DEL

• **Well-delivered**

- Organizes and manages class time and students' activities in and out of class to yield high levels of student learning.
- Models and fosters critical, analytical, and creative thinking.
- Communicates content clearly with students
- Has constructive and timely written and oral communication with students.
- Gives timely, useful, and fair feedback on activities and assignments tied to performance criteria and learning objectives.
- Fosters the mindset that growth is possible and ability is not fixed.
- Engages, challenges, and supports students cognitively, emotionally, and behaviorally.

IE

• **Inclusive and ethical**

- Designs instruction to be inclusive and equitable (including practices of Universal Design for Learning<sup>5</sup>), ensuring that every student can participate fully and their presence and participation is valued.
- Represents the diversity of the field's practitioners, content, and applications.
- Acts with integrity and respect toward students.
- Implements course policies that are equitable and applied fairly.

RE

• **Reflective and evolving**

- Implements regularly-revised content and pedagogical techniques that are current, research-informed, rigorous, and relevant.
- Uses teaching development activities, evaluations, and other indicators to inform teaching practice and student learning experiences.
- Creates an environment that invites constructive student feedback to the instructor.

---

<sup>4</sup> Note that “learning objectives” should be interpreted very broadly, and not just as a list of factual content to be conveyed.

<sup>5</sup> Information about UDL practices is at <https://www.disability.illinois.edu/academic-support/accessible-it-group/udl/teach-learn>.

## Appendix 5: Draft Student Feedback on Courses and Instruction

### Why provide ratings?

- Universities evaluate teaching for a number of reasons, such as to:
  - help instructors improve;
  - evaluate faculty and other instructors for promotion and merit;
  - identify problems with courses and instruction.
- Your feedback is part of a comprehensive system at the University of Illinois to help ensure that you and future Illinois students receive the best instruction possible.
- The faculty and administration take your feedback seriously, and so should you.
- On average, providing this feedback takes less than ten minutes, so you can do a lot of good with a very small investment of time.
- At no point will your instructor see the name of a student who provided feedback, and your instructor will not see the results until the final grades have been submitted.

### Important reminder<sup>6</sup>

- As we learn more about the errors and biases we all bring to our assessments, research shows that even the best intentions to provide accurate evaluations can be affected by factors quite independent of actual teaching standards and teaching ability. For example:
  - We might rate easier courses as better than harder ones now, even though we may think differently a few years later.
  - We can be influenced by irrelevant or superficial features of the instructors, like [appearance] [gender or race]<sup>7</sup>, when what matters is how much we learned.
- At Illinois, we want our evaluations to be fair and unbiased so we can reward and retain the best teachers. Unbiased ratings will assist the university and future generations of students.
- Please make sure that, as you rate your instructor and the class, you focus on the quality of instruction and of the class.
- **Please note that you are evaluating the instruction of <<instructor name>> in <course>>**<sup>8</sup>
- **You will be answering questions about:**
  - Your level of engagement with the course (2 multiple choice or scale questions)
  - Your feedback on specific features of the course and instruction (12 multiple choice questions).
  - Your open-ended opinions about the course and instruction (4 short open ended questions).

\*\*\*\*\*

---

<sup>6</sup> The Task Force recommends that research be conducted to identify the impact that this kind of anti-bias statement has on results, and to inform the best way to construct such a statement if it proves useful.

<sup>7</sup> Research could experimentally vary which phrase is used here.

<sup>8</sup> Actual instructor name and course name will go here.

### Part 1: Your engagement with the course

- 1) Think about the class sessions (if in person or synchronous online) or recordings (if asynchronous). About what percent of those classes did you attend or view? Move the orange slider to indicate:



- 2) Think about the time you spent on this course outside of face-to-face class time or watching the equivalent class recordings.

About how many hours did you spend each week on that extra work and activity?

- 0-1 hours each week
- 2-4 hours each week
- 5-7 hours each week
- 8-10 hours each week
- More than 10 hours each week

### Part 2: Ratings of specific features of the course and instruction<sup>9</sup>

Please read each statement below and indicate how true you think that statement was for this course and instruction.

DES

- 1) The course materials and instructor communicated the course objectives and policies very clearly.

