

Final Report

**Georgia Southern University
ad hoc Committee
on Student Ratings of Instruction¹**

**Nan LoBue (CLASS)
Trent Maurer (CHHS, Chair)
Eudiah Ochieng (SGA)
Cordelia Zinskie (COE)**

**Submitted
February 17th, 2017**

¹ The Committee recognizes Delena Gatch (COSM) and Errol-Anthony Spence-Sutherland (SGA) for their work on the Committee Spring semester 2015-Spring semester 2016.

Table of Contents

Section I: Introduction, Timeline, and Overview	2
Section II: Background	5
Section III: Design of the Pilot Student Ratings of Instruction Instrument.....	13
Section IV: Methodology of the Pilot Study and Feedback Survey of Faculty and Chairs	16
Section V: Results of the Pilot Study.....	19
Section VI: Results of the Feedback Survey of Faculty and Department Chairs.....	46
Section VII: Recommendations	54
Section VIII: References.....	59
Appendix A: Pilot Student Ratings of Instruction Instrument	64
Appendix B: Faculty Questionnaire.....	67
Appendix C: Department Chair Questionnaire.....	68
Appendix D: Recommended Student Ratings of Instruction Instrument	69

Section I: Introduction, Timeline, and Overview

Introduction: Committee Charge and Composition

In the Spring of 2015, the Senate Executive Committee of the Faculty Senate at Georgia Southern University created an ad hoc Committee on Student Ratings of Instruction [SRI] to identify an instrument to replace the current Student Rating of Instruction form in use at Georgia Southern University. The Committee was charged to develop or choose for purchase an instrument that

- i. would function primarily as a formative assessment to inform and improve teaching effectiveness (per BOR policy), and
- ii. was consistent with the recommendations of the 2014 ad hoc Committee on SRIs, namely
 - a. that it incorporate best practices from the research literature on SRIs,
 - b. that it focus on student learning, learning behaviors, and formative feedback, and
 - c. that it provide opportunities for students to specify in writing how the instructor promoted student learning.

Additionally, the Committee was charged to

- i. pilot test the new SRI instrument in classes from every college and of various sizes and levels,
- ii. make final revisions to the new SRI instrument based on the results of the pilot testing,
- iii. present the new instrument to the Faculty Senate for adoption, and
- iv. propose methods to make the evaluation of teaching effectiveness more equitable and consistently defined, assessed, and used across the university including 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.

The ad hoc Committee was initially composed of four faculty representatives nominated by the Senate Executive Committee and one student representative from the Student Government Association [SGA]. Committee members included²

- i. one former department chair,
- ii. one expert in questionnaire construction and survey methodology,
- iii. one expert in academic assessment of student learning,
- iv. two experts on the Scholarship of Teaching & Learning [SoTL], one of whom had specific expertise in the research on Student Ratings of Instruction,
- v. two members of the 2014 ad hoc Committee on SRIs, and
- vi. at least one teaching award winner at each of the departmental, college, university, and university system levels.

Timeline

In the Fall 2015 semester, the ad hoc Committee explored two options for SRIs for purchase, ultimately rejecting both for failure to meet the guidelines set in the Committee's charge. Instead, the Committee developed a measure of their own following those guidelines for pilot

² In the Fall of 2016, one faculty member of the committee resigned because of a change in position, and the student representative was replaced (because of turnover in SGA).

testing in Spring 2016. The development of this measure followed best practices in questionnaire design (Artino, La Rochelle, Dezee, & Gehlbach, 2014), including 1. Conducting a thorough literature review on research on SRIs and best practices, 2. Synthesizing the literature review with other data and considerations (e.g., BOR policy, logistical requirements, the results of the 2014 ad hoc Committee on SRI's findings from faculty and chair questionnaires, etc.), 3. Developing items (DeVellis, 2016; Fowler, 2013; McCoach, Gable, & Madura, 2013), 4. Assessing response process validity (e.g., asking students how they interpret the items and response options), and 5. Pilot testing the measure.

In early Spring 2016, the Committee assessed response process validity by alpha testing the pilot measure in the courses of two of its members at midterm, and querying students in those courses about the pilot measure; this test suggested students were interpreting the items as intended and revealed no major problems with the instrument. Subsequently, the Committee was informed that the software necessary to process the pilot SRI they had developed was not in use in every department on campus but had been approved for purchase by the Provost's Office. However, the Committee was informed that training on how to use the software for the relevant departmental administrative assistants would not be completed in time to run the pilot test in Spring 2016 and that pilot testing would need to be postponed until Fall 2016. As a result, the Committee postponed the pilot testing of the proposed SRI measure in face-to-face courses until Fall 2016, but proceeded with pilot testing the measure in online courses in Spring 2016.

