

4-10-2014

Final SRI Report

Georgia Southern University

Ad Hoc Committee on Student Ratings of Instruction

Follow this and additional works at: <https://digitalcommons.georgiasouthern.edu/faculty-senate-index>



Part of the [Higher Education Administration Commons](#)

Recommended Citation

Georgia Southern University and Ad Hoc Committee on Student Ratings of Instruction, "Final SRI Report" (2014). *Faculty Senate Index*. 204.

<https://digitalcommons.georgiasouthern.edu/faculty-senate-index/204>

This other is brought to you for free and open access by the Faculty Senate at Digital Commons@Georgia Southern. It has been accepted for inclusion in Faculty Senate Index by an authorized administrator of Digital Commons@Georgia Southern. For more information, please contact digitalcommons@georgiasouthern.edu.

Final Report
Georgia Southern University
Ad Hoc Committee
on Student Ratings of Instruction

Helen Bland (JPHCOPH)
James Harris (CEIT)
James Jupp (COE)
William Levernier (COBA)
Nan LoBue (CLASS)
Trent Maurer (CHHS)
James Reichard (COSM)
Sonya Shepherd (Library)
Rebecca Ziegler (Library, Chair)

Submitted
April 10th, 2014

Table of Contents

[Section I: Introduction](#)

[Section II: Background](#)

[Section III: SRI Faculty Results](#)

[Section IV: SRI Administrator Results](#)

[Section V: Recommendations](#)

[Section VI: References](#)

[Appendix: Updated Skowronek et al. \(2011\) Measure](#)

Section I: Introduction

In the Fall of 2013, the Senate Executive Committee of the Faculty Senate at Georgia Southern University created an Ad Hoc Committee to conduct a comprehensive review of the Georgia Southern University Student Ratings of Instruction [SRI] instrument and its use. The committee was charged to

- i. Determine the purpose of Student Ratings of Instruction (e.g., what is their role in evaluating teaching effectiveness?).
- ii. Review the effectiveness of the current instrument in meeting that purpose.
- iii. Determine how student ratings of instruction fit with other assessment measures.
- iv. Evaluate how frequently the instrument is administered.
- v. Evaluate how the instrument is administered to students (e.g., the merits of electronic versus paper submissions).
- vi. Examine how the instruments are processed after being collected.
- vii. Recommend ways the instrument and its administration might be improved.

The Ad Hoc Committee was composed of one faculty representative nominated by the Senate Executive Committee from each College and the Library. Committee members included two former department chairs, an expert in survey construction and administration, an expert on the Scholarship of Teaching & Learning [SoTL] and specifically the research on SRIs, and at least one teaching award winner at the departmental, college, university, and university system levels.

During the Fall 2013 semester, the Ad Hoc Committee created two surveys, one for faculty and one for department chairs, to assess faculty and administrator opinions. In January of 2014, the Ad Hoc Committee announced via email the two online surveys they had created and invited all Georgia Southern University faculty and department chairs to provide their feedback via the respective surveys. A total of 234 faculty members (out of an estimated 950) and 21 department chairs (out of 37) responded to the surveys. This reflects response rates of 25% and 57%, respectively, both greatly in excess of what is typical for online surveys. This response rate is likely an indicator of high levels of motivation to comment on the SRI form and its use.

Both surveys asked participants a series of open-ended questions about the SRI form and its use as well as 1-3 demographic questions. Results were analyzed for emergent themes with exemplar quotations identified for each question in each version of the survey. Results are presented by survey version; first, the faculty version, next, the administrator version. Within each section, a general summary of results is presented first, followed by demographic data and finally summaries and exemplar quotations for each question in the survey. For each question, some of the completed surveys did not address any of the themes and some addressed more than one. The number of times each theme was addressed is reported as a raw number and a percentage of all respondents to that question.

Section II: Background

Research on Student Evaluations of Teaching [SET], as they are known in the pedagogical literature, is extensive. Over 2,500 peer-reviewed pieces of scholarship on the topic have been published, making it one of the most prolific areas of research in all of the Scholarship of Teaching & Learning. Although controversy exists, and some questions have not yet been settled, there is sufficient evidence to draw several key conclusions about SETs.

SETs are typically used for two purposes: a) formatively, to improve teaching effectiveness, and b) summatively, to evaluate faculty performance as teachers. These two purposes cannot always be effectively assessed with a single measure (Hoyt & Pallett, 1999). The University System of Georgia [USG] Board of Regents [BOR] [Policy Manual Section 8.3.5.1](#), paragraph 1, states, “Each institution, as part of its evaluative procedures, will utilize a written system of faculty evaluations by students, *with the improvement of teaching effectiveness as the main focus of these student evaluations.*” [emphasis added] Thus, it is the explicit goal of the USG BOR that student evaluations of teaching should be used in a primarily formative way with only secondary attention to their value in summative evaluations of faculty. This BOR mandate, within which all USG institutions must operate, also guided the work of the Ad Hoc Committee.

To evaluate or improve teaching effectiveness, it is first necessary to define what constitutes an “effective teacher” (Gibbs, 1995). Skowronek, Friesen, and Masonjones (2011) proposed a six-part definition explicitly for the purposes of designing an effective SET instrument:

. . . an effective teacher: 1) creates an active learning environment to engage students (Angelo, 1993), 2) makes an attempt to identify students’ prior knowledge about a topic and goals for a course (Perry, 1970), 3) attempts to make course content meaningful to the “real-world”, 4) attempts to develop deep levels of understanding and help students reflect on that understanding (i.e., critical thinking) (Halpern, 1999), 5) should remain excited and enthusiastic about the material they are teaching (Voss & Gruber, 2006) and 6) is committed to personal growth within the discipline (Lowman, 1995). (p. 3)

Additionally, because the end goal of teaching is ultimately student learning, “the best criterion of effective teaching is student learning” (Cashin, 1988). Further, “teaching effectiveness depends not just on what the teacher does, but rather on what the student does” (McKeachie & Hofer, 2001, p. 6). Thus, any meaningful evaluation of teaching effectiveness, by definition, must also assess student learning and learning behaviors. Using SETs to ask students questions about their own learning has the added benefit of focusing students’ attention on the ultimate purpose of all teaching activities (Titus, 2008).

Unfortunately and ironically, most SETs do not actually assess teaching effectiveness. As Titus (2008) notes, “Although the term *student evaluation of teaching* (SET) is commonly used, most researchers agree that the rating scales solicit *student opinions* (e.g., Powell 1978) and provide indications of *student satisfaction* (e.g., Abrami, d’Apollonia, and Cohen 1990).” (p. 416). Further,

Many students report that rather than reading the actual rating items, they locate a column on the form to reflect their general level of enjoyment in the course and then mark all of the rating items in that same column at that same value (Titus, 2008, p. 402)

Even more problematically, “What these students define as teaching excellence, then, may be what reinforces their unexamined conventional assumptions that professors actually aim to question through teaching.” (Titus, 2008, p. 409).

As the American Educational Research Association (2013) notes,

To evaluate teaching, focus on student learning outcomes. Ideally, a system to evaluate education faculty as teachers will do three things:

- Help institutions define “teaching quality” based on student learning outcomes;
- Help faculty members improve their teaching by identifying where they need professional development; and
- Help evaluators determine a faculty member’s relative strengths and weaknesses as a teacher.

The evaluation method we most often turn to—student ratings, sometimes supplemented by measures of teaching productivity such as the number of advisees—does none of these things well. Student ratings. . . do not promote student-centered learning, and they do not identify and reward the most effective teaching practices (Healey & Jenkins, 2003; Hutchings, Huber, & Ciccone, 2011; Singer et al., 2012). (p. 3)

The psychometric quality of most SETs is thus questionable at best (Abrami, d’Apollonia, & Rosenfield, 1997; Kulik, 2001; Wachtel, 1998), and SETs take a simplistic approach to teaching effectiveness (McKeachie, 1997), requiring students to judge elements of teaching that they lack the background or knowledge to evaluate, including: a) the appropriateness of class objectives, b) instructor knowledge of the material, c) the fairness of graded materials and assessments, and d) and the relevance of course materials (Seldin, 2006).

Further, multiple variables that have nothing to do with teaching effectiveness have been identified as biasing influences on SET scores (Aleamoni, 1999; Cashin, 1995; Feldman, 1993; Pratt, 1997; Wachtel, 1998). Examples of such biasing influences include:

- Course Characteristics:
 - Subject Discipline (Cramer & Alexitch, 2000; Franklin, 2001; Nerger, Viney, & Reidell, 1997)
 - Required vs. Elective Courses (Algozzine et al., 2004; Cashin, 1995; Franklin, 2001)
 - Course Level (Algozzine et al., 2004; Cashin, 1990; Nerger et al., 1997; Schlenker & McKinnon, 1994)
 - Class Size (Algozzine et al., 2004; Cashin, 1995; Cramer & Alexitch, 2000; Nerger et al., 1997; Schlenker, McKinnon, 1994)
 - Course Difficulty or Rigor (Addison, Best, & Warrington, 2006)
- Student Characteristics
 - Expected or Actual Grade (Cashin, 1995; Franklin, 2001; Heckert, Latier, Ringwald, & Silvey, 2006)
 - Student Motivation (Heckert et al., 2006)
- Instructor Characteristics
 - Instructor Gender (Feldman, 1993; Centra & Gaubatz, 2000)
 - Instructor Race (Hammermesh & Parker, 2005; Hendrix, 1998; Rubin, 1998)
 - Instructor Attractiveness or Expressiveness (i.e., the “Dr. Fox” effect) (Naftulin, Ware, & Donnelly, 1973; Marsh & Roche, 1997)

The findings on the influence of course difficulty/rigor and student motivation on SET scores is important to examine within the context of federally mandated standards for collegiate level work and students' self-reported study behaviors. The United States Department of Education Office of Postsecondary Education defines a credit hour as:

An amount of work represented in intended learning outcomes and verified by evidence of student achievement that is an institutionally established equivalency that reasonably approximates not less than—(1) One hour of classroom or direct faculty instruction and a minimum of two hours of out of class student work each week for approximately fifteen weeks for one semester or trimester hour of credit (U.S. Department of Education, 2011, 34 CFR 600.2)

Thus, a student taking a 15-credit hour semester course load should be spending a minimum of 30 hours (15 x 2) per week out of class on academic work. Data from Georgia Southern University students on the 2011 National Survey of Student Engagement [NSSE] (Georgia Southern University, n.d.) reveals that less than 7% of first-year students and less than 12% of seniors are spending that much time out of class on academic work. The modal student is spending less than one third of the required time on out of class academic work. Yet, 88% of first-year students and 81% of seniors report spending “quite a bit” or “very much” time “studying and on academic work.”

These findings bleed into the relationship between student grades and SETs, which is particularly problematic because evidence from the literature suggests that some faculty members may “water down” content and rigor and inflate student grades in order to receive higher evaluation scores. Recent research has established that SET scores and grades are positively correlated, but SET scores are unrelated to learning outcomes after controlling for grades (Weinberg, Hashimoto, & Fleisher, 2009). Additionally, students taught by instructors with higher SET scores demonstrated higher performance on common course examinations, but follow up investigations with the same students revealed such students performed *worse* in later courses in the same sequence than students taught by instructors with lower SET scores (Carrell & West, 2010).

Taken together, these findings reveal that most SETs do not assess what they purport to assess, do not provide reliable summative data on teaching effectiveness, and are known to be vulnerable to racism, sexism, and other forms of discriminatory bias against protected classes. To use SETs in summative ways to evaluate faculty performance would introduce such biases in to the evaluation processes themselves, which would violate both Georgia Southern University policy and state and federal law.