- Not at all true
- Slightly true
- Moderately true
- Very true
- Completely true

DES

- 2) The course activities, including assignments and assessments, were aligned well with the course focus and objectives.

DES

- 3) This course was appropriately challenging. It took a reasonable amount of effort to master the content and keep up with the workload.

DEL

- 4) The course was well run.

[For example, the instructor was timely, it was usually easy to access the materials, regular course activities ran smoothly, course communication was frequent and clear.]

<sup>9</sup> We show the proposed rating scale for question (1). The same scale is proposed for questions (2) through (12) and is not repeated here. This draft shows the element of the Definition of Teaching Excellence each question aligns to; those indicators will not be on the actual student surveys.

- DEL** 5) **The instructor’s feedback to me was valuable.** For example, comments on my work were helpful and encouraging, feedback came quickly after tests and assignments, feedback helped me learn.
- DEL** 6) **The instructor provided support and encouragement, and was so engaging that I really wanted to learn.**
- DEL** 7) **I think I learned a great deal in this course.**
- DEL** 8) **I had many opportunities in this course to be creative or practice critical/analytical thinking.**
- IE** 9) **In my experience, the instructor was always respectful.**
- [NOTE: This means the instructor did not use disrespectful language including inappropriate comments regarding individual characteristics such as intellectual ability, race or ethnicity, religion, gender, political affiliation, sexual orientation, or place of origin. The instructor did not engage in disrespectful actions like mocking facial expressions and gestures, sexual harassment, or asking students to do things that are humiliating or denigrating.]
- IE** 10) **The instructor and course design made it easy for me to participate in all elements of the class, and I felt that my presence was welcome.**
- IE** 11) **Course policies, including grading procedures, were applied fairly in the class.**
- RE** 12) **The instructor encouraged students to provide feedback on the course and/or instruction.**

### Part 3: Your open-ended opinions about the course and instruction

- 1) Do you think your instructor’s teaching was excellent? (YES/NO)  
- **(If yes, comment box opens)** Please say a few words about why.
- 2) Did you experience problems with this class or your instructor you want the department to know about? (YES/NO)  
- **(If yes, comment box opens)** Please say a few words about your experience.
- 3) What are the major strengths of the instructor and/or the course design?
- 4) What do you suggest to improve instruction and/or the course design?

## Appendix 6: Draft Peer Review of Instruction<sup>10</sup>

(Use with guidance sheet with details to evaluate criteria for this type of class in this unit)

Faculty member reviewed: \_\_\_\_\_  
Reviewer: \_\_\_\_\_  
Course Name and Number: \_\_\_\_\_  
Date: \_\_\_\_\_

### Before observation:

- The instructor will comment on personal goals for instruction before the observation.
- The instructor and reviewer will agree on any elements that are N/A for the course.

### During observation:

- The reviewer will indicate “Area of strength,” “Acceptable,” “Area for improvement” for each of the specific elements
- The reviewer will provide some comments for each category of criteria. Comments must be provided for any criterion indicated as an “Area for improvement”.

### After observation:

- The reviewer will comment on the instructor’s achievement of their goals

**Note for reviewer:** Research shows that even the best intentions to provide accurate evaluations can be affected by irrelevant or superficial features of an instructor, when what matters is the quality of their teaching. At Illinois, we want our evaluations to be fair and unbiased so we can reward and retain the best teachers. Please make sure that, as you review the instructor, you focus on quality of instruction.

**Pre-observation (instructor):** Personal goals for instruction in this course:

**Post-observation (reviewer):** Comments on goal achievement:

<sup>10</sup> Colored markers indicate the elements that parts of the rubric reflect upon. They will not be included in the final instrument.

