Evaluating Teaching As If Learning Mattered:

A Guide to Good Practices at Georgia State University

Center for Instructional Innovation

Georgia State University

Purpose of this Booklet

This booklet is about good practices for evaluating teaching effectiveness and is intended to provide guidance to department chairs, promotion and tenure committees, and individual faculty. It should also be helpful to faculty who prepare graduate students to teach in higher education. The recommendations are drawn from the professional literature on faculty evaluation, the policies of the University (including the 2011-2016 Strategic Plan and Faculty Handbook), and from conversations with Georgia State University chairs and former chairs about evaluating teaching effectiveness. **Our conversations with department chairs emphasized the need to include faculty in discussions about changing policies and practices in how teaching evaluated. The following material is designed to assist in those conversations, as well as in to help set some goals to improve our practices.**

At the heart of these suggestions is the recognition that the goal of the instructor is to promote students' mastery of the content, skills, and values and ethics that have been identified by the department and University. As noted in Goal 1 of the University's Strategic Plan*, Our goal is to continue to position the University as a model of academic success for diverse populations and to lead among research institutions in producing engaged citizens of the 21st century.*

Evaluations of teaching effectiveness have formative and summative purposes: formative in that evaluations provide feedback to instructors for improving their teaching (and student learning), and summative because teaching evaluations are used to judge instructors for the purposes of merit raises, promotion, and continued employment. With this in mind, the goals of this booklet are to suggest strategies for evaluating teaching effectiveness (e.g., in annual evaluations, third-year pre-tenure reviews, promotion and tenure evaluations, and post-tenure reviews) in ways that:

- Promote deep and enduring student learning;
- Document teaching effectiveness in a fair, accurate, and efficient manner; and
- Encourage instructors to explore and adopt innovative and effective strategies that improve student learning.

Because teaching effectiveness is nearly always defined in terms of the impact of the instructor on student learning, the ideal way to evaluate teaching would be through evidence of student learning. However, determining the impact of an instructor on student

learning can be difficult. Experts in documenting teaching effectiveness (e.g., Fink, 2008) recommend using the following four sources for the most complete and accurate measure of teaching effectiveness: How instructors engage students (Student Evaluations of Instruction), the course design (syllabi and course materials), evidence of attempts to improve teaching (instructors' descriptions of using feedback from students and colleagues, as well as classroom and scholarly resources), and evidence of the resulting student learning.

Student Evaluations of Instruction and Beyond

At Georgia State University, as in nearly all universities, the student evaluations of instruction (SEI) are the core of how we evaluate teaching. Unfortunately, focusing primarily on student perceptions of instructors' effectiveness does little to promote effective teaching (Fink, 2008). SEIs are at best, only a proxy for the fundamental question, "Did student learn what they needed to learn"? Research indicates that SEIs are only modestly related to how much students learn (e.g., Arthur, Tubre, Paul, & Edens, 2003; Stark-Wroblewski, Ahlering, & Brill, 2007). That is to say, student evaluations are not a proxy for how much students learn in a class. Of course, SEIs should always be part of the evaluation process, but are best used as a measure of instructor/student engagement. In discussions at the Center for Instructional Innovation our panels of GSU department chairs emphasized that teaching evaluations need to look beyond the numbers.

While SEIs are best viewed as an assessment of how the instructor has engaged the students and engaged the course content, SEIs are an important **formative assessment** to guide instructors. To make the best use of SEI scores as a **summative assessment**, some GSU chairs examine a "basket" of items they considered essential to promoting student learning and for which students can provide valid information. For example, on the College of Arts & Sciences SEI form, items such as, "*Returned test results and evaluations of my work in a reasonable period of time; Spoke in a way that communicated the subject in an understandable manner; Responded constructively and thoughtfully to questions and comments";* and "*Motivated me to learn*" provide more focused and helpful information than the single item of rating the effectiveness of the instructor. Likewise, on the Robinson College of Business SEI form, an instructor's rating on "*Motivated me to do my best work*" and "*Made me work harder than in most courses*" help to assure appropriate rigor in courses. For some other SEI items, however, students are not likely to have the ability to make an accurate judgment, e.g., "The instructor cares about the quality of his/her teaching".