During the first week after the Spring 2016 semester had concluded, the Committee sent the individual results for courses that had been selected for the online pilot test to the affected faculty and department chairs. Additionally, those selected faculty and department chairs were invited to participate in two online Qualtrics questionnaires (one for faculty, one for chairs) requesting their feedback about the utility of the data in improving teaching and learning and any other comments they might have about the proposed SRI form.

In the Fall 2016 semester, the Committee pilot tested the measure in face-to-face courses. Individual departments were responsible for conducting and processing the pilot SRI forms for selected courses in their departments and returning the data to the Committee and the selected faculty by the first day of the Spring 2017 semester (January 9th). On the second day of the Spring 2017 semester, all faculty and department chairs for which the Committee had received usable data were invited to participate in two online Qualtrics questionnaires (one for faculty, one for chairs) requesting both their feedback about the utility of the data in improving teaching and learning and any other comments they might have about the proposed SRI form.

By January 13th, 2017, the Committee had completed the initial data analysis for the pilot test. By January 27th, 2017, the Committee had completed the data analysis for the faculty and department chair feedback questionnaires and submitted an official request for an agenda item to the Faculty Senate so that the Committee could present its draft report and solicit feedback from the Faculty Senate. On January 30th, 2017, the Committee met to finalize the Draft Report in preparation for the February 7th, 2017 Faculty Senate meeting. The Committee submitted the Draft Report to the Faculty Senate Moderator for distribution to the Faculty Senate on January 31st, 2017. On February 3rd, 2017, the Committee was informed that the Faculty Senate Executive Committee had declined to put the Committee on the agenda for the February 7th, 2017 Faculty Senate meeting and was instead going to recommend to the Faculty Senate that the Committee be dissolved before being allowed to complete its charge and make recommendations. The Committee objected to this decision and requested the Faculty Senate Executive Committee reconsider allowing the Committee to present its report to the Faculty

Senate. The Senate Executive Committee agreed to allow the Committee to complete its charge and write and present its Final Report at the March 6th, 2017 Faculty Senate meeting. On February 8th, 2017, the Committee met to make its recommendations and write the Final Report. On February 17th, 2017, the Committee submitted an official request for an agenda item to the Faculty Senate for the March, 6th, 2017 meeting so that the Committee could present its Final Report and recommendations. The Final Report was attached to the agenda request and was also submitted to the Faculty Senate Librarian for inclusion in the Librarian's Report so that the work of the Committee could be officially documented in Faculty Senate records. The Committee also submitted three motion requests related to its recommendations for consideration at the March 6th, 2017 meeting.

Overview

The remaining document is organized into six sections: 1. Background information from the literature on SRIs and best practices, which serves to contextualize the work of the Committee, 2. Information on the design of the proposed SRI instrument, 3. A description of the methodology of the pilot testing of the proposed SRI instrument and survey of faculty and chair perspectives about the proposed SRI instrument, 4. The results of the pilot testing of the proposed SRI instrument, 5. The results of the survey of faculty and chair perspectives of the pilot instrument, and 6. The recommendations of the ad hoc Committee.

Section II: Background³

“I no longer think [student evaluations] should be used in any formal way by any institution, especially not as a measure of teaching quality and especially not for the purposes of hiring, merit evaluations, firing, tenure, et cetera. They do not measure what they purport to measure.”

Philip B. Stark (Associate Dean and Professor of Statistics at UC-Berkeley),
as quoted in Flaherty (2015)

“We don’t measure teaching effectiveness. We measure what students say, and pretend it’s the same thing. We calculate statistics, report numbers, and call it a day.”

Stark and Freishtat (2014, p. 9)

“Learning results from what the student does and thinks and only from what the student does and thinks. The teacher can advance learning only by influencing what the student does to learn.”

Nobel Laureate Herbert A. Simon,
as cited in Ambrose, Bridges, DiPietro, Lovett, & Norman (2010, p. 1)

“[Y]ou count only as you add to a sum into which you disappear without a trace.”