	N/A	Area for improvement	Acceptable	Area of strength	<b>Instruction is well-delivered</b>
<b>DEL</b>					Organizes and manages class time and students' activities in class to yield high levels of student learning.
<b>DEL</b>					Models and fosters critical, analytical, and creative thinking.
<b>DEL</b>					Communicates clearly and constructively with students.
<b>DEL</b>					Fosters a mindset that growth is possible.
<b>DEL</b>					Engages, challenges, and supports students cognitively, emotionally, and behaviorally.
<b>DEL</b>					Uses pedagogical techniques that are current and appropriate for the content and learning objectives.
Comments:					

	N/A	Area for improvement	Acceptable	Area of strength	<b>Instruction is inclusive and ethical</b>
<b>IE</b>					Conducts instruction such that every student can participate fully.
<b>IE</b>					Conducts instruction such that every student's presence and participation is facilitated and valued.
<b>IE</b>					Recognizes the power differential between instructor and student and acts with integrity and respect toward students.
Comments:					

Any additional formal comments:

Confidential comments just for the instructor:

## Appendix 7: Draft Peer Review of Course Materials and Design<sup>11</sup>

(Use with guidance sheet with details to evaluate criteria for this type of class in this unit)

**Definition:** For purposes of this review, the “**course designer**” is the person who was in charge of final decisions regarding the course design (e.g. syllabus, reading list, assignments, assessments) for the semester in which the course is being evaluated. The course designer may also be the instructor who delivers the course, but that is not always the case.

Course designer: \_\_\_\_\_ Reviewer: \_\_\_\_\_  
Course Name and Number: \_\_\_\_\_ Date: \_\_\_\_\_

Before review:

- The course designer will comment on their goals for the course design.
- The course designer will explain to the reviewer what changes they have made to the course in this cycle, and how they used evaluations and other indicators to inform those changes.
- The course designer and reviewer will agree on any review elements that are N/A for the class.

During review:

- The reviewer will indicate “Area of strength,” “Acceptable,” “Area for improvement” for the remaining specific elements.
- The reviewer will provide some general comments for each category of criteria.

After review:

- The reviewer will comment on the course designer’s achievement of the goals for the course.

**Note for reviewer:** Research shows that even the best intentions to provide accurate evaluations can be affected by irrelevant or superficial features of an employee, when what matters is the quality of their work. At Illinois, we want our evaluations to be fair and unbiased so we can reward and retain the best teachers. Please make sure that, as you review, you focus on the quality of the design of this course.

<b>Pre-review (course designer):</b>
What changes have been made (if any) since the previous offering, and how did feedback influence those choices?
Goals for course design:
<b>Post-review (reviewer)</b> Comments on goal achievement:

<sup>11</sup> Colored markers indicate the elements that parts of the rubric reflect upon. They will not be included in the final instrument.

		N/A	Area for improvement	Acceptable	Area of strength	<b>Learning objectives and course materials are well-designed</b>
DES						Course materials establish clear learning objectives that are both challenging and attainable and intersect with overall course goals.
DES						Course materials establish student-centered learning and assessment experiences that support the learning objectives.
DES						Course materials, including the syllabus, are readily available and coherently organized.
DES						Course materials establish student workload, performance criteria, grading procedures, and class policy expectations.
DES	RE					Course content is current, rigorous, and relevant.
Comments:						

		N/A	Area for improvement	Acceptable	Area of strength	<b>Course design is inclusive and ethical</b>
IE						Course design ensures every student can participate fully.
IE						Course design ensures every student’s presence and participation is valued.
IE						Course content represents the diversity of the field's practitioners, content, and applications.
IE						Course policies are equitable.
Comments:						

Any additional formal comments:
Confidential comments just for the course designer:

## Appendix 8: Sample Guidance for Peer Evaluation of Teaching<sup>12</sup>

*This sheet gives examples of concrete elements for peer reviewers to identify in evaluating <in person> course design and instruction. Elements that need tailoring for the department are marked with a \**