Chairs at GSU concurred with the professional literature (e.g., Seldin, 2006) that when evaluating SEIs, it is important to take into account the level of the course (core, upper-division, graduate), the student composition in the course (majors vs. non-majors) and the general range of SEIs for others who have taught the same course. Likewise, if over a three-year period one instructor taught

new preparations each semester while a colleague repeatedly taught the same course, consideration should be given for the mitigating circumstances of having many, new preparations. Chairs also talked about uncovering the story behind the SEI numbers. For example, when most students have rated an instructor in the 4-5 range but a subset of students scored all items as 1 or 2, there is likely a reason for the outlier scores that should be explored with the instructor (perhaps low evaluations were by a few students who were not majors in the department). In the case of widely discrepant ratings by students, the median SEI score often provides a more accurate measure of "central tendency" than using the mean.

The professional literature (e.g., (Arreola, 2000; Pallett, 2006) suggests that when we use SEIs, it is a good practice to:

- Use no more than three to five categories when rating multiple instructors' SEI scores (e.g., Excellent, Above Average, Meets Expectations, and Fails to Meet Expectations"). Trying to make fine discriminations, such as ranking instructors on SEIs, almost certainly goes beyond the technical limits of the instrument.
- Use a sufficient number of classes--six to eight classes taught by an instructor and consider SEIs changes over time, rather than a snapshot view.
- Limit Student Evaluations of Instruction to account for no more than a third to a half of the overall evaluation of teaching. **See the first paragraph of Final Thoughts (page 7) for suggestions on adding some of the components which follow.**

As our instructional practices evolve from a primarily lecture format to more student-centered, field-based, online, and hybrid models, we will need to continually revise items on our SEIs and even consider alternative forms that would align with our instructional formats. For example, an SEI item such as "Speaks clearly" is certainly not relevant for an online class.

Urban Myths: Contrary to what many instructors believe, there is no evidence in the literature that making a course easier results in higher student evaluations. Likewise, although there is a modest correlation between the grades students earn and how they rate the instructor, the converse of giving more "A" grades doesn't raise SEI scores (Marsh & Roche, 2000).

Practical Considerations

In spite of the obvious limitations of using SEIs for summative evaluations of teaching, there are heuristic reasons for continuing to use them. The main benefit is the efficiency of using a few numbers to judge a very complex set of skills. For the annual review of a tenured instructor whose career is marked by excellent teaching evaluations supported by high-quality student learning, the summary scores on SEIs may provide enough information to document continued good work. However, for a post-tenure review, we would expect closer scrutiny, including reviewing course design materials, evidence of student learning, and efforts to improve student learning since the previous review. On the other hand, for a non-tenured instructor or a lecturer or clinical instructor whose major role is teaching or the tenured faculty member whose SEIs have been below program standards, we would expect all reviews to include assessing course design, evidence of student learning, and documentation of efforts to improve, described in the sections which follow.

Course Design for Learning

The course syllabus, especially the student learning objectives, course assignments, assessments, and materials, give a picture of how an instructor has designed the learning activities and the level of rigor in the course. Section 401.01 of the GSU Faculty Handbook specifies the elements to be included in each course syllabus. For example:

- Are there student learning outcomes that define the content and skills they are to master?
- Are the learning outcomes stated as observable behaviors? Rather than stating "students will learn about the social impact of the advancement of modern technology", the learning outcomes specify what students will do to demonstrate what they have learned (e.g., "students will write a research paper comparing and contrast advantages and disadvantages of the the advancement of modern technology for working education).