Additionally, using SETs may prove ineffective or even harmful in promoting teaching effectiveness if the SETs themselves are not designed with formative goals as their primary purpose:

Researchers, though, have yet to find any direct evidence that the institutionalization of SETs results in an improvement in higher education overall or produces more excellent teachers or more learned students (Olivares 2004). There is no evidence that SETs are sufficient in themselves to create any measurable improvement in instruction (e.g., Kember, Leung, and Kwan 2002; Marsh and Hocevar 1991). . . Ironically, some researchers have found student ratings to have unintended negative effects on educational quality through decreasing faculty morale and inducing lowered academic standards and grade inflation (Greenwald and Gillmore 1997a; V. E. Johnson 2003; Ryan, Anderson, and Birchler 1980). (Titus, 2008, p. 398)

When standard rating forms are used to assess teaching, “they become de facto the operational definition of effective instruction” (d’Apollonia and Abrami 1997b: 51) and thereby, as Kolitch and Dean (1999) observe, can militate against forms of teaching concerned with critical thinking or transformative pedagogy. (Titus, 2008, pp. 401-402)

However, SETs specifically designed to focus on student learning and to provide formative feedback as their primary goal represent “best practices” in their use and can be highly effective. For example, the American Educational Research Association (2013) recommends “that evaluations of faculty teaching focus on what and how students learn, and that they use evidence-based criteria for assessment.” (p. 3). The University of Wisconsin-La Crosse (2007) notes,

Student evaluation of instruction should be only one of several measures of teaching effectiveness. Additionally, SEI scores should be interpreted within the context of variables known to be related to evaluations (e.g. student motivation, class size, discipline, etc.) and in general, it is recommended that SEI scores not be compared across instructors. . . it is recommended that absolute cut-off scores (e.g. SEI scores must be above 3.5 for a candidate to be considered for tenure) be avoided.

If SET scores should not be compared across instructors nor to “cut-off” scores, it is clear that they are being interpreted qualitatively with respect to student learning. As referenced above, Skowronek et al. (2011) developed an SET form specifically to evaluate “effective teaching” from a student learning perspective. Their work represents the most recent thinking on best practices in the use of SETs and their description of their development process beautifully illustrates the modern shift in thinking about how SETs are conceptualized.

Along with the set of new assessment items, a new rating scale was created. This scale was adapted from a model used at the University of California-Berkley in assessing “student learning gains” (UC Regents, 2000). Rather than asking if students agreed or disagreed with a statement (on a five point scale ranging from strongly disagree to strongly agree), students are now being asked whether or not a certain component helped their learning (on a five point scale ranging from “did not help my learning” to “helped my learning a great deal”). . . This was a dramatic shift in the student evaluation instrument as focus was shifted from emotive responses regarding instructional methods to a focus on what the instructor does to facilitate learning (i.e., a student might not agree with the presentation style an instructor used, but he or she could still learn in such an environment). . . To separate potential confounding or multidimensional issues of teaching, no questions assumed any quality or component was present in the classroom. Instead, additional items were added to allow students an opportunity to first provide information about the level of certain components. (pp. 3-4)

Skowronek et al.’s (2011) results were equally promising:

We would also expect a relationship to grade with this new survey; the higher the grade the more learning has presumably taken place. Even though numerous items are predictors of teaching effectiveness, it is important to note that the most variance of the teaching effectiveness score that can be explained across both time points is small at best. More importantly, although significant, the unique variance explained for by grade was a very small ($sr^2 = .046$ at Time 1 and $.013$ at Time 2) predictor of the teaching effectiveness score. This suggests that, even though grade was a significant predictor, student ratings were not largely driven by the grade they believed they

were going to receive in the course, which could reduce the need to “dumb-down” a course (Huemer, 2005) or artificially inflate grades to get high ratings. (p. 10)

This approach to SETs, then, more accurately assesses teaching effectiveness by refocusing SETs on what teaching effectiveness is all about: student learning. Further, “faculty participants believed this survey [compared to a more traditional version that did not focus on student learning] would be much more useful in a summative format as an instructor progresses through the tenure and promotion process.” (Skowronek et al., 2011, p. 11). Such student-learning focused SETs would also more closely align with the USG [Policy Manual Section 8.3.5.1](#) and its explicit focus on the use of SETs for the improvement of teaching effectiveness. They would also better reflect the Georgia Southern University Faculty Handbook definition of superior teaching as “focused on student learning outcomes” (205.01). It is within this broader context that the results of the faculty and administrator surveys are presented and discussed below.

Section III: SRI Faculty Results

General Summary

Faculty responses indicated a high level of frustration and anger with the current SRI form and its use. Many faculty expressed multiple concerns about lack of validity of the current measure, its vulnerability to factors that are irrelevant to teaching, its insufficient focus on student learning, and overuse and misuse by administrators in evaluating faculty teaching. Many faculty comments echoed findings from the literature reviewed in the Background section.

Demographic Data (N = 233)

Participants' College (N = 217, 16 unreported)

COBA: 31
COE: 28
CEIT: 17
CHHS: 29
CLASS: 71
COSM: 23
JPHCOPH: 11
Library: N/A
Other: 7

Academic Rank (N = 217, 16 unreported)

Full Professor: 47
Associate Professor: 74
Assistant Professor: 53
All Other Ranks: 43

Tenure Status (N = 219, 14 unreported)

Tenured: 126
Tenure track: 46
Non-tenure track: 47

Questions, Result Summaries, & Exemplar Quotations

- 1. Generally speaking, what are your thoughts about the current Student Ratings of Instruction instrument and how it is used in your department/unit and college?**

The following seven general themes emerged from the surveys (N = 224 respondents). Most respondents offered an opinion about the effectiveness of the current SRI instrument in evaluating teaching effectiveness and/or critiques of its use, which were subsequently elaborated in later questions.

Theme 1 (N = 9, 4%): Very Good or Good

“I am pleased with it.”

“I find the SRI instrument to be very effective in improving my course structure and teaching methods. I rely heavily on student responses when preparing for the next semester. I am unsure how my department handles the evaluations once collected.”

“It is good.”

Theme 2 (N = 52, 23%): Fairly Good / Fine / Adequate / Somewhat Effective / Moderately Effective / or some other similar description.

“I don't have any complaints about the content of the rating system.”

“I think it works fine for its purpose.”

“I think it is sufficient to do the job it is supposed to do.”

Theme 3 (N = 63, 28%): Not Effective / Limited Effectiveness / Minimally Effective / or some other similar description.

“Generally speaking, they are relied on heavily to assess teaching. Currently there is no other method of reviewing teaching, although peer reviews of teaching are "recommended." Numbers are compared between faculty instead of taken into context of course level, type of course (methods, etc.), size of course, etc. Comments from students are focused on in letters for annual review or tenure (pretenure) review without any opportunity for a faculty member to respond, explain or defend. These letters are kept on file while tenure review statements are not.”

“I believe it carries too much weight because it is used as the singular measure”

“In my department, it's used to punish faculty. Mean scores are calculated when a class size doesn't allow for statistical significance. Negative comments from students are used against faculty, but positive comments are ignored. A secretary types up the comments and if a student has something in the negative column such as, "Nothing bad, I like the class," it is still used as negative.”

Theme 4 (N = 13, 6%): Completely Ineffective / Very Ineffective / Worthless /or some other similar description.

“I think it is a colossal waste of time. It is purely a popularity contest. 18-21 year old young people are completely unqualified to evaluate me as an instructor. In my 15 years of teaching here at GSU, I have only received 3 or 4 constructive comments that ended up changing the way I taught the class. The remainder of the comments were purely personal.”

“I think that the current instrument is really stupid. It does not provide useful feedback. I have been teaching at Georgia Southern for 25 years. With the current instrument, my student evaluations have *NOT* changed a bit. In fact, it is possible that they have actually gotten better. Note: Generally, as an instructor ages, one expects the

student evaluations to decline. I do not think the student evaluations have any meaningful use in our department or college, unless they are at the extreme. However, I have served on tenure/promotion committees where an instructor has received consistently pitiful student evaluations, yet are rated as excellent teachers by their department chair, tenure/promotion committee, etc... I have seen poor student evaluations used to force a faculty member to resign, too. The current instrument is used by the administration in whatever way it wants that makes them fulfill the administrations agenda.”

“It is a ‘cop out’ in the sense that everyone I have spoken with agrees (faculty and administrators) it is a terrible way to evaluate professors and has provoked a significant dumbing down of much of the curriculum. However, because people like numbers and have nothing else, it is heavily relied up as some way to evaluate teaching. Most faculty, though, seem to find little value in reading over comments and do not learn much from student responses.”

“It's a methodologically unsound instrument that is misinterpreted and overinterpreted to ‘evaluate’ teaching. Faculty scores are inappropriately computed, without adequate controls, and idiosyncratically compared to other faculty members' scores, again without adequate controls or the use of appropriate statistical tests.”

Theme 5 (N = 55, 25%): The SRI is vulnerable to factors that have nothing to do with teaching effectiveness (is too influenced by course difficulty, the professor’s standards, the level of the course, and so on). It is a popularity contest.

“As usual, what seems to be a very functional tool has too many variable factors to have much reliability and validity. If you are a faculty who sets high standards for class performance and student expectations (not unreasonable at all) and don't give extra credit to bump up the grade or allow a cushion of free points in the system, then you will be crucified on the form no matter how good your teaching may have been. I have tested it out in various semesters and my teaching ratings are directly related how much fluff is allowed in the grade. To then have these form the entire or mostly entire basis for teaching evals makes it a very conflictual experience on the whole.”

“I believe there is too much weight on a professors teaching ability being utilized by SRI. A professor can be the greatest person and give ‘easy’ grades to students just to get high SRI. This does not mean students actually learned anything in the course.”

“While I personally gain some insight into my teaching after reading the results, data and comments, I feel as though these ratings are a popularity contest. Students that like the professor rate the professor high, and those that do not like the professor, rate the professor low. They do not really think about the teaching style of the professor.”

Theme 6 (N = 21, 9%): The SRI should not be the only component of the evaluation process; it should either not be used at all or weighted less than it currently is in evaluating faculty.

“Currently, this is the only formal structure in place to evaluate faculty in my unit. It should be supplemented with other measures.”

“I do think the SRIs help faculty to better plan for classes. I take more time reviewing the comments to see what I can do to improve instruction, class activities, etc. I think we should use the evaluations in this way, but I do have a problem with departments and colleges putting such emphasis on the scores as tied to tenure and promotion.”

“I don't agree with using the student ratings for tenure and promotion. In my opinion student feedback seems highly correlated to their current grade in the course. For example, if a student is failing, they have negative comments and rate things low. Students who are making an A, rate everything high.”

Theme 7 (N = 9, 4%): The written comments are more useful than the questions with numbered answers for improving teaching effectiveness.

“I do believe that it is important to obtain feedback from our students. However, I personally find the "Comments" section much more useful than the Likert scale questions and wish that there was more encouragement for students to write the comments.”

“I really only pay attention to the students' written comments. We look at them in the department, but we don't find them useful enough to have much weight in the teaching effectiveness decisions---except for the written comments. They are more telling than the scantron.”

“Written comments are probably the most useful part of the process and are used effectively by me and my department.”

“The students' written comments are of considerably greater value than the numerical scores in the various categories. By reading the written comments, one can at least get some insight into what is fueling the numerical scores. Also, with written comments, it is easy to distinguish a thoughtful, constructive criticism (or compliment) from the rantings of an immature student who is disgruntled over having to work hard to obtain a particular grade.”

2. **How should the data collected from Student Ratings of Instruction be used?**

The following six general themes emerged from the surveys (N = 218 respondents). Over one third of respondents stated that SRI data should be used in a formative way for faculty development to improve faculty teaching. Approximately equal numbers stated that SRI data should be used in a summative way to formally evaluate faculty, but nearly one quarter of respondents explicitly stated that SRI data should not be used as the sole or primary measure of evaluating faculty teaching.

Theme 1 (N = 16, 7% of respondents): The SRI is vulnerable to factors that have nothing to do with teaching effectiveness (is too influenced by course difficulty, the professor's standards, the level of the course, and so on). It is a popularity contest.

“I think it should be used but with better insights. There are too many factors involved in teaching evaluation. Seems like the easier the class, the higher the score with our institution's students.”

“Very little at all. The evaluations are too dependent upon the mood of the class and the timing of the evaluation. If the instructor has just given a difficult exam, then the students will give a less favorable evaluation than if the instructor had just bought goodies to the students. Timing is everything, but sometimes the timing just doesn't work.”

“Sparingly since evaluations are really based on likeability and warmth. Good professors, who are challenging, may receive lower evaluations just because of their personalities.”

Theme 2 (N = 77, 35% of respondents): The SRI should be used by the professor to improve his/her teaching and to make adjustments to the course where needed.

“It should be used by the faculty to improve instruction methods in future courses.”

“I think it should be used by the instructor to revise the course for future offerings.”

“For personal use by the instructor, to identify ways to improve a class.”

Theme 3 (N = 9, 4% of respondents): The SRI should allow the department chair to provide feedback to the professor regarding how to improve the course.