### Instruction:

#### **Instruction is well-delivered and engaging:**

- \*Organizes and manages class time and students' activities in class to yield high levels of student learning.
  - DOES: Use active learning, class discussion, or other course activities effectively; encourage student interaction and broad participation; make good use of class time.
  - DOES NOT: Arrive late, use poorly planned course activities.
- Models and fosters critical, analytical, and creative thinking.
  - DOES: Provide students with opportunities to think creatively and solve problems themselves; model analytical thinking themselves when asked questions.
  - DOES NOT: Discourage divergent thinking; encourage students only to memorize.
- Communicates clearly and constructively with students.
  - DOES: Speak such that students can hear and/or understand; reinforce verbal communication with materials such as prepared notes, slides, audio-visual artifacts, or embodied demonstrations; use verbal statements and other materials that are clear and encourage creative thought.
  - DOES NOT: Speak indistinctly; fail to re-explain communication that students did not understand; criticize without guidance for improvement.
- Fosters a mindset that growth is possible.
  - DOES: Celebrate student improvement; turn student contributions and even mistakes into positive learning opportunities.
  - DOES NOT: Violate FERPA by publicly sharing grades; rank students by quality of performance; compare student work to an unrealistic or ambiguous standard.
- Engages, challenges, and supports students cognitively, emotionally, and behaviorally.
  - DOES: Ask students to make appropriately non-trivial contributions to the course such as solving original problems, designing a project, closely observing a demonstration; provide support for all students to be able to contribute; give constructive feedback.
  - DOES NOT: Instruct at a level that is much too easy or too hard; provide non-constructive and/or discouraging feedback.
- \*Uses pedagogical techniques that are current and appropriate for the content and learning objectives.
  - DOES: ---
  - DOES NOT: ---

---

<sup>12</sup> Different versions of guidance sheets can be developed for different types of classes (online/in-person; lecture/discussion; etc.) and then tailored by units.

**Instruction is inclusive and ethical:**

- Conducts instruction such that every student can participate fully
  - DOES: Use accessibility best practices during instruction such as describing images used in presentation and using closed captions for videos.
  - DOES NOT: Use pedagogical techniques that exclude students with disabilities.
- Conducts instruction such that every student’s presence and participation is facilitated and valued.
  - DOES: Implement opportunities for students of varied abilities and opinions to take part in class with their whole selves; treat all students equitably.
  - DOES NOT: Publicly reveal students’ disabilities; embarrass or denigrate students in the classroom; allow students to be excluded from participation due to reticence or ideological isolation.
- Recognizes the power differential between instructor and student, and acts with integrity and respect toward students.
  - DOES: Provide opportunities for students to engage in constructive dialogue; address students with respect; encourage students to treat each other with respect; make no assumptions about students.
  - DOES NOT: Make suggestive comments to students; use inappropriate or offensive gestures and/or speech; use discriminatory, dismissive, or other abusive language; ignore offensive behavior on the part of students; encourage students to work personally for the instructor.

**Course materials and design:**

**Learning objectives and course materials are well-designed:**

- Course materials establish clear learning objectives that are both challenging and attainable and intersect with overall course goals
  - DOES: Clearly state learning objectives in syllabus that are neither trivial nor unreasonably ambitious; list aspirational goals separately with context in the syllabus
  - DOES NOT: Fail to state learning objectives in course materials; have aspirational goals listed as core learning objectives; choose objectives that are trivial for the students in the course; emphasize learning objectives that are tangential to the mission of the course.
- Course materials establish student-centered learning and assessment experiences that support the learning objectives.
  - DOES: Include an articulated logical sequence for learning the course content; include opportunities to develop skills relevant to the field through course activities; use pedagogical approaches to engaged learning that are current state-of-the art in the field; have assessments that further student mastery of course content
  - DOES NOT: Award course credit solely for attendance; use dated pedagogical approaches that do not afford opportunities for student engagement; design assessments that are not closely related to student learning activities

- Course materials, including the syllabus, are readily available and coherently organized
  - DOES: Make all materials easily available in a timely way for students; make clear which materials should be accessed at different points in time; avoid technical problems with materials.
  - DOES NOT: Have materials that are missing or unclear; suffer from frequent technical problems in accessing and using course materials.
- Course materials clearly establish student workload, performance criteria, grading procedures, and class policy expectations.
  - DOES: Describe all assignments that will be expected throughout the term; include clear descriptions and grading rubrics for assessments and final grades; state expectations for class participation and behavior; establish consequences for things like tardiness and academic integrity violations.
  - DOES NOT: Fail to describe important elements of the course.
- Course content is current, rigorous, and relevant.
  - DOES: Use learning materials (readings, exercises, etc.) that are up to date with advances in the field, reasonably challenging, and closely aligned with learning objectives.
  - DOES NOT: Use learning materials that are outdated; include irrelevant material that personally benefits the instructor