The syllabus can also show how an instructor has incorporated current and appropriate materials and resources, and has engaged students in learning activities and prepared assessments that match the learning outcomes and standards of rigor. For example, are there multiple opportunities throughout the course for students to demonstrate their knowledge and skills and receive feedback from assessments of their progress? Likewise, examples of course assessments, including examinations, papers, and projects can provide an insight into the extent to which students are expected to integrate and apply the information in the course. An example of a rubric for judging syllabi is at **http://www.gsu.edu/images/CTL/Syllabus_Rubric.pdf**. However, to see the extent to which the instructor's plans in the syllabus have been executed, we need to examine student learning.

Evidence of Student Learning

Research indicates that SEIs are only modestly related to how much students learn (e.g., Arthur, Tubre, Paul, & Edens, 2003; Stark-Wroblewski, Ahlering, & Brill, 2007). That is to say, how much students learn in a class is an independent variable, not directed reflected in their evaluations of the course. Therefore, when instructors add information that documents learning, they provide an additional, essential element about teaching effectiveness. Faculty might show pre-post test results; students' performance on departmental examinations, and examples of improvement in students' work across a semester (perhaps shown by a rubric that assesses complex skills, such as critical thinking), including examples of how faculty feedback to students on a project or paper resulted in improved work by the end of the semester (Montgomery, 2002)**.**

Student learning portfolios with archives of learning products and students' reflections about learning experiences can be a core element in documenting evidence of learning for both programs and individual instructors.

Evidence of Efforts to Improve Teaching

Efforts to improve teaching fall into two categories. An instructor may endeavor to increase low student SEIs and/or improve student performance, e.g., remediating teaching/learning problems by assessing students' prior knowledge. In this case, changes in SEIs as well as student learning might be monitored. Peer and chair observations can play an important role in mapping a path for improvement and documenting changes. On the other hand, an instructor may be trying to further enhance teaching effectiveness and student learning through the use of innovative pedagogy, including enhanced technology and strategies from the professional literature (scholarly teaching). Assuming that this instructor has demonstrated effective pedagogical skills, evidence of improved learning would become the key measure of effectiveness, even though SEIs might temporarily fluctuate. A set of tools to provide formative feedback to instructors during a course is available at **http://www.gsu.edu/cii/assessment.html**. **Our role as a research university should extend to supporting and rewarding instructors for systematically examining and sharing evidence of innovative attempts to improve student learning, including recognizing that SEIs may fluctuate while innovations are being implemented.**

Other information to consider
While evidence from the classroom is at the core of teaching responsibilities, there are other responsibilities related to teaching which should be considered, if relevant. For example, credit should be given for redesigning a course or program (especially when undertaken at the request of the program), for advising and mentoring students (including guiding students doing research and completing theses, dissertations, or special projects). Similarly, recognition should be given for mentoring graduate students and colleagues in improving their class skills as well as for awards and recognition for teaching excellence.

Teaching Portfolios
An ideal way for an instructor to organize and present information on teaching effectiveness is through a teaching portfolio (Seldin, Miller, & Seldin, 2010). A well-written teaching portfolio is not just a collection of artifacts but an integrated document that includes:

- A philosophy of teaching that might describe how classroom experiences (teaching and learning), the professional standards and practices of the discipline, pedagogy conferences, literature on teaching and learning, and colleagues helped to shape what she/he understands about motivating, engaging, and assessing students;
- Course materials (syllabi, assignments, assessments, and activities);
- Description of courses taught, developed, and revised;
- Reflections and descriptions of efforts to improve; and
- Archives of students' work and how the instructor guided them to deep and enduring mastery of content and skills.

The portfolio would show how the philosophy had driven the course design through syllabus, course materials and resulting student work. For information on writing a teaching philosophy, see **http://chronicle.com/article/How-to-Write-a-Statement-of/45133/**
Portfolio integration becomes even more evident if it is posted to a website so that others can follow the hyperlinks showing the relationship between the philosophy, course materials, student products, and efforts and results of improvement. Online teaching portfolios can even include video clips of classroom activities.