“First and foremost, it should be used as feedback to the professor. It should also be used as a vehicle for discussion with the chair of the department and other faculty who may be able to offer advice to a faculty member who is struggling in one or more areas. From an administrator’s viewpoint, it can also help in “matching” faculty to the type of courses they are best suited to teach. The purpose should be to improve teaching and learning – not to inform personnel actions.”

“The data should be used to help faculty get a reading of their teaching. The department head should provide feedback to faculty on the ratings and make suggestions for how to improve ratings. They should carry minimal weight in annual review and P&T.”

“They should be used to allow me, along with my department chair, to explore what appears to have gone well in my class, and what might be improved.”

Theme 4 (N = 52, 24% of respondents): The SRI should not be the only component of the evaluation process.

“The data should be used as a small component of overall faculty evaluation. Peer evaluation of quality of instruction would be a much more informative indicator of quality of instruction.”

“As part of a more comprehensive evaluation that would include a visit by a colleague trained in instructor evaluation (not necessarily in the same college), a self-reflection, and some kind of student learning outcome achievement measure.”

“As one component of evaluating teaching. There should be a review of the syllabus, the assignments, the tests, etc. There should also be a peer review of the class.”

Theme 5 (N = 8, 4% of respondents): The written comments are more useful than the questions with numbered answers for improving teaching effectiveness.

“The “rank 1-5” questions on the SRI instrument should be de-emphasized, perhaps by putting them on the back, to be completed last. For me, the comments are the most helpful in improving my courses.”

“It can be used as a tool for improving teaching but should not be the only instrument used to measure effective teaching. Personally, I will look at the student comments more than the actual numbers.”

“The comments students make are the most useful to me. It helps determine if I need to think about changing anything or not.”

Theme 6 (N = 27, 12% of respondents): The SRI should be used as a measure of faculty performance in teaching, should be part of the annual evaluation process, or should be used in promotion and tenure decisions.

“They should be used to determine whether faculty members are doing a good job in the classroom. Everyone is aware of the problems inherent in student ratings. But along with some issues come some strengths. Students are consumers. They have experience. Although their responses may be shallow sometimes or their priorities are misplaced, they are the best indicators of instruction. By the same token, they should not be the only metric. Peer review of teaching is a nice supplement.”

“For annual evaluations, P&T and post-tenure decisions. It need not be the only data regarding teaching effectiveness considered in these decisions and may not be sufficient in all cases, but it should be included in all cases.”

“Used to evaluate the instructor’s performance in the classroom.”

3. How effective is the current Student Ratings of Instruction instrument in evaluating faculty teaching? Why?

The following six general themes emerged from the surveys (N = 221 respondents). Approximately 25% of respondents considered the existing instrument to be adequate or better, whereas almost half considered it to be ineffective or worse.

Theme 1 (N = 11, 5% of respondents): Very Good or Good.

“I think it works well but has too many questions. I think the comments sections is very valuable for help with course improvement.”

“Studies show they are pretty good. The questions ask about the level of work, whether the student was engaged, and whether the instructor was an effective communicator. That is the information we need to evaluate teaching.”

“Very good. Don’t change it. There are always people that get bad evaluations (usually because they can’t teach) and want to change the evaluation process rather than improve their teaching.

Theme 2 (N = 45, 20% of respondents): Fairly Good / Fine / Adequate / Somewhat Effective / Moderately Effective / or some other similar description.

“I believe the instrument is somewhat effective. Student perceptions and emotions can help the faculty know to some extent how well the course is structured. However, feelings and subjective perceptions are only part of the

story. Most 18-21 year old undergraduates do not have a basic knowledge of pedagogy, and therefore do not fully understand what factors contribute to teaching effectiveness. They may, for example, rate an instructor highly who ends class early most of the time and gives homework assignments on which an A can be easily earned.”

“I think it is reasonably effective over a period of time. One class or one semester of student ratings isn’t very helpful, but several of them over a couple of years tend to show meaningful trends. My experience is that students usually attempt to share constructive criticism on student ratings of instruction.”

“Moderately effective. The numbers provide one indication, but the student comments are much more powerful and should be regarded as important to a faculty member’s success in teaching.”

Theme 3 (N = 95, 43% of respondents): Not Effective / Limited Effectiveness / Minimally Effective / or some other similar description.

“Not effective. In my opinion student feedback seems highly correlated to their current grade in the course. For example, if a student is failing, they have negative comments and rate things low. Students who are making an A rate everything high.”

”Not effective at all. Too many students aren’t evaluating based on valid criteria. They tend to evaluate based on whether they had fun, got their grade they wanted without working hard, had too much homework, etc. Students often read the question in unreasonable ways – perceiving that an instructor is unavailable if the instructor wasn’t in the office when the student dropped by (as opposed to being there for office hours or an appointment).

“Limited effectiveness. Students don’t take these seriously- usually eager to finish and get out of class. And, odd as it may seem, students in the final days of a class may not be the best persons to evaluate teaching – esp while emotions are high b/c of grade, exams, etc. Sometimes it’s only later that students come to appreciate how much they learned in a course. Students are also easily swayed by personalities. They may learn more from a tougher, less personable instructor than one who is an entertainer and less demanding.”

“It may help to show general trends or performance, but overall it is not a very effective measure of successful teaching. It doesn’t measure how much the students have learned at all. It only focuses on whether or not the student are satisfied customers, and that is a very different matter from how much they have been challenged to learn. And the few negative comments I receive from students usually come from students who were irresponsible, often absent, poorly prepared, who don’t even have an accurate view of how the course played out from week to week because they weren’t there.”

Theme 4 (N = 14, 6% of respondents): Completely Ineffective / Very Ineffective / Worthless /or some other similar description.

“Completely ineffective. Most of the time only about 30% of the students in my class will fill out a survey and it has been my experience that those are the ones with an axe to grind. Students who are satisfied with their grade RARELY fill out a review.”

“It’s worthless. It doesn’t evaluate faculty teaching. It investigates students’ idiosyncratic perceptions of faculty members and courses, with all the adherent prejudices, biases, and inaccuracies that come with untrained, unqualified individuals giving their anonymous opinions with impunity.”

“Very ineffective. Students tend to not take them very seriously or try to use them as revenge against a disliked professor.”

Theme 5 (N = 8, 4% of respondents): Poor for on-line courses.

“It is poor, especially for online courses. As mentioned earlier, a professor can make the class easy just to get high SRI. A big problem with online SRI is the return rate. I barely get 50% response rate. These are students who did poorly in my courses so my numbers are skewed. It is not a true measurement of effective teaching.”

“I teach in an all online program. The instrument does not address online courses well. It asks many questions about things over which the instructor has no control in an online course.”

“Elements of the SRI are not as effective when it comes to online teaching and have more applicability with on campus classes. We maybe need an SRI that is specific for the dynamics of online class instruction. The other concern is low return numbers so the results can be skewed to a specific population who maybe didn’t represent either the positive or negative impressions from the learning experience.”

Theme 6 (N = 52, 24% of respondents): Not sure or the effectiveness was not specifically stated in the response, but an opinion related to the SRI was offered.

“I’m not sure how effective they are. I get the sense that it is a popularity contest, rather than a real measure of how much the students learn or are engaged. I teach all upper division classes and most of these courses are ones that students in my program would rather not take. It is evident that students are NOT prepared for the rigor (content and time/project management) of an upper division class. This then leads to lower SRIs. I continually hear student make comments about easy ‘A’s’ in other classes. Because I don’t give easy A’s and expect student to be able to apply their knowledge in my courses (my courses are project based), students struggle, which is then reflected in their attitude and evaluations of my courses.”

“Ratings are often higher for faculty who grade easily and are usually higher in the small classroom setting. As such, they are more a measure of popularity and close interaction with faculty than actual quality of instruction.”

“They may give a sense of how well students like their professors, but that can be popularity, as well as actual engagement. Questions should reflect concrete issues such as whether students are being taught what they need to learn (SLOs), and should take into consideration self-reported student effort. There should also be an accounting of grade distribution so if an instructor gives 90% A’s they have to justify it (perhaps in their reflection or a questionnaire).

“Results vary depending on teaching performance, but also time and day of class, amount of homework or papers assigned or not assigned, and difficulty or easiness of the class. While poor professor are outed by students on the

current instrument, less than glowing results are often due to things out of the professor's control.”

4. A) In your department/unit, what other measures of faculty teaching are required? B) What other measures of faculty teaching are not required, but are available to evaluate faculty teaching?

Combining the responses to parts A and B, the following four general themes emerged from the surveys (N = 215 respondents). By far, the most common other measure of teaching effectiveness (both required and optional) was peer observations of teaching. Some faculty reported that peer observations (one or multiple) were required. Others reported that peer observations were available, but not required. Still others did not even mention peer observations of teaching as an option. This suggests substantial inequities in the ways teaching is evaluated from department to department, particularly in the emphasis placed on SRI data as a measure of teaching effectiveness.

A) other measures of faculty teaching are required (N = 35, 16%):

- Peer Evaluation (N = 22, 12%)
- Syllabus (N = 10, 5%)
- Administration Observation (N = 4, 2%)

B) other measures of faculty teaching that are not required but available to evaluate your teaching

- None (N = 64, 30%)
- Peer Evaluations (N = 61, 28%)
- Observation (N = 8, 4%)
- Syllabus (N = 7, 3%)
- Self-evaluation, narrative (N = 7, 3%)

Other types of faculty teaching evaluations mentioned:

- Student Learning Outcomes
- Administration Observation
- Faculty Development, Workshops
- Tests, Assignments
- Grades & Grade Distribution
- Exit Interviews
- Annual Performance Reviews
- Faculty Workload
- Student accomplishments

Theme 1 (N = 73, 34%): None/I don't know

“I don't know as I have never had any feedback on teaching for the last several years.”

“None, except at pre-tenure, tenure, promotion and post-tenure review. We can certainly make informal arrangements among ourselves, but I don't think people do that very often.”

“None, we depend solely on the data from the student rate as the Holy Grail.”

Theme 2 (N = 47, 22%): Additional faculty member-provided items (e.g., syllabus, assignments, exams)

“syllabi”

“grades and self evaluation”

“Faculty member's Teaching portfolio where syllabi, handouts and overview of work created are evaluated. Evaluation of work produced in the course.”

“We're required to write an annual narrative, part of which describes an instructor's teaching.”

“we submit syllabi and grades every semester”

Theme 3 (N = 67, 31%): Additional non-faculty member-provided items (e.g., peer review, national certification passing rates)

“Peer review of instruction by other faculty. Department Assessments. Department course assessment each semester (other than evaluations, our online assessment tool)”

“We do peer-evaluation, but not very effective. Some of us are evaluated by instructors outside of our programs, and some of us are called upon to evaluate our supervisory administrators, which seems an obvious short-coming.”

“Evaluation of teaching material and text is evaluated each semester by departmental faculty.”

Theme 4 (N = 83, 39%): not required, but available to evaluate faculty teaching

“faculty may seek peer review of courses”

“we have been told that we can have peer observations, but there is no mechanism for doing that. There is no formal procedure to evaluate faculty teaching using any of the items listed in the faculty handbook on which faculty teaching evaluations are supposed to be based.”

“Student preparedness in next level course work. Faculty engagement in curricular committees and other department committees. Creation an private course evaluation specific to the course, by the professor, for students to fill out, put in an envelope that is given it to the secretary to release to faculty member at the start of the next semester.”

“All other measures must be implemented by faculty. No support is offered to assist in this process.”

5. In your department/unit, what weight (percentage or otherwise) are Student Ratings of Instruction given compared to other teaching measures in evaluating faculty teaching during annual reviews?

The following three general themes emerged from the surveys (N = 208 respondents). Two out of every five respondents indicated that they did not know the weight SRIs were given in evaluating faculty teaching during annual reviews. Responses also indicated substantial differences between departments, with SRIs given as little as 10% weight in some departments and as much as 100% weight in others.

Theme 1 (N = 79, 38%): not sure/I don't know:

“ask the chair”

“I am unsure”

“I do not know exactly as I am new faculty”

Theme 2 (N = 47, 23%): Specific numeric/percentage weight (ranged from 10%-100%):

“100%”

“25%”

“In years past, 100% of our annual teaching evaluation has been based on our Item #18 average for the year, without any controls for disciplinary differences, class type or level, class time, student or instructor gender, course rigor, etc. One year, we were compared to the departmental mean. If we were below the mean at all, regardless of whether or not such a difference was statistically significant (I don't think our chair knew how to compute that), we were ‘below expectations’ in teaching. Again, without any controls for any of the factors known to influence SRIs that have no basis in teaching (e.g., course difficulty).”