**Course design is inclusive and ethical:**

- Course design ensures every student can participate fully
  - DOES: Use content that can be accessed by all students in the class; states clearly what may be technology requirements with instructions; ensure all required resources are equitably available to all students
  - DOES NOT: Have assignments that unduly burden some students due to issues of access or accommodation.
- Course design ensures every student's presence and participation is valued
  - DOES: Establish expectations in course materials for civility and inclusion in the classroom; constructs safe opportunities for all students to add their own perspectives to class activities.
  - DOES NOT: Structure activities such that they are likely to make particular students uncomfortable or isolated.
- Course content represents the diversity of the field's practitioners, content, and applications
  - DOES: Include diverse perspectives, authors, or applications in the field.
  - DOES NOT: Include materials that unreasonably reflect an instructor's personal bias
- Course policies are equitable
  - DOES: Establish fair grading procedures; state clear decision trees for student requests regarding problems such as missed work.
  - DOES NOT: Establish inappropriate rewards or consequences for course participation; have unclear, inappropriate, or unfair grading procedures or rubrics for assessments; expect work outside of class that would disproportionately burden some students; refuse to accommodate religious holidays.

## Appendix 9: Draft Self-Reported Annual Inventory of Teaching Contributions

Name: \_\_\_\_\_

AY: \_\_\_\_\_

	Y/N/NA	If yes, describe extent of activity
<b>Curriculum research and development</b>		
Developed new course		
Developed new learning materials		
Engaged in research on pedagogy		
Increased accessibility and equity in a course or learning materials		
Substantially revised existing course		
<b>Educational service</b>		
Development or delivery of training, certificate, or continuing education programs		
Led discussion or gave presentation about teaching		
Mentored colleague or graduate student in teaching		
Observed/peer reviewed other instructors' teaching		
Taught audiences outside traditional Illinois students (e.g., high school club, student/community groups)		
Participated in or led committee related to teaching		
<b>Personal development</b>		
Earned instructional certification (e.g., from CITL)		
Participated in instructional development training or workshop		
<b>Student mentoring and supervision</b>		
Contributed to student learning outside courses (e.g. clubs, study abroad)		
Engaged in informal graduate mentoring		
Engaged in informal undergraduate mentoring		
Served as formal graduate student advisor		
Served as graduate student committee member		
Supervised graduate student research/projects		
Supervised undergraduate student research/projects		
<b>Other contributions or comments</b>		

### Appendix 10: Draft Personal Self-Evaluation of Instruction

Reflect on the course you just finished teaching. For each category, please check the one of the three boxes that best describes your own assessment of the course and your instruction. “Category” descriptions show which criteria in the Definition of Teaching Excellence applies: Well-designed DES, Well-delivered DEL, Inclusive and ethical IE, and Reflective and evolving RE. [This rubric was inspired by that used at U. Kansas.]