Final Thoughts
If current practices for evaluation of teaching effectiveness primarily focus on the use of student evaluations of instructors, the steps to broadening the process to include information suggested above probably best begins with conversations with program faculty. Individual instructors might be invited to set annual personal goals for improving/renewing their practice and enhancing student learning (e.g., addressing the performance of students who are struggling in a class or enhancing an aspect of the learning of all students). The intent is to put a premium on reflection, innovation, and improvement while minimizing any risk of temporary drops in SEIs. While revised course materials is the first step in documenting this process, documenting improvement in students' mastery

of content and skills is the ultimate goal. An example of how this might be negotiated with an instructor could be to individualize the elements and weighting for evaluating teaching: SEIs (40%), evidence of excellence in course design and efforts to improve (25%), evidence of high-quality student learning (25%) and evident of advising and mentoring students and colleagues (10%). Over time, the emphasis on student learning could increase and SEIs decrease.

Most instructors shutter when students approach a course asking, "What's on the test"? Likewise, if the primary measure of teaching effectiveness is students' perceptions, as expressed in the Student Evaluation of Instruction forms**, we risk encouraging faculty to teach to the test of the SEI**. If student learning (and the high-quality teaching which produces it) is important to us, we need to use multiple measures that assess and encourage instructors to teach classes that have high standards for deep and enduring mastery of content and skills and are based on our best understanding of teaching and learning in higher education.

References and further Reading

Arreola, R. A. (2000). *Developing a comprehensive faculty evaluation system: A handbook for college faculty and administrators on designing and operating a comprehensive faculty evaluation system* (2nd ed.). Bolton, MA: Anker.

Arthur, W. Jr., Tubre, T., Paul, D. S. & Edens, P. S. (2003) Teaching effectiveness: the relationship between reaction and learning evaluation criteria, *Educational Psychology,* 23(3), 275–285.

Fink, D. (2008). Evaluating teaching: A new approach to an old problem. *To Improve the Academy: Resources for Faculty, Instructional, and Organizational Development* , Vol. 26, pp. 3-21.

Marsh, H. W. & Roche, L. A (2000). Effects of Grading Leniency and Low Workload on Students' Evaluations of Teaching: Popular Myth, Bias, Validity, or Innocent Bystanders? Journal of Educational Psychology, 2000, Vol. 92, No. 1. 202-228.

Montgomery, K. (2002). Authentic tasks and rubrics: Going beyond traditional assessments in college teaching. *College Teaching.* 50: 34-39.

Pallett, W. (2006). Uses and abuses of student ratings. In P. Seldin (Ed.), *Evaluating faculty performance: A practical guide to assessing teaching, research, and service* (pp. 50-65). Bolton, MA: Anker.

Seldin, P. (2006). *Evaluating faculty performance: A practical guide to assessing teaching, research, and service*. Boston: Anker.

Seldin, P, Miller, J. E., & Seldin, C. A. (2010). *The teaching portfolio: A practical guide to improved performance and promotion/tenure decisions—4th ed.* San Francisco: Jossey-Bass.

Stark-Wroblewski, K., Ahlering, R. F., & Brill, F. M. (2007). Toward a more comprehensive approach to evaluating teaching effectiveness: supplementing student evaluations of teaching with pre-post learning measures. *Assessment & Evaluation in Higher Education*, *32*(4), 403-415.

NOTES:

Point of Contact	Phone: (404) 413-2541 Fax: (404) 413-2558 Email: cii@gsu.edu Website: www.gsu.edu/cii
Dr. George Pullman, Director gpullman@gsu.edu Dr. Harry Dangel, Director Emeritus hdangel@gsu.edu Associate Directors: Dr. Brennan Collis, Writing Across the Curriculum Unit, brennan@gsu.edu Dr. Jennifer Lawrence, Critical Thinking through Writing Unit, Jlawrence2@gsu.edu Dr. Sharon Cavusgil, English Language Support Services, scavusgil@gsu.edu	Campus Location: 75 Piedmont Avenue, Suite 502, CTB Atlanta, GA 30303 Mailing Address: PO Box 3987 Atlanta, GA 30302-3987 Please visit our website for Faculty and GTA events and services.