Theme 3 (N = 55, 26%): Non-numeric weight:

“all the weight is given on the student ratings”

“do not know percentage, but strongly weighted”

“Since they are the only official measure, full weight can be given to them in evaluating faculty.”

“Item 18 is the ONLY thing we use...”

“I don't think there is any exact percentage. It's one measure among many. My impression is that if scores are good, they will help a faculty member (as long as other measures reinforce the faculty member's effectiveness), and if they are not good, that won't necessarily hurt the faculty member as long as there is a legitimate reason and not a repeating pattern of poor pedagogy.”

It should also be noted that a subset of respondents (N = 14, 7%) chose to answer the question instead by stating that too much weight was placed upon SRIs or that they were arbitrary or subjective. Because this did not address the question asked, it was not coded as a primary theme, but is noted here.

“Too much. Evaluations are just impossible in this way. It's like trying to give a teacher a grade for a whole semester by showing up to one class. This is not a good measure of ensuring good teaching. Nor certainly is evaluating personality via students a good way either. The measuring of effectiveness should be focused on the outcomes. Unfortunately sometimes the outcomes are out of a professor's control, but if students from a particular professor are consistently turning out bad work, then one could evaluate the teaching material and review with the professor the classroom approach.”

“There is no specific number but I am supposed to get over a 4 to 5 rating on overall evaluation. Getting a 4 is achievable but chair wants it to go up each year. This is impossible if my average is currently a 4.4. How higher can SRI go? I do not expect to get a 5 on every question but based on annual review from last year, that is what is expected of me.”

6. In your department/unit, what weight (percentage or otherwise) are Student Ratings of Instruction given compared to other teaching measures in evaluating faculty for promotion and/or tenure?

The following three general themes emerged from the surveys (N = 206 respondents). Almost half of the respondents indicated that they did not know the weight SRIs were given in evaluating faculty teaching for promotion and/or tenure.

Theme 1 (N = 96, 47%): not sure/I don't know:

“???”

“For some reason we have to provide all of our student evaluations in our package, but I am not sure how much weight they carry. In my department, being a researcher is much more valued than being a teacher.”

“not sure”

“I don't have the answer to this. Our P & T guidelines are under review and there has been much discussion regarding teaching effectiveness and methods for evaluating.”

Theme 2 (N = 37, 18%): Specific numeric/percentage weight (ranged from 10%-100%):

“100%”

“10-20%”

“Student ratings get 100% because there really is nothing else.”

Theme 3 (N = 43, 21%): Non-numeric weight:

“Student ratings of instruction are the predominant measure for teaching effectiveness for promotion and tenure. I do not believe our department has a specific metric/weight for each as it is more subjective, but every annual review of teaching effectiveness begins with SRI's.”

“item 18 is the ONLY thing we use, at all.”

“no weight”

“SRIs are given heavy weight. Each P&T committee weighs the SRI assessments for each faculty member reviewed so it is difficult to say how much weight the SRI is given each year.”

“Ratings of Instruction are one tool used to evaluate teaching, they are not given more weight than other methods. Each method is used in conjunction with other methods to provide an overview. It would be hard to assign a specific

percentage since we look at faculty portfolios and reviews of instruction and work produced holistically. For some the ratings may hold more value than for others depending on comments repeated from course to course and strength of other methods. As Department Chairs change, this could change since our department chair guides the tenure and promotion process/ evaluation.”

As with Question 5, a subset of respondents (N = 15, 7%) chose to answer the question instead by stating that too much weight was placed upon SRIs or that they were arbitrary or subjective. Because this did not address the question asked, it was not coded as a primary theme, but is noted here.

“Too much. I don't know what the exact percentage is but it's one of a few measures that is weighed in the ‘teaching’ part of teaching, scholarship and service. Student evaluations below 4.0 are frowned upon. With small classes, it is particularly difficult to maintain such a high score, especially when we tell our students not to expect an A. They can say the same to us. The numbers mean something different to the students than they do to us. This gap is a problem.”

“No stated policy. Weight varies. If the faculty like the professor then low scores are rationalized away as indicative of the courses the faculty member teaches, etc. High scores can be used as a measure to show too much emphasis in this area so lower evaluations can be stated in other areas. Bottom line: You need to be liked by your peers and maybe your students!”

7. In your department/unit, what information do faculty receive about how to appropriately interpret a) the statistical summary data provided from their Student Ratings of Instruction? b) anonymous student comments on their Student Ratings of Instruction?

The following three general themes emerged from the surveys (N = 204 respondents). Given that over 80% of respondents indicated that they were given no information on how to interpret either the quantitative or qualitative data from SRIs, it appears that Georgia Southern is in violation of BOR Policy Manual Section 8.3.5.1, paragraph 1, which states, “Each institution, as part of its evaluative procedures, will utilize a written system of faculty evaluations by students, *with the improvement of teaching effectiveness as the main focus of these student evaluations.*” [emphasis added] If faculty are not provided with information about how to interpret SRI data, it is unlikely that the “main focus of these student evaluations” can be on “the improvement of teaching effectiveness,” and suggests instead, consistent with other faculty and administrator responses to these surveys, that the primary focus of SRIs at Georgia Southern is for faculty performance evaluations, in direct contradiction of the express intentions of the BOR.

Theme 1 (N = 167, 82%): None/zero/I don't know, or information indicating no information provided

“none. none.”

“Absolutely none. In fact, I have had to repeatedly argue with past chairs that they were inappropriately interpreting and overinterpreting the summary data and making fallacious comparisons to supposed departmental means. We also receive no information about how to interpret student comments. This lack of information is particularly hard on junior faculty when they face their annual reviews, 3rd year reviews, and P&T reviews, as there are no guidelines on how to respond to the comments.”

“None on either account. It is simply assumed that the higher the quantitative score the better and the more positive the qualitative comments the better. There has never been any discussion at the department or college level regarding interpretation of the results. What do the numbers on the quantitative section mean? Without statistical analysis related to significance, we speak out of ignorance when we talk about a 4.5 being better than a 4.1. How do we know from a statistical standpoint if there is truly any difference between these scores? Likewise, there has never been any conversation regarding data transformation of the qualitative data. People cherry pick the best or worst comments depending on their perspective and agenda. Rather than engaging in analysis, the SRI's encourage overarching generalizations and anecdotal responses.”

“None, except that we are to measure up to a department average, whether we are teaching creative writing, linguistics, first-year composition, or technical writing.”

“all stats and comments (anonymously)”

Theme 2 (N = 13, 6%): Additional information is provided, but nothing on how to interpret

“a) We receive frequency plots for each question. b) I have not received students comments”

“Mostly up to faculty - A department average is communicated to determine where a faculty member stands in relation to peers.”

“None. We simply receive our averages and the departmental averages for the course.”

“We get two numbers each semester (course average and instructor average) and we are told what the departmental mean and standard deviation are. This informs us about whether we are in the norm, statistically significantly above or below the norm. Again, this comparison is faulty given that it groups together all courses across 6 academic areas in the department and all courses from Intro courses to Senior Seminars.”

Theme 3 (N = 21, 10%): Some information on how to interpret provided

“An email is sent from the secretary about the summary data. Verbal instruction about student comments has come from the department chair.”

“basic definitions of means”

“Chair's evaluation, and that is about it.”

8. Student Ratings of Instruction can sometimes be administered electronically instead of face-to-face with a paper version. If you could choose, which method of administration would you pick and why?

The following three general themes emerged from the surveys (N = 211 respondents). Slightly more faculty respondents preferred a face-to-face paper version than an online version. Proponents of a face-to-face version thought that student response rates would be significantly higher, significantly more representative of the diversity of responses in the course, and that a monitored in-class environment would reduce collusion in responses. Additionally, respondents expressed concern that electronic/online versions would allow students who had stopped

attending class to complete evaluations, would be vulnerable to fraud (e.g., someone other than the student completing the evaluation), would have to employ highly coercive tactics such as withholding final grades until they are completed in order to get students to complete them (which would render the results meaningless), and would make it easier for students to “flame” instructors (e.g., writing highly critical, cruel, or even criminal threats because of the perception of greater anonymity online). Proponents of online versions thought that online administration would be “greener” (using less physical resources) and more efficient (because support staff would not need to scan/type student responses). Additionally, many proponents of online versions mentioned that with online versions, they would not need to use instructional time to administer SRIs.

Theme 1 (N = 93, 44%): Face-to-face paper version:

“F2F. The literature is replete with examples of massive problems with on-line evaluations, most notably the pathetic response rate (often single digits). Only through coercive measures can institutions get the majority of their students to take on-line valuations, and those measures corrupt any data that might be collected.”

“Definitely, the best approach is face-to-face. My experience at other institution, where this data was collected electronically, indicates that the amount of collected data decreases considerably. Much more participation is attained by collecting the data face-to-face.”

“Electronic means ‘optional.’ We already have examples of ‘hybrid’ class of well over 100 students who do this online. Maybe 10% of the students actually participate. Face-to-face classes (i.e., paper forms) at least get a response from a useful percentage of students. 10% is useless information, even for someone like me who actually reads and responds to the data.”

“F@F gets a better response rate. Electronic is worthless - tends to only get responses from angry/disgruntles/failing students”

“paper. I teach in an only online program. unless you require the student to complete the evaluation, they will not do it. Also, if you require them to do it they almost ALWAYS either choose right down the middle, or lower-- it does not appear they choose the best scores. If they feel required to complete it, students have told me that they feel that isn't right and therefore complete it when they are angry -- since they don't know their final grades.”

Theme 2 (N = 76, 36%): Electronically/online:

“Electronic re: cost & eliminating the hassle of doing them in class. I just don't think we'll get much student response if we go electronic unless there is some kind of sanction imposed for students NOT doing them.”

“electronically”

“electronically is better, but the computers are not available in each classroom. One of the reason's for electronic is to have things easily readable and accessible rather than the paper format after the evaluation is completed.”

“I prefer electronic, but have no solution to the low response rate issue.”

Theme 3 (N = 23, 11%): No preference/unsure/don't know/either/it depends:

“Does not matter to me. It is likely that fewer students would use an electronic version unless it was done in class.”

“I do not have a preference”

“It depends. If they are electronic, but optional, then you will only get feedback from the very few students who are either very happy or very unhappy, and this is not a representative sample at all. I also fear that online evaluations, taken in isolation, will have that anonymous feeling that leads students to have less integrity. (i.e. anonymous online trolling)”

“undecided”

9. Which items on the current Student Ratings of Instruction instrument are the most valuable to you as a teacher and why?

The following four general themes emerged from the surveys (N = 206 respondents). Many faculty respondents explicitly mentioned student learning in their responses and their desire to use SRI feedback to improve their teaching effectiveness for the purpose of increasing student learning, consistent with BOR Policy Manual Section 8.3.5.1, paragraph 1’s stated intentions for how SRI data should be used.

Theme 1 (N = 42, 20%): Student/Course items (#1-7)

“How much effort did you put in, workload, and to what degree were you intellectually challenged questions help me keep a balance between content difficulty and student ability.”

“I look at how much they learned, how much effort they put into their class and how much they were challenged. I personally think when these values are high in combination that I have done my job and the students took something away from this course that ay stick around for a lifetime. The appreciation for the knowledge often comes much later in life.”

“Questions 2 and 3 are important to me as I want to ensure my students are gaining conceptual knowledge (they can apply what they learn) from my classes as opposed to on the surface learning (simply memorizing definitions and formulas). I think it is important to stimulate and challenge the way students think so they hopefully gain better understanding of the sciences. Question 5 is somewhat important as I do not want to make my courses too easy nor too difficult. However, it is also a difficult question to gauge as students (mostly freshmen) tend to find everything difficult.”

Theme 2 (N = 98, 48%): Instruction/Instructor items (#8-18)

“1. Instructor's preparation. Anyone who has ever had me knows I prepare extensively for courses and it is obvious every day. Students who are out to ‘get’ me will often ding me on all items, including this one, but no reasonable person would conclude that a low score on this item for me is valid. It's an easy was to demonstrate student animus. 2. Tests/graded activities. Typically, students who do poorly in the course rate me very lowly on this item to deflect blame for their own failures. Again, great way to demonstrate student animus. 3. Instructor's interest. Same as #1.”

“Did the instructor provide enough information. Did the class content meet the expected course objectives.”