COURSE: \_\_\_\_\_ INSTRUCTOR: \_\_\_\_\_

Category	Area for improvement	Acceptable	Area of Strength
<p><span style="background-color: #90EE90; padding: 2px;">DES</span> <i>Are learning objectives clear?</i></p> <p><span style="background-color: #90EE90; padding: 2px;">DES</span> <i>Are learning objectives reasonably ambitious?</i></p> <p><span style="background-color: #90EE90; padding: 2px;">DES</span> <i>Are course learning objectives aligned with curriculum?</i></p>	<ul style="list-style-type: none"> <li>Students were not clear on the learning objectives.</li> <li>Learning objectives were not clearly stated in the syllabus.</li> <li>Learning objectives were either not at all challenging or unattainable</li> </ul>	<ul style="list-style-type: none"> <li>The course materials and I clearly articulated learning objectives to students</li> <li>Learning objectives were both challenging and attainable</li> </ul>	<ul style="list-style-type: none"> <li>Students demonstrated that they understood the course goals and how they related to their program or study.</li> <li>The learning objectives for this course were clearly connected to overarching program or curricular goals</li> </ul>
<p><span style="background-color: #90EE90; padding: 2px;">DES</span> <i>Is the course appropriately challenging and focused?</i></p> <p><span style="background-color: #FF8C00; padding: 2px;">RE</span> <i>Is the course up to date?</i></p> <p><span style="background-color: #ADD8E6; padding: 2px;">IE</span> <i>Is the course inclusive of the field's diversity?</i></p> <p><span style="background-color: #90EE90; padding: 2px;">DES</span> <i>Is the content aligned with the learning objectives?</i></p>	<ul style="list-style-type: none"> <li>Based on student performance or feedback, the content and materials were too easy or difficult and/or too narrow or broad.</li> <li>Content and materials needed to be modified to be up to date</li> <li>Content and materials needed to be modified to be inclusive of the field's diversity</li> <li>Course content was not clearly aligned with its learning objectives.</li> </ul>	<ul style="list-style-type: none"> <li>Based on student performance or feedback, the content was appropriately challenging and focused.</li> <li>Course materials represented current standards and were intellectually sound.</li> <li>Course content included some diversity of voices or content.</li> <li>Course content was somewhat aligned with its learning objectives.</li> </ul>	<ul style="list-style-type: none"> <li>Based on student performance or feedback, the content was appropriately challenging and focused, with integration across topics</li> <li>Course content was innovative and/or related to the cutting-edge of the field.</li> <li>Course content represented the full diversity of the field's practitioners, content, and applications.</li> <li>Course content was very well-aligned with its learning objectives.</li> </ul>
<p><span style="background-color: #FFFF00; padding: 2px;">DEL</span> <i>How well is time used in and out of class?</i></p> <p><span style="background-color: #FFFF00; padding: 2px;">DEL</span> <i>How well are assignments, assessments, and learning activities implemented to help students learn?</i></p> <p><span style="background-color: #FFFF00; padding: 2px;">DEL</span> <i>Is there evidence of student engagement with the course?</i></p>	<ul style="list-style-type: none"> <li>My course was disorganized, and I had trouble effectively implementing in class or out-of-class activities.</li> <li>I would benefit from learning more about best practices in teaching delivery.</li> <li>I was not sure how to engage students in this course.</li> </ul>	<ul style="list-style-type: none"> <li>I approached each class with a plan.</li> <li>I understand and follow standard course delivery practices in my discipline and institution.</li> <li>I provided a few opportunities for students to practice skills and learn knowledge embedded in course goals.</li> <li>I found that my students were often engaged.</li> </ul>	<ul style="list-style-type: none"> <li>Activities in my class were well planned, integrated, and reflected commitment to providing meaningful assignments and assessments.</li> <li>I used effective, high-impact, or innovative methods to improve student understanding.</li> <li>I used in- and out-of-class activities to provide coordinated opportunities for practice and feedback on important skills and concepts.</li> <li>My students showed high levels of engagement.</li> </ul>

Continued on next page

<p><b>Achievement of learning outcomes</b></p> <p><b>DEL</b> <i>Are students effectively learning content?</i></p> <p><b>DEL</b> <i>Are students developing critical, analytical, and creative thinking?</i></p> <p><b>IE</b> <i>Is learning monitored fairly and well?</i></p> <p><b>IE</b> <i>Are learners of all types and abilities benefiting from this course?</i></p>	<ul style="list-style-type: none"> <li>• The students achieved only a low level of skill/understanding.</li> <li>• This course didn't provide much opportunity for critical or creative thinking.</li> <li>• This course didn't do much to analyze or monitor student learning, and/or the assessments may sometimes be inconsistent.</li> <li>• Quite a few low- or high-achieving students were not really able to benefit from this course.</li> </ul>	<ul style="list-style-type: none"> <li>• Students' achievement was typical for a course like this.</li> <li>• Students showed some development of their critical, analytical, or creative thinking capacity.</li> <li>• I had clear and fair standards and approaches for evaluating the quality of student skill and understanding</li> <li>• Most students in the course were able to improve their knowledge and skills relative to the beginning of the term.</li> </ul>	<ul style="list-style-type: none"> <li>• Students' achievement was high for a course like this, and the quality of learning in this course supports success in other contexts (e.g., subsequent courses, non-classroom venues)</li> <li>• Students really improved their critical, analytical, or creative thinking.</li> <li>• I connected my goals for evaluating student understanding to program or curriculum expectations, or used authentic assessments</li> <li>• I took positive, successful steps to support learning in students or all abilities.</li> </ul>
<p><b>Course climate and inclusion</b></p> <p><b>IE</b> <i>Is the classroom climate respectful?</i></p> <p><b>IE</b> <i>Is the course designed to be inclusive of students with disabilities?</i></p>	<ul style="list-style-type: none"> <li>• I did not set expectations for civility and respect in my classroom; the environment might not have been welcoming for all students.</li> <li>• This course didn't take steps to be inclusive of students with disabilities.</li> </ul>	<ul style="list-style-type: none"> <li>• Through the syllabus and my daily interactions, I shared my expectations for civility and inclusion in the classroom; most students probably felt comfortable.</li> <li>• This course took some steps to support students with disabilities.</li> </ul>	<ul style="list-style-type: none"> <li>• My classroom climate was respectful and cooperative; nearly all students appeared comfortable, motivated, and engaged.</li> <li>• This course used best practices to support students with disabilities.</li> </ul>
<p><b>Reflection and evolution</b></p> <p><b>RE</b> <i>How has your teaching changed over time?</i></p> <p><b>RE</b> <i>How has this been informed by self-reflection, feedback and teaching development activities?</i></p> <p><b>RE</b> <i>How do you foster student feedback to inform your teaching?</i></p>	<ul style="list-style-type: none"> <li>• My teaching has changed very little over time.</li> <li>• I have not used self-reflection on prior teaching or feedback from students or peers to inform course development or teaching practices in subsequent offerings</li> <li>• I did not welcome students who wanted to provide constructive feedback on the course.</li> </ul>	<ul style="list-style-type: none"> <li>• My teaching has changed somewhat over time to improve.</li> <li>• I occasionally use self-reflection, feedback from peers and/or students, and/or instructional training to improve my course and my teaching.</li> <li>• I was receptive to students who wanted to provide constructive feedback on the course and my teaching.</li> </ul>	<ul style="list-style-type: none"> <li>• I have regularly made adjustments to improve my teaching.</li> <li>• I have often and systematically used self-reflection, evidence of student learning, and/or feedback from peers and students to improve my teaching and/or course design.</li> <li>• I actively encouraged student feedback and will use it to inform my teaching.</li> </ul>

## Appendix 11: Draft Formal Self-Evaluation of Instruction

Course name and number: \_\_\_\_\_

Instructor name: \_\_\_\_\_

First complete the Personal Self-Evaluation of Instruction rubric for this course. Then please answer each of the following questions in a couple of sentences with insights from the rubric in mind. This document will be part of your official teaching evaluation record.

Refer to the ***Campus-Wide Definition of Teaching Excellence*** when asked about relevant criteria:<sup>13</sup> they are well-designed, well-delivered, inclusive and ethical, reflective and evolving.

1) What were **two of your primary goals** in teaching this course, and why were those goals important to you? What criteria in the Definition of Teaching Excellence do they best apply to?

Goals	Criteria <sup>14</sup>
1)	
2)	

2) Specifically, what were your **two greatest successes** in teaching this course this term? What criteria in the Definition of Teaching Excellence do they match?

Success	Criteria
1)	
2)	

<sup>13</sup> The form will have a hyperlink to the Definition.

<sup>14</sup> In each column that asks for relevant Criteria, there will be a drop-down menu with the four criteria to choose from. This draft uses colored markers to indicate questions that themselves are markers of reflective teaching; those markers will not be on the final instrument.

3) What are the **two biggest challenges** you had in teaching this course this term? What criteria in the Definition of Teaching Excellence do they match?

Challenge	Criteria
1)	
2)	

**RE** 4) What did you do to try to overcome those challenges, and how did those efforts work?

Attempted solutions and outcomes
1)
2)

**RE** 5) What are **two things you would like to change** to improve this course and/or your instruction if you were to teach it again, and why? What criteria in the Definition of Teaching Excellence do they match?