“I use #9 (preparation), #12 (clarity), #15 (helpfulness), and #18 (overall) as they directly relate to things I can do and how I interact with the students.”

Theme 3 (N = 34, 17%): Items #19-22 (pre/post interest, required, major)

“How much the students rate their level of interest before the class and after. This really helps me see how much I have affected their world view.”

“I value the questions that ask students how interested in the material they were before and after the course, and how much they learned”

“Level of interest in the subject before and after the course, how much students learned”

Theme 4 (N = 71, 34%): Comments

“Comments!!!! -- because that actually means something to me. A 4.3 in organization does not.”

“I find the comments most valuable especially the thoughtful comments concerning class and assignments. I am not concerned if a student writes, the class is too hard. This is college, it should be hard but doable.”

“I read the comments, and if there is a theme, I take it into consideration. However, students say outrageous things, such as, I never got a syllabus.”

10. Which items on the current Student Ratings of Instruction instrument are the least valuable to you as a teacher and why?

The following five general themes emerged from the surveys (N = 182 respondents). A consistent remark from faculty respondents was that items which did not help them improve their teaching were not valuable, again suggesting that faculty view the value of SRI data in terms of improving teaching effectiveness and enhancing student learning, consistent with Policy Manual Section 8.3.5.1, paragraph 1’s stated intentions for how SRI data should be used.

Theme 1 (N = 45, 25%): Student/Course items (#1-7)

“#1, 5 and 6. I expect that student invest their time in studying and learning instead complaining about the load of reading and assignments”

“Questions that pertain to workload and intellectual challenges. Students don't have the post-graduate experiences to understand how much work is necessary to succeed in a competitive environment. The same with intellectual challenges; the academic stress in high schools is typically not the same level as a university. Students don't understand that by definition, Professors —profess, which is different by definition than teaching. I think students too often place the responsibility of learning on faculty not themselves.”

“Any question that begins with "Compared to other courses of similar credit value" is not helpful. This often comes down to comparing apples with oranges. You are asking students to compare online course to F2F courses to required courses to elective courses to courses in their major to courses outside their major to upper level courses to core courses.”

“Anything that is not question 8-18 because we are not evaluated on that.”

Theme 2 (N = 43, 24%): Instruction/Instructor items (#8-18)

“#13 - students do not understand that application and synthesis of the course material is relevant on an exam and are just looking for rote memorization questions”

“Instructor availability, helpfulness: these are items that are based on each individual students' experiences. If students never come for help, they have no idea if you're available or if you'd be of help if they did come by. In my large classes in particular, it is very rare that any students ever come by, so how do they know how to answer these questions at the end of the semester?”

“Overall how would you rate this professor? This is basically a "how do you like this professor" question. I often find that the rating for this question is completely contrary to the ratings for all the other questions. I score very high on the questions about being prepared, knowledgeable, available to students, tests reflecting the material in class, etc. and then score very low on this question. This is the popularity question and it should be removed.”

Theme 3 (N = 26, 14%): Items #19-22 (pre/post interest, required, major)

“Required course, and other non-instructional assessment items -- these provide little information about instruction, but could be useful for understanding why some ratings may be lower than others (required course, workload and difficulty level, etc.).”

“The item like if this course is requirement.”

“what was your level of interest in course before and after?”

Theme 4 (N = 22, 12%): #23/Expected Grade

“question 23”

“questions 23 - what grade do you expect to get. I don't know what this is really going to tell me or how this is going to help me improve my teaching”

“The "What grade do you expect in this course" question.”

Theme 5 (N = 10, 5%): All/Most of the items

“Most of them. Students are not qualified to objectively judge most of the items and their responses to some of them indicate that they are not interpreting the questions in the way they are written. For example, the instructor availability question is one many faculty consistently get dinged on, not because we are not available, but because students expect 24/7 access to us and no matter how many times we tell them that such expectations are unrealistic and need to change, they insist on it and punish us for not being willing or able to drop everything any time of day or night to respond to them. Similarly, the helpfulness item--helpfulness does not mean telling them what they want to hear or just ‘giving’ them easy answers. Helpfulness means teaching them, helping them to help themselves, but most students just resent the effort we make them put forward. Also, #16 course objectives. Students are just not qualified to judge this. If they could determine what is and what is not relevant to the course objectives, they could

teach the class. Finally, the open-ended questions, though well-intentioned, are often worthless. They typically yield pabulum, whether positive or negative they are almost never objective or concrete, which makes it impossible to understand them or respond to them. For example, students may claim an instructor is 'bias' (note the misspelling), but do not provide a single specific example of what they are interpreting as 'bias.' Is it a reliance on facts? Is it a refusal to allow prejudicial comments towards others in the class? How can we know?"

"Most of it is not particularly valuable."

"The entire instrument"

11. If you could add/delete/change items/questions on the current Student Ratings of Instruction instrument, what would you do and why?

The following five general themes emerged from the surveys (N = 164 respondents). No predominant pattern emerged.

Theme 1 (N = 40, 24% of respondents): no changes, N/A, "I don't know," or leave it as is:

"None"

"N/A"

"They are, I think, a necessary evil in this age of accountability."

Theme 2 (N = 13, 8% of respondents): scrap the entire measure and start over from scratch:

"I would revisit the entire instrument and examine each question carefully, determine what it is determine, and how it is used. For example, the current Summary of Means supplied in the evaluation for promotion/tenure/post-tenure review is absolute statistical nonsense and totally meaningless."

"I would scrap the existing instrument and start over. Given administrative misinterpretation of SRI data, I would strongly suggest the new form have no quantitative questions at all (that way, there is nothing for administrators to misinterpret)."

"Redo the instrument completely. Decide if the SRI is for evaluation or as a tool for improvement. Then design an appropriate instrument."

Theme 3 (N = 8, 5% of respondents): add items to assess what *students* did to be successful in the course (e.g., study hours, reading assigned texts, etc.):

"I think students should have to indicate how many hours they study for the course in the average week. I think students should have to indicate how many classes they have missed over the course of the semester. I think students should have to answer open-ended questions about what THEY did to learn the material and be successful in the course. I think they should have to report on how much of the assigned reading they did. "

“I would like to add questions related to student involvement ... eg how often did you attend, did you meet with the instructor if you were having problems, did you read the text etc.”

Theme 4 (N = 7, 4% of respondents): more/more specific open-ended questions:

“Add open ended questions asking students to describe ways in which the instructor was: organized/prepared, interesting and thought provoking, challenging, effective at presenting course material, innovative in teaching, effective in make class interest in subject, etc. I would take away the scores all together.”

“I would make the open ended questions more specific, break up course/instructor.”

Theme 5 (N =6, 4% of respondents): add questions that focus on student learning/SLOs/assessment:

“Questions should be added to evaluate student centered teaching. Students should be asked to evaluate both in class and out of class activities in terms of benefit on learning. Current instrument focuses too much on instructor as deliver of information and not enough on instructor as facilitator of learning.”

“Eliminate questions 8-18. Revise the evaluation to be an evaluation of student learning; then, use it as a tool to inform the Assessment Committee about how instruction is meeting SLOs.”

12. If you could change anything about how the current Student Ratings of Instruction instrument is used in evaluating faculty teaching, what would you change and why?

The following three general themes emerged from the surveys (N = 175 respondents). The largest group of respondents recommended reducing or eliminating the use of SRIs in evaluating faculty teaching.

Theme 1 (N = 26, 15% of respondents): no changes, N/A, “I don’t know,” or leave it as is:

“Currently I am okay with the form.”

“THAT SHOULD NOT THE BUSINESS OF ANY UNIVERSITY COMMITTEE. Every department (and profession) has the right to define it.”

“None.”

Theme 2 (N = 46, 26% of respondents): reduce or eliminate their use in evaluating faculty teaching:

“Although I know BOR regulations require SRIs, I would be in favor of eliminating them completely. They are worse than worthless--they stand in the way of instructors implementing rigor, discipline, professional expectations for students' behavior, and innovative teaching methods that might result in greater student learning gains at the cost of student ‘satisfaction.’ Students are not customers and we should not treat them as such. Their opinions about faculty teaching are largely worthless when it comes to actually evaluating that teaching and we should not lie to them and pretend otherwise. I would change to a form that focuses more on students' reflection on their own behaviors and learning over the semester and one that would require students to provide specific, concrete examples for any comment about the course or the instructor to be considered valid. . . I would also create university-wide

guidelines that would force all chairs, deans, and other administrators to stop misinterpreting the data and set clear standards for how to interpret and/or compare SRI data and would make violations of these guidelines cause for faculty grievances. . . The data from current SRIs is worthless and the way it is currently used by administrators is harmful and inappropriate. It prevents great teaching and ultimately harms student learning. With all of our focus on assessment lately, I would think that the need to ACCURATELY assess faculty teaching would be important and would be cause enough to get rid of these meaningless forms.”

“I place considerably less emphasis on it at college and university levels. I know administrators like to quantify but the numbers on evaluations are not necessarily an accurate measure of good teaching.”

“I would use them to give feedback to instructors; I would not weigh them heavily in evaluating teaching.”

Theme 3 (N = 8, 5% of respondents): SRIs should be administered electronically

“Do it electronically”

“If you are going to waste time asking students, Put it online.”

“Go online”

13. What are your thoughts about how often the Student Ratings of Instruction instrument should be administered (e.g., in each course each semester, in only one semester/academic year, etc.)?

The following three general themes emerged from the surveys (N = 195 respondents), with an overwhelming majority of respondents endorsing the status quo.

Theme 1 (N = 126, 65%): Status quo, each class, each semester:

“end of semester per each class”

“every semester”

Theme 2 (N = 37, 19%): Less often, such as once/course/year or one semester/year:

“Every semester in every course is too much. I would support once per year in each course.”

“At many other institutions, faculty have the option of selecting one course each semester to NOT be evaluated. This is particularly important if the faculty member is trying a new technique or strategy that might meet with student resistance. I would be in favor of allowing a faculty member to select one course each semester to be free from evaluation, as long as it isn't the same course every semester.”

Theme 3 (N = 13, 7%): More often, such as twice per course/semester or more:

“If we had Student Ratings of Student Learning rather than of ‘Instructor’ instruction, I would be in favor of administering them, electronically, at the beginning, the middle, and the end of the semester. Even, perhaps, every

three weeks, so that students can evaluate their own progress. Teachers could then use the data to look at patterns and search for ways to improve instruction to fill the gaps.”

“At the very minimum they should be administered at the mid-term and end of course. However, I think instructors should take it upon themselves to seek more student evaluations and input.”

14. If you have any additional thoughts or comments about this subject, please type them below.

The following three general themes emerged from the surveys (N = 95 respondents). Respondents primarily re-iterated prior responses in this question.

Theme 1 (N =16, 17%%): N/A, none, no response:

“N/A”

“None.”

Theme 2 (N = 13, 14%): Eliminate or replace the existing measure:

“Change is good...Try another instrument-----“

“I recommend you move beyond the instrument and look at implementing additional approaches to measuring teaching effectiveness.”

Theme 3 (N = 8, 8%): Don’t change the existing SRI:

“I do not see a need to change the current form.”

“I think that the current evaluations can be properly used and the problem is there are no guidelines for their use.”

Section IV: SRI Administrator Results

General Summary

Many chairs indicated dissatisfaction with the current SRI and its use in evaluating teaching, often echoing the concerns raised by faculty and the literature presented in the Background section. One clear pattern that emerged from the data was that chairs appeared to be just as frustrated with the current forms and their use as faculty and were very supportive of finding better and more appropriate ways to evaluate faculty teaching effectiveness.

Also note that because the sample size for this survey was less than one tenth of that of the faculty survey, there were significantly fewer total responses. This resulted in the identification of fewer themes for responses each question, but many of those themes reflected sentiments expressed by a significant number of the respondents (e.g., 20-75%). As a result, the data presentation in this section is slightly different.

Demographic Data (N = 21)

Participants' College (N = 19, 2 unreported)

COBA: 4 of 5

COE: 1 of 3

CEIT: 2 of 5

CHHS: 3 of 3 (100%)

CLASS: 4 of 11

COSM: 4 of 5 (80%)

JPHCOPH: 1 of 5

Library: N/A

Questions, Result Summaries, & Exemplar Quotations

1. Generally speaking, what are your thoughts about the current Student Ratings of Instruction instrument and how it is used at Georgia Southern University?

Overall, most responding chairs (N = 15, 71%) identified multiple problems with the current SRI instrument and its use in evaluating faculty teaching. A subset of respondents (N = 5, 24%) indicated some utility to the current SRI form.

Example comments:

- “Constructive feedback is rarely offered.”
- “often inappropriately averaged across items and often comparisons made between course averages inappropriately, as courses and students differ.”

- “It is generally misused by administrators as they go higher up the administration chain of command. They look at overall averages and often a single measure.”
- “I really feel that the SRI should be considered ONE part of several things used to support teaching”
- “I am wary of giving too much weight to such ratings in T&P and other decisions.”
- “I believe that it is only marginally effective and too much weight is assigned to it for faculty evaluations.”

2. How should the data collected from Student Ratings of Instruction be used?

Overall, 38% (N = 8) of responding chairs suggested that SRI data should be used in formative ways by faculty members to improve their own teaching. Multiple respondents (N = 4, 19%) suggested that SRI data either not be used at all for evaluation/T&P, or used only minimally and with great caution.

Example comments:

- “As one of several perspectives on the ability of an instructor to effectively deliver instruction. It also has value in assessing the enthusiasm that the instructor brings to the classroom .whether there is sufficient communication between students and th instructor, and student perceptions of the fairness of testing and weights assigned to different facets of the course.”
- “With caution for annual review purposes, but I believe are being used less cautiously and have largely impacts on annual reviews.”
- “Mainly to determine patterns across faculty for faculty improvement.”
- “For faculty development and course improvement”
- “Primarily for faculty feedback and discussions with chairs, rather than for comparison to other faculty.”
- “Chairs should use it to evaluate all faculty relevant to departmental norms.”
- “The data collected should be used exclusively by the instructor to improve the classroom experience. Use of ratings for purposes of Instructor effectiveness or for promotion and tenure decisions is not appropriate.”

3. How effective is the current Student Ratings of Instruction instrument in evaluating faculty teaching? Why?

Overall, responding chairs (N = 14, 67%) identified many specific problems in using the current SRI instrument to evaluate faculty teaching that compromise its effectiveness as an evaluative tool and may even lead to erroneous conclusions about faculty teaching.

Example comments:

- “Fair to poor. It only reflects one point of view. Students tend to blame professors for their own poor performance. I find that students rate professors the highest who give the least assignments, lecture to the test, offer study guides to the test and give extra credit.. I consider those to be my poorest performing professors.”
- “The instrument was designed and implemented when traditional delivery methods (e.g. lecture) was the norm. For those that continue to do this, it is okay. It comes up short for evaluating online delivery, flipped classes, hybrid courses, and other pedagogical advances.”
- “VERY limited. It is a single dimension of measure – ‘student perceptions of teaching’.”
- “Personally, I do not think that it is very effective (maybe 6 or 7 on a scale of 10, with 10 being the best). SRI should be ONE part of the overall evaluation of teaching.”

- “Somewhat effective. It does not measure learning outcomes of students. Too much emphasis is placed on the nuances of class organization and structure and not enough on content.”
- “its useful if evaluated appropriately. It reflects only one point of view and very often that point of view is how well the student can navigate the course with the least amount of effort for a good grade. Usually it does not reflect learning outcomes. I have found that student expectations centers around the faculty telling the student what he/she needs to know, providing a study guide that re-inforces that information and then testing only what is on that study guide. This only reflects short term memory and is not the type of teaching that I would encourage - but whenever the student is made responsible for participating in his/her education the evaluations become very low. I have found that some of the best teaching is done by professors receiving the lowest ratings and some of my poorer professors receive the highest ratings (not always though - this is variable- the student comments have to be taken into consideration - they provide very useful information as to why they rated a professor the way they did)”
- “I have seen students make zig-zag patterns on the scantrons. No one in their right mind should think that the ratings are pure measures of teaching effectiveness.”
- “Very poor. There is not universal student mindset on which we can depend. Some students give serious and valuable feedback and others do not. This lack of consistency makes the ratings nearly useless for evaluating faculty.”

4. **A) In your department/unit, what other measures of faculty teaching are required? B) What other measures of faculty teaching are not required, but are available to evaluate faculty teaching?**

In most departments (N = 13, 62%), SRIs were clearly used as the primary (or only) evaluative measure of faculty teaching not produced by the faculty member (e.g., syllabi, summary of curricular revisions, etc.). This was particularly true for annual reviews as compared to T&P, pre-tenure, etc., where peer reviews were more likely to be required. In other departments (N = 5, 24%), SRIs were only a small part of evaluations of faculty teaching. Such inconsistencies suggest substantial inequities in the ways SRIs are used to evaluate faculty teaching from department to department.

Example comments:

- “No other "measures" but materials (digital and hardcopy) and syllabi.”
- “Syllabi are evaluated; test and assignment content and results and peer observation is used sporadically. In courses where student learning outcomes are assessed, assessment data patterns across time and adjustments made in response to previous assessment results are not overlooked.”
- “Nothing else is required...however peer evaluations are strongly encouraged...as well as participating in teaching workshops either on or off campus.”
- “Faculty are evaluated annually by their peers. These become a permanent part of the faculty member's personnel file.”
- “None.”
- “Our department has a peer evaluation instrument. It is required of faculty going through any type of review (pre-tenure, tenure, promotion, etc.). It is optional to any faculty at any time. We have a peer review committee elected each year who conducts the reviews.”

5. In your department/unit, what weight (percentage or otherwise) are Student Ratings of Instruction given compared to other teaching measures in evaluating faculty teaching during annual reviews?

Some units had clearly assigned percentage weights (N = 6, 29%); other units were unable to specify the specific influence of SRIs on annual evaluations of teaching (N = 13, 62%). Again, this suggests substantial inequities in the ways SRIs are used to evaluate teaching from department to department.

Example comments:

- “About 10%”
- “25%”
- “30%, Peer reviews 40%, Self-assessment 30%”
- “They carry about 50%”
- “We have no weights. It is all used subjectively by the chair. A weighted system for something so subjective would be useless.”
- “We do not assign a percentage (to my knowledge). We compare the ratings to the department mean for undergraduate lower, undergraduate higher, and graduate courses. The APR document suggests that tenure-track teaching faculty assign a minimum of 50% of ‘workload’ to teaching and are asked to provide multiple indicators of teaching effectiveness.”
- “No specific numerical weighting or percentage, which is the way it should be. The chair and peer review committees see the numbers and reach a holistic judgment of teaching based on them and many other considerations.”
- “There isn't even a place in the evaluation tool in which to list the SRI.”
- “Unfortunately these ratings are still used as the primary measure of teacher quality.”

6. In your department/unit, what weight (percentage or otherwise) are Student Ratings of Instruction given compared to other teaching measures in evaluating faculty for promotion and/or tenure?

Responses to this question varied widely, from nearly 100% to “barely looked at.” Some units (N =4, 19%) had clearly assigned percentage weights; other units (N = 15, 71%) were unable to specify the specific influence of SRIs on T&P. Again, this suggests substantial inequities in the ways SRIs are used to evaluate teaching from department to department.

Example comments:

- “25%”
- “30%, Peer reviews 40%, Self-assessment 30%”
- “No specific percentage but typically more than it should -- probably 60% because its an easy to access summary which is prone to inappropriate comparisons because it is identical across courses, yet not equally useful across courses.”
- “They used to carry almost 100% but during the past two years with change in P&T guideline, other measures/documents are required.”
- “They are typically rolled-up by year so they are given a strong weight and must be included with materials to the Provost.”
- “I don't have a percentage.”

- “SRI numbers are barely looked at since we don't have a good tool nor are the numbers submitted in a consistent manner.”

7. As an administrator, what education/training have you had about how to appropriately interpret a) the statistical summary data provided from Student Ratings of Instruction? b) anonymous student comments on Student Ratings of Instruction? Please describe.

Responses to this question were the most similar of any question in the survey. Without exception, *none* of the responding chairs reporting having received any specific education/training *from Georgia Southern* about how to use GSU's SRI instrument for evaluating faculty. Many chairs (N = 6, 29%) referenced their graduate education, especially in statistics and/or methods courses, which would provide a basic understanding of the issues involved, but no specific understanding of the unique issues and confounding variables known to affect SRIs. A number of chairs (N = 4, 19%) referenced their own experience with receiving and/or reviewing SRIs in reference to this question. Only four chairs (19%) indicated any specific training or education in interpreting SRIs (e.g., workshops in their fields, reading the SRI literature, training at other institutions). These responses suggest that many chairs may have an overconfidence in their ability to identify known issues in interpreting SRIs, and potential “blindness” to the need for specific training on how to interpret (and avoid misinterpreting and over interpreting) SRIs. These responses further suggest that Georgia Southern is in violation of BOR Policy Manual Section 8.3.5.1, paragraph 2, which states, “Institutions will ensure that the individuals responsible for conducting performance evaluations are appropriately trained to carry out such evaluations (BoR Minutes, 1979-80, p. 50; 1983-84, p. 36; May, 1996, p. 52).”

Example comments:

- “none”
- “None formally from the university.”
- “I am confident in my interpretive skills for survey data results.”
- “No formal training. Only experience over time.”
- “No formal training. Only years of experience doing my best to interpret the data.”
- “I have never received any education/training specific to this form or dataset”
- “None. Just basic understanding of statistics and common sense when reading student comments.”
- “My graduate work and research is quantitative, social science based, and have constructed measures or scales many times for my published research. I apply that understanding to the question of how to interpret or understand SRIs. I do not recall having received any training specifically on SRIs and their use as an administrator, except for a session in a conference of department heads in my field.”
- “a) none at GSU; b) none at GSU. Received great deal training at other institutions”

8. Student Ratings of Instruction can sometimes be administered electronically instead of face-to-face with a paper version. If you could choose, which method of administration would you pick and why?

Responses to this question were evenly split between respondents. Half of respondents favored electronic administration, primarily because of the cost of administering and entering data from paper versions. Half of the respondents favored paper versions, primarily because of documented issues with online versions (response rate, validity, etc.).

Example comments:

- “F2F due to greater participation and likely more valid responding.”
- “Electronically for convenience”
- “Face to face... Better represent all students and better comments. Online evaluations have sampling bias unless ALL students are required to complete them”
- “Electronically please. What are we waiting for?”
- “Paper version. My former institution wen to the electronic version and the rate of completion dropped about 60%. We need evaluations to assist junior faculty and it was extremely hard to get students to do the online version.”
- “face to face - it the only way we can get students to fill out the form. Online evaluations are usually only filled out by those who either feel very positive or negative about the professor.”
- “Electronically”

9. Which items on the current Student Ratings of Instruction instrument are the most valuable to you as someone who evaluates faculty and why?

Responses to this question varied significantly among respondents. Many (N =10, 48%) focused explicitly on item #18 (instructor rating), either alone or in conjunction with other items. Some chairs (N = 6, 29%) mentioned items from the “course” section (#1-7), whereas others (N = 9, 43%) focused on items from the “instructor” section (#8-18). Again, this suggests a significant lack of consistency in how SRI data is used across departments.

Example comments:

- “Questions 5,6, & 18”
- “I zero in on items 1-3 to get a better understanding of student engagement. Item 5 provides an acceptable gage of course difficulty, especially when considered alongside item 3. I always consider item 6 to get a sense of student perceptions of how easy he course is and item 7 to get a sense of student perceptions of the overall value of the course. I pay close attention to student responses to item 10.”
- “3 and 5 help gauge rigor of the course, 9,11,12,14,15,18 demonstrate diligence of the instructor, 21 & 22 indicate degree of affinity for the course”
- “I look first at the summary ratings of the course and of the instructor. These are most important in my examination of the SRIs. Then I examine the individual items, to see if any are particularly high or low, compared to those of the department. But the written responses are the second-most helpful because students will write comments that indicate strengths or weaknesses that are not captured well by the scantron items.”
- “The two main items would be the overall rating of the instructor and the expected grade in the course as it give an idea of the student's motivation at the time of filling out the form.”

10. Which items on the current Student Ratings of Instruction instrument are the least valuable to you as someone who evaluates faculty and why?

Responses to this question demonstrated a similar pattern to student responses to SRIs: many items identified in the previous question as “most valuable” were identified here by other individuals as “least valuable.” Six respondents

(29%) identified the contextual information about the course (#1-7) as least valuable, whereas only three (14%) identified instructor questions (#8-18) as least valuable.