Intended changes	Criteria
1)	
2)	

## Appendix 12: Draft Single-Year Holistic Evaluation of Instruction

The evaluator should check one box for each cell in this rubric using all student ratings, peer reviews, and self-evaluations for the year. Note:

- Instructor self-evaluations should be used to provide context for negative responses in student or peer inputs.
- Comments should provide reasons for findings and suggestions for improvement, and are especially important where input responses are negative or are mixed (either for a single course or between courses).
- An instructor need not have excellence in all sub-parts of an element for that to be an area of strength.
- Instructors need not exceed “acceptable” in the elements of teaching excellence in order to earn tenure (for tenure-track faculty) or contract renewal (for other instructors).

\*\*\*

Instructor: \_\_\_\_\_

Year: \_\_\_\_\_

List all courses taught by this instructor this year and check which materials were available and used to create holistic evaluation:

Course	Term	Student ratings?	Self-evaluation?	Peer review?

For peer reviews, give course and name of evaluator; add lines if more than one):

\_\_\_\_\_

Instructor: \_\_\_\_\_

Year: \_\_\_\_\_

Area for improve	Acceptable	Area of strength	<b>Criteria</b>
			<div style="text-align: center;"><span style="background-color: #90EE90; padding: 2px;">DES</span> <b>Well-designed</b></div> <ul style="list-style-type: none"> <li>Creates clear learning objectives that are both challenging and attainable, and intersect with the overall course goals.</li> <li>Designs student-centered learning and assessment experiences that support the learning objectives.</li> <li>Provides readily available, coherently organized, and high quality course materials (including syllabi) that establish student workload, learning objectives, performance criteria, grading procedures, and class policy expectations.</li> </ul>
<u>Comments:</u>			
			<div style="text-align: center;"><span style="background-color: #FFFF00; padding: 2px;">DEL</span> <b>Well-delivered</b></div> <ul style="list-style-type: none"> <li>Organizes and manages class time and students' activities in and out of class to yield high levels of student learning.</li> <li>Models and fosters critical, analytical, and creative thinking.</li> <li>Has constructive and timely written and oral communication with students.</li> <li>Gives timely, useful, and fair feedback on activities and assignments tied to performance criteria and learning objectives.</li> <li>Fosters the mindset that growth is possible and ability is not fixed.</li> <li>Engages, challenges, and supports students cognitively, emotionally, and behaviorally.</li> </ul>
<u>Comments:</u>			
			<div style="text-align: center;"><span style="background-color: #ADD8E6; padding: 2px;">IE</span> <b>Inclusive and ethical</b></div> <ul style="list-style-type: none"> <li>Designs instruction to be inclusive and equitable, ensuring that every student can participate fully and their presence and participation is valued.</li> <li>Represents the diversity of the field's practitioners, content, and applications.</li> <li>Recognizes the power differential between instructor and student and acts with integrity and respect toward students.</li> <li>Implements course policies that are equitable and applied fairly.</li> </ul>
<u>Comments:</u>			

Area for improvement	Acceptable	Area of strength	<b>Criteria</b>
			<div style="display: inline-block; background-color: orange; padding: 2px 5px; font-weight: bold;">RE</div> <b>Reflective and evolving</b> <ul style="list-style-type: none"> <li>Implements regularly-revised content and pedagogical techniques that are current, research-informed, rigorous, and relevant.</li> <li>Uses teaching development activities, evaluations, and other indicators to inform teaching practice and student learning experiences.</li> <li>Creates an environment that invites constructive student feedback to the instructor.</li> </ul>
<u>Comments:</u>			
			<b>Contributions in Self-Reported Inventory</b> <p><u>Area for improvement:</u> Makes very few contributions beyond required courses taught and/or poor execution on activities assigned.</p> <p><u>Acceptable:</u> Makes appropriate contributions beyond required courses taught for rank; work is acceptably well done.</p> <p><u>Area of strength:</u> Makes good contributions to teaching beyond required courses taught that exceed expectations for rank and/or work is exceptionally well done.</p>

**Summary table<sup>15</sup>**

	Area for improvement	Acceptable	Area of strength
<b>Criteria</b>			
Well-designed			
Well-delivered			
Inclusive and ethical			
Reflective and evolving			
Inventory			
<p><b>Summary comments:</b></p> <p>What was this person’s greatest strength in teaching this year?</p> <p>What is the area in which they have most potential for improvement, and how could they improve?</p>			

<sup>15</sup> All cells in this table except comments will automatically be populated from the entries in the rubric above.