Example comments:

- “I don't pay much attention to items 8 and 9 because I don't think that students are positioned to know if important points are sufficiently stressed or how prepared instructors are. I don't pay much attention to item 4, unless there are red flags for items 1-3, 5, & 6. I don't pay much attention to item 10, unless there are red flags for items 1-3, 5, 6, 11 & 12.”
- “Effort in learning material (#1) How often did you seek outside help (#4), How was the workload (#6) level of interest (#19&20) - too subjective and not really rated to the learning objectives”
- “23 is anecdotally unreliable. I have seen students inflate their expectations on this item time and again.”
- “Questions 1-8 do not provide useful information about the instructor and how well he/she conducted the class. Comments about the course itself do not help me evaluate the faculty member.”

11. If you could add/delete/change items/questions on the current Student Ratings of Instruction instrument, what would you do and why?

Responses to this question also varied widely: some chairs suggested no changes (N = 5, 24%), others suggested changes, but didn't provide any specifics (N = 4, 19%), still others recommended specific additions, changes, or deletions (N = 8, 38%). One theme that emerged from a subset of respondents who made specific suggestions was a focus on student learning.

Example comments:

- “I would recommend we ask no more than 10 questions and provide more space for open responses. For example asking students to give an example of how the faculty member showed that they were concerned that the students were learning the necessary content”
- “I have no idea. I would just not use it as the ultimate assessment of a faculty members teaching effectiveness”
- “I would want to know well the course prepared the student for clinical application of the content; the degree the course encouraged critical analysis of the content; the degree to which the professor used problem solving techniques instead of lecture; the degree to which the student improved writing skills, analytical skills, problem-solving skills; the degree to which the professor interacted with the student in a professional manner; the degree to which the professor held the student accountable for his/her learning”

12. If you could change anything about how the current Student Ratings of Instruction instrument is used in evaluating faculty teaching, what would you change and why?

Responses to this question suggested that many respondents (N = 5, 24%) favored changes to the way the SRI forms are currently used, specifically a decrease in the focus and importance given to them in evaluating teaching.

Example comments:

- “I'd like to see less emphasis on the ‘Score’”
- “Train chairs, deans, provost etc. to understand evaluation of faculty teaching is multidimensional and "student perceptions" of teaching effectiveness is one data point among many.”

- “Increasing the objective data over the subjective data”
- “I am wary of the SRI summary accompanying the person's CV as virtually the only things that go forward beyond the college in T&P decisions. Whether or not administrators place too much weight on SRIs, the fact that this is all that goes forward gives the impression that they do.”
- “I would prefer to evaluate how a faculty member assesses their student evaluations and what they plan to do to address any weaknesses. The faculty member should be in the best position to interpret specific comments and then as an administrator, you can better address the planned efforts to improve and optimize the faculty's teaching.”
- “Would require its use only with other assessment measures - not as a stand alone assessment.”

13. What are your thoughts about how often the Student Ratings of Instruction instrument should be administered (e.g., in each course each semester, in only one semester/academic year, etc.)?

Responses to this question yielded a modal pattern endorsing the status quo (N = 13, 62%): every course, every semester. A subset of respondents suggested more or less often, with the primary pattern for those who suggested more often being including a new mid-semester course evaluation.

Example comments:

- “Every course, every semester”
- “every course-maybe even midterm and then at the end”
- “Randomly administered to select classes - maybe 25% of all class in a department per semester (with no repeats if in the first round); however, tenure-track faculty should have a minimum number of evaluations prior to submitting the portfolio for review, say 18 evaluations across the various classes”
- “For repeat courses once a year. Other courses once each semester is adequate”

14. If you have any additional thoughts or comments about this subject, please type them below.

Responses to this “catch all” question yielded a wide array of diverse suggestions. Some respondents used it as space to reiterate prior opinions; others used it to provide “summary comments” about SRIs.

Example comments:

- “Students' opinions of evaluations are a good thing; however, students may not be the individuals who are experts in teaching. Some look for easy instructors, classes that may give them high grades, or entertaining professors. They base evaluations on many preconceived notions that have nothing to do with the content they must learn in order to be competitive in the job market. Evaluations come at the end of the semester. In order to help faculty, they should be given in the beginning of the semester, maybe called early assessment of the course, and they provide little help to faculty to improve the course for that group of students. Many students do not even read the evaluations. Students who have poor attendance and poor grades, take their frustrations out on the professors. So what are we measuring - how well students like the professors or at what level did they learn the content?”
- “We absolutely should not move to a one size fits all numerical cut off for personnel, merit, or any other kinds of decisions affecting the lives of faculty. Professional peers in the discipline are the best judge of quality teaching. Students are not, because they will often negatively score rigor, the necessity of assigned reading, and the

necessity of oral class participation. Faculty know better than students what constitutes good teaching, and any system seeking the support of faculty must reflect this reality.”

- “I would hate to see Georgia Southern go to all online evaluations. I think this would be a HUGE mistake and weaken the evaluation system and make it harder for department chairs to evaluate faculty. I also would hate to see Georgia Southern go to fewer valuations than every class every semester. This is the only way to have a complete picture of a faculty member's teaching.”
- “Many of the courses in the sciences (and I am sure in other disciplines as well) are very hierarchical. As such, the latter courses require the knowledge of previous courses as a foundation to the upper level courses. I have often had discussions with students who express the thought that at the time of the class, they did not like the difficulty in which the course was taught and assessed but now appreciate the difficulty in that they are better prepared than some of their other classmates for the upper level courses. Therefore, it could be useful to find a way follow up on initial course evaluations after the course has been completed.”

Section V: Recommendations

Because of the myriad problems with the current SRI measure used by Georgia Southern University and its use, both as identified in the Background section and as identified by faculty and department chairs, the committee makes the following recommendations:

1. That the Faculty Welfare Committee be charged with:
 - a. composing a new SRI that incorporates best practices from the research literature and focuses on student learning, learning behaviors, and formative feedback (e.g., the Skowronek et al., 2011 measure, see Appendix); the new SRI should provide multiple opportunities for students to specify in writing how the instructor promoted student learning;
 - b. pilot testing the new SRI form in classes from every college and of various sizes and levels;
 - c. making final revisions to the new SRI measure based on the results of the pilot testing and presenting the new measure to the Faculty Senate for adoption; and
 - d. proposing methods to make the evaluation of teaching effectiveness more equitable and consistently defined, assessed, and used across the university. This would include developing guidelines for how SRIs should be used and objectively valued in annual reviews and in promotion and tenure (and pre/post tenure) decisions for all faculty.
2. That the comparison of individual faculty SRI data to other faculty members (e.g., department means) or to a specific “cut point” (e.g., 4.0) be immediately prohibited and that the faculty handbook be amended to reflect this change. Such comparisons are contrary to “best practices” in the use of SRIs (University of Wisconsin-La Crosse, 2007), are not statistically valid (University of Washington, 2005), and are based upon the erroneous presupposition that the current SRI form is not vulnerable to biasing influences from variables unrelated to teaching effectiveness, including biases against protected classes. Further, in a university that emphasizes commitment to excellence in teaching and learning as a hiring criterion, it should be expected that the large majority of faculty are already good teachers.
3. That the use of SRI data as either the sole or majority criterion for evaluating teaching effectiveness be immediately prohibited and that the faculty handbook be amended to reflect this change until such time as the Faculty Welfare Committee can develop more specific guidelines for the use of SRI data. The existing SRI forms do not assess teaching effectiveness in any meaningful way, and weighting them so heavily runs contrary to best practices in their use (University of South Dakota, n.d.; University of Washington, 2005). Additionally, the Georgia Southern University Faculty Handbook (205.01) already lists multiple other methods for assessing teaching effectiveness: “examination of course syllabi and other course materials, peer evaluations when available, critical review and dissemination of teaching products, performance of students in subsequent venues, follow-up of graduates in graduate school or in their employment”.
4. That faculty members should be given an opportunity to respond to SRI results. “Faculty should have an opportunity to discuss the objectives of the course, how the teaching methods were used to meet that objective,

and how circumstances in the course might have affected” SRIs (University of Washington, 2005). These responses should be permanently appended to any future reports of that SRI data.

5. That Georgia Southern University immediately discontinue the practice of forwarding a one-page summary of the SRIs to the Provost’s Office along with major reviews (e.g., promotion and tenure, post-tenure) and amend the faculty handbook to reflect this change. Such out-of-context summaries can have neither use nor purpose unless the data is being inappropriately compared to other faculty or to “cut points”, both of which violate best practices for the use of SRI data (University of Wisconsin-La Crosse, 2007), and are not statistically valid (University of Washington, 2005).
6. That SRI administration match the method of delivery for the course: online courses should use online course evaluations; face-to-face courses should use face-to-face evaluations. The existing literature documents extremely low participation rates for online course evaluations in face-to-face courses which can only be ameliorated by the implementation of costly, logistically complicated, and draconian measures to coerce student compliance (the use of which would entirely negate any value of such evaluations on improving teaching effectiveness).

Section VI: References

- Abrami, P. C., d'Apollonia, S., & Cohen, P. A. (1990). Validity of student ratings of instruction: What we know and what we do not. *Journal of Educational Psychology, 82*, 219–31.
- Abrami, P. C., d'Apollonia, S., & Rosenfield, S. (1997). The dimensionality of student ratings of instruction: What we know and what we do not. In R. P. Perry and J. C. Smart (Eds.), *Effective teaching in higher education: Research and practice* (pp. 321-67). New York, NY: Agathon Press.
- Addison, W. E., Best, J., & Warrington, J. D. (2006). Students' perceptions of course difficulty and their ratings of the instructor. *College Student Journal, 40*, 409-416.
- American Educational Research Association. (2013). *Rethinking faculty evaluation: AERA report and recommendations on evaluating education research, scholarship, and teaching in postsecondary education*. Washington, DC: American Educational Research Association. Retrieved from:
<http://www.aera.net/Newsroom/NewsReleasesandStatements/AERAIssuesRecommendationsforRethinkingFacultyEvaluationforthe21stCentury/tabid/15254/Default.aspx>
- Aleamoni, L.M. (1999). Student rating myths versus research facts from 1924 to 1988. *Journal of Personnel Evaluation in Education, 13*, 153-166.
- Algozzine, B., Beattie, J., Bray, M., Flowers, C., Gretes, J., Howley, L., Mohanty, G., & Spooner, F. (2004). Student evaluation of college teaching: A practice in search of principles. *College Teaching, 52*(4), 134-141.
- Angelo, T. A. (1993). "Teacher's dozen": Fourteen general, research-based principles for improving higher learning in our classrooms. *AAHE Bulletin, 45*, 3-13.

- Carrell, S. E., & West, J. E. (2010). Does professor quality matter? Evidence from random assignment of students to professors. *Journal of Political Economy*, 118, 409–432.
- Cashin, W. E. (1995). *Student ratings of teaching: The research revisited (IDEA Paper No. 32)*. Manhattan, KS: Kansas State University, Center for Faculty Evaluation and Development.
- Cashin, W. E. (1990). *Student ratings of teaching: Recommendations for use (IDEA Paper No. 22)*. Manhattan, KS: Kansas State University, Center for Faculty Evaluation and Development.
- Cashin, W. E. (1988). *Student ratings of teaching: A summary of the research (IDEA Paper No. 20)*. Manhattan, KS: Kansas State University, Center for Faculty Evaluation and Development.
- Centra, J. A., & Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching? *Journal of Higher Education*, 71, 17-33.
- Cramer, K. M., & Alexitch, L. R. (2000). Student evaluations of college professors: Identifying sources of bias. *The Canadian Journal of Higher Education*, 30, 143-164.
- d'Apollonia, S., & Abrami, P. C. (1997). Scaling the ivory tower, Part I: Collecting evidence of instructor effectiveness. *Psychology Teaching Review*, 6, 46–59.
- Feldman, K. A. (1993). College students' views of male and female college professors: Part II— Evidence from students' evaluations of their classroom teachers. *Research in Higher Education*, 34, 151–211.
- Franklin, J. (2001). Interpreting the numbers: Using a narrative to help others read student evaluations of your teaching accurately. In K. Lewis (Ed.), *Techniques and strategies for interpreting student evaluation* (pp. 85-100). San Francisco, CA: Jossey-Bass.
- Georgia Southern University. (n.d.). *National Survey of Student Engagement (NSSE)*. Retrieved from <http://em.georgiasouthern.edu/osra/nsse/>
- Gibbs, G. (1995). *Assessing student-centered courses*. UK: Oxford Brookes University.

- Greenwald, A. G., & Gillmore, G. M. (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist*, 52, 1209–17.
- Halpern, D. F. (1999). Teaching for critical thinking: Helping college students develop the skills and dispositions of a critical thinker. *New Directions for Teaching and Learning*, 80, 69-74.
- Hamermesh, D. S. & Parker, A. M. (2005). Beauty in the classroom: Instructors' pulchritude and putative pedagogical productivity. *Economics of Education Review*, 24 369-376.
- Healey, M., & Jenkins, A. (2003). Educational development through the disciplines. In R. Macdonald & H. Eggins (Eds.), *The scholarship of academic development* (pp. 47–57). Buckingham, UK: Open University Press.
- Heckert, T. M., Latier, A., Ringwald, A., & Silvey, B. (2006). Relation of course, instructor and student characteristics to dimensions of student ratings of teaching effectiveness. *College Student Journal*, 40, 195-204.
- Hendrix, K. G. (1998). Student perceptions of the influence of race on professor credibility. *Journal of Black Studies*, 28, 738-764.
- Hoyt, D. P., & Pallett, W. H. (1999). *Appraising teaching effectiveness: Beyond student ratings (IDEA Paper No. 36)*. Manhattan, KS: Kansas State University, IDEA Center.
- Huemer, M. (2005). *Student evaluations: A critical review*. Retrieved from http://home.sprynet.com/~owl1/sef.htm#_14
- Hutchings, P., Huber, M. T., & Ciccone, A. (2011). *The scholarship of teaching and learning reconsidered: Institutional integration and impact* (Jossey-Bass Higher and Adult Education Series). San Francisco, CA: Jossey-Bass.
- Johnson, V. E. 2003. *Grade inflation: A crisis in college education*. New York, NY: Springer-Verlag.
- Kember, D., Leung, D. Y. P., & Kwan, K. P. (2002). Does the use of student feedback questionnaires improve the overall quality of teaching? *Assessment and Evaluation in Higher Education*, 27, 411–25.
- Kolitch, E., & Dean, A. V. (1999). Student ratings of instruction in the USA: Hidden assumptions and missing conceptions about 'good' teaching. *Studies in Higher Education*, 24, 27–42.

- Kulik, J. A. (2001). Student ratings: Validity, utility, and controversy. In M. Theall, P. C. Abrami, & L. A. Mets (Eds.), *The student ratings debate: Are they valid? How can we best use them?* (pp. 9-25). San Francisco, CA: Jossey-Bass.
- Lowman, J. (1995) *Mastering the techniques of teaching* (2nd ed.). San Francisco: Jossey-Bass.
- Marsh, H. W., & Hocevar, D. (1991). Students' evaluations of teaching effectiveness: The stability of mean ratings of the same teachers over a 13-year period. *Teaching and Teacher Education, 7*, 303–41.
- Marsh, H. W., & Roche, L. A. (1997). Making students' evaluation of teaching effectiveness effective: The critical issue of validity, bias, and utility. *American Psychologist, 52*, 1187-1197.
- McKeachie, W. (1997). Student ratings; the validity of use. *American Psychologist, 52*, 1218-1225.
- McKeachie, W.J., & Hofer, B.K. (2001). *McKeachie's teaching tips: Strategies, research, and theory for college and university teachers* (11th ed.). Boston, MA: Houghton Mifflin.
- Naftulin, D. H., Ware, Jr., J. E., & Donnelly, F. A. (1973). The Doctor Fox lecture: A paradigm of educational seduction. *Journal of Medical Education, 48*, 630-635
- Nerger, J.L., Viney, W., & Riedel, R.G. II (1997). Student ratings of teaching effectiveness: Use and misuse. *The Midwest Quarterly, 38*, 218-233.
- Olivares, O. J. (2004). Student evaluations of teachers: Intended and unintended social consequences. *Journal on Excellence in College Teaching, 15*, 105–27.
- Perry, W. G. (1970). *Intellectual and ethical development in the college years: A scheme*. New York, NY: Holt, Rinehart, & Winston.
- Powell, R. (1978). Faculty rating scale validity: The selling of a myth. *College English, 39*, 616–29.
- Pratt, D. D. (1997). Reconceptualizing the evaluation of teaching in higher education. *Higher Education, 34*, 23-44.

- Rubin, D. L. (1998). Help! My professor (or doctor or boss) doesn't talk English. In J. N. Martin, T. K. Nakayama, L. A. Flores (Eds.), *Readings in Cultural Contexts* (pp. 149 – 160). Mountain View, CA: Mayfield Publishing Company.
- Ryan, J. J., Anderson, J. A., & Birchler, A. B. (1980). Student evaluations: The faculty responds. *Research in Higher Education, 12*, 317–33.
- Schlenker, D., & McKinnon, N. (1994). Assessing faculty performance using the student evaluation of instruction. Atlantic Baptist College. (ERIC Document reproduction Service No. HE027508).
- Seldin, P. (2006). *Evaluating faculty performance: A practical guide to assessing teaching, research, and service*. Boston, MA: Anker Publishing Company, Inc.
- Skowronek, J. S., Friesen, B. K., & Masonjones, H. (2011). Developing a statistically valid AND practically useful student evaluation instrument. *International Journal for the Scholarship of Teaching and Learning, 5*(1), Article 11. Retrieved from: <http://digitalcommons.georgiasouthern.edu/ij-sotl/vol5/iss1/11>
- Singer, S. R., Nielsen, N. R., & Schweingruber, H. A. (Eds.) (2012). *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. Washington, DC: National Academies Press.
- Titus, J. (2008). Student ratings in a consumerist academy: leveraging pedagogical control and authority. *Sociological Perspectives, 51*, 397-422.
- U.S. Department of Education. (2011). *Title 34: Education* (34 CFR 600.2). Washington, DC: U.S. Government Printing Office. Retrieved from: <http://www.gpo.gov/fdsys/granule/CFR-2011-title34-vol3/CFR-2011-title34-vol3-sec600-2>
- UC Regents. (2000). *Assessment of student learning gains*. Retrieved from: http://mc2.cchem.berkeley.edu/Evaluation/class_ev.html

University of South Dakota. (n.d.). *Frequently asked questions: The IDEA evaluation process*. Retrieved from:

<https://www.usd.edu/continuing-and-distance-education/upload/IDEA-FAQ.pdf>

University of Washington. (2005). *Recommendations for the use of student ratings of instructors in merit and promotion decisions*. Office of Educational Assessment. Retrieved from:

<http://www.washington.edu/oea/resources/recommendations.html>

University of Wisconsin—La Crosse. (2007, October). *Student evaluation of instruction: Background information and guidelines*. Retrieved from:

<http://www.uwlax.edu/facultysenate/Archived%20Documents/SEI%20Background%20Info%20&%20Gdlns.htm>


Voss, R., & Gruber, T. (2006). The desired teaching qualities of lecturers in higher education: a means end analysis.

Quality Assurance in Education, 14, 217- 242.

Wachtel, H. K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment and Evaluation in Higher Education, 23*, 191–212.

Weinberg, B. A., Hashimoto, M., & Fleisher, B. M. (2009). Evaluating teaching in higher education. *Journal of Economic Education, 40*, 227–261.

Appendix: Updated Skowronek et al. (2011) Measure

Class Climate	The University of Tampa Classroom Survey	
		

Mark as shown: Please use a ball-point pen or a thin felt tip. This form will be processed automatically.
 Correction: Please follow the examples shown on the left hand side to help optimize the reading results.

WHY YOU SHOULD COMPLETE THIS EVALUATION

The university is dedicated to continuously improving classroom instruction. As a way of furthering this mission, we value your input regarding your direct experience in this course. **Your responses are part of the overall faculty evaluation process and can help both the university and your professor better understand your classroom experience and the impact it has on your learning.**

INSTRUCTIONS

Please read the instructions at the beginning of each section carefully. **All your responses will be kept anonymous. Faculty will not see your responses until AFTER final grades have been submitted.**
Thank you for completing this survey!

1. THE COURSE

Indicate below how each aspect of the course **impacted your learning**, ranging from "Did not help at all" to "Helped a great deal" or by indicating the level present for that aspect. If you are unable to evaluate a particular aspect in anyway, please choose "NA" ("Not Applicable").

- 1.1 The class assignments/ projects/ activities were:

Easy

Difficult

NA



1.2 The class assignments/ projects/
activities:

1.3 Comments on class activities:

Did not help at all

Helped a
great deal

NA



1.4 Class discussions occurred:
Frequently

Never
NA

1.5 The class discussions:

Did not help at all

Helped a great deal NA

1.6 Comments on class discussions:

1.7 The exams/ quizzes/ tests were:
NA

Easy

Difficult

1.8 The exams/ quizzes/ tests:

Did not help at all

Helped a great deal NA

1.9 Comments on exams/ quizzes/ tests:

1.10 The way this course was organized: Did not help at all

Helped a great deal NA

1.11 Comments on course organization:

1.12 The pace of this course was:
NA

Slow

Fast

1.13 The pace at which this course progressed:

Did not help at all

Helped a
great deal

NA

1. THE COURSE [Continue]

1.14 Comments on course pace:

1.15 Overall, the course:

Did not help at all	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Helped a great deal NA

- 1.16 I know more about this subject now than I did before I took this course.
- 1.17 My skills in this area have improved as a result of taking this course.
- 1.18 My appreciation of this subject increased as a result of taking this course.
- 1.19 The learning objectives of the course were met.

Strongly Disagree Strongly Disagree Strongly Disagree

Strongly
Disagree

Strongly	NA Agree							
Strongly	NA Agree					<input type="checkbox"/>		<input type="checkbox"/>
Strongly	NA Agree					<input type="checkbox"/>		<input type="checkbox"/>
Strongly	NA Agree				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2. THE PROFESSOR

Indicate below how each aspect of the course **impacted your learning**, ranging from "Did not help at all" to "Helped a great deal" or by indicating the level present for that aspect. If you are unable to evaluate a particular aspect in anyway, please choose "NA" ("Not Applicable").

		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.1	The professor's presentations/ explanations:						

Did not help at all

Helped a great deal NA

2.2 Comments on presentations/ explanations:

2.3 The professor's enthusiasm for the subject was:

- Low High NA
- 2.4 The professor's level of enthusiasm for the subject:
- 2.5 Comments on enthusiasm:



Did not help at all

Helped a
great deal

NA

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
				<input type="checkbox"/>

2.6 The professor stimulated my interest in the subject:

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

Never

Frequently

NA

2.7 The level at which the professor stimulated my interest in the subject:

2.8 Comments on interest stimulated:



Did not help at all

Helped a
great deal

NA



2.9 The professor's interactions with me:

Did not help at all



Helped a great deal

NA



2.10 Comments on interactions:



2.11 The professor provided feedback on my work:



Never Frequently NA
2.12 The professor's feedback on my work:

Did not help at all

Helped a
great deal

NA



Clas

2. THE PROFESSOR [Continue]

2.13 Comments on feedback:



2.14 The professor challenged me to learn:



Never

Frequently

NA

2.15 The level at which this professor challenged me to learn:

2.16 Comments on challenge:



Did not help at all

Helped a
great deal

NA

2.17 Overall, the professor:

Did not help at all

Helped a
great deal

NA



3. Additional comments

3.1 What aspect(s) of your **classroom experience (course, professor, etc.)** helped your learning most?

3.2 What aspect(s) of your **classroom experience (course, professor, etc.)** could have been changed to help your

learning?

4. THE STUDENT

The information in this section is important for the purposes of improving teaching and grouping responses into similar classifications. **Your responses below will NOT impact the validity of your responses** in the previous sections.

Please answer each statement honestly.

4.1 Are you taking this course as part of your major/
minor?

4.2 How many class meetings did you miss in this course?

4.3 Approximately how many hours per week did you spend on this course outside of the classroom?

4.4 How often did you seek the professor's assistance?

4.5 Based on the professor's expectations, how often were you fully prepared for class?

No Yes Not Sure

None 1-2 3-4

None 1-3 4-6
7-9 10-12 13 or more

Never Rarely Periodically

Frequently Always

4.6 I believe my final grade in this course will be:

A AB
B BC C
C D D F

S U
P

4.7 I am:

Female Male Prefer not to answer

4.8 I consider myself to be:

Asian Black
Caucasian/White Hispanic/Latino
Multi-Ethnic Native American Pacific Islander
Prefer not to answer

4.9 My current status at UT is:

Freshman Sophomore Junior
Senior Graduate Student