Platt (1993, p. 2),
critiquing the message that “traditional” Student Ratings of Instruction
send to students about the value of their feedback

The Student Ratings of Instruction instrument currently in use at Georgia Southern University is considered a “traditional” Student Evaluation of Teaching [SET]⁴ instrument in that it is similar to those that have been in use at many universities for several decades: 1. There are multiple quantitative items on a common ordinal response scale (often Likert-type), 2. Those quantitative items are often separated into two or more distinct areas of evaluation thought to reflect different underlying factors associated with teaching effectiveness (e.g., the course and the instructor), and 3. Because of assumptions associated with the first two items on this list, it is common practice to take the arithmetic mean of either all the items on the entire SET, or a subset of items (e.g., all of the items on the “instructor” section) and treat that single number as a meaningful indicator of teaching effectiveness. Unfortunately, as has been thoroughly and repeatedly documented in vast literature on SETs, there are multiple significant problems with “traditional” SETs like the kind currently in use at Georgia Southern University.

Problems with “Traditional” Student Evaluations of Teaching

The content of “traditional” Student Evaluations of Teaching [SETs] has been criticized in the research literature for many reasons. The two most important criticisms share a common

³ Some of the text presented in this document was previously presented in the 2014 *Final Report of the Georgia Southern University Ad Hoc Committee on Student Ratings of Instruction*, Statesboro, GA: Georgia Southern University. Retrieved from <http://academics.georgiasouthern.edu/facultysenate/files/Final-SRI-report.pdf>. It is reproduced here with permission from the author.

⁴ “Student Evaluations of Teaching” is the preferred term in the research literature for what Georgia Southern University calls “Student Ratings of Instruction.” To ensure the use of a single, consistent term, which is often directly quoted, throughout this section we use the term Student Evaluations of Teaching or SETs.

theme: SETs do not accurately assess what they are supposed to assess. In other words, SET scores are poor indicators of *student learning* and poor measures of *teaching quality*.

Relationship between SET scores and student learning. Because the end goal of teaching is ultimately student learning, “the best criterion of effective teaching is student learning” (Cashin, 1988, p. 2; cf. Cohen, 1981). “Traditional” SETs have poor correlations between ratings and student learning: “Whatever it is the students are responding to, it’s certainly not what they’re learning,” (Philip B. Stark, as quoted in Pettit, 2016, para. 4). In fact, there is evidence of an *inverse* correlation between student learning and SET scores. Students taught by instructors with higher SET scores have 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 (Braga, Paccagnella, & Pellizzari, 2014; Carrell & West, 2010). As noted by Uttl, White, and Gonzalez (2016, p. 19),

Despite more than 75 years of sustained effort, there is presently no evidence supporting the widespread belief that students learn more from professors who receive higher SET ratings. If anything, the latest large sample studies show that students who were taught by highly rated professors in prerequisites perform more poorly in follow up courses (Weinberg, Hashimoto, & Fleisher, 2009; Yunker & Yunker, 2003).

It is important to examine that finding within the context of federally mandated standards for collegiate level work and students’ self-reported study behaviors, because “teaching effectiveness depends not just on what the teacher does, but rather on what the student does” (McKeachie & Hofer, 2001, p. 6). The United States Department of Education Office of Postsecondary Education (2011) defined 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. (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 2015 National Survey of Student Engagement [NSSE] (Georgia Southern University, n.d.) revealed that only 6% of first-year students and 8% of seniors reported spending that much time out of class on academic work. The modal student was spending less than one third of the required time on out-of-class academic work (i.e., 6–10 hours). Although such numbers are nearly identical to the national average within NSSE, other data has suggested that nationally the average student spends between 12 and 15 hours per week on academic work (Arum & Roksa, 2011; Pascarella, Blaich, Martin, & Hanson, 2011). Additional NSSE data revealed that 82% of both first-year students and seniors at Georgia Southern University reported at least sometimes attending class without having completed readings or assignments.

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. Course grades are one of the strongest predictors of SET scores, accounting for nearly 10% of the variance (Wright & Jenkins-Guarnieri, 2012). Further, the least academically competent students (i.e., the ones most likely to earn low grades) are actually more likely to make external attributions for academic failure (i.e., blame the instructor for “poor teaching”) instead of identifying how their own behaviors (e.g., study time, attendance, etc.) could have contributed to their grades (Kruger & Dunning, 1999). Yet, at best, SET scores are unrelated to actual learning after controlling for grades (Weinberg et al., 2009), and at worst SET scores are actually negatively related to learning outcomes (Braga et al., 2014; Carrell & West, 2010).

As Braga et al. (2014, p. 85) explained, “good teachers are those who require their students to exert effort; students dislike it, especially the least able ones, and their evaluations reflect the utility they enjoyed from the course” instead of what they actually learned. Davidson and Price (2009) argued,

[C]ollege students today adopt an increasingly consumerist goal for their education (Baker and Copp 1997; Delucchi and Smith 1997a, 1997b; Shepperd 1997; Smith 2000). Students perceive themselves primarily as customers (Higher Education Research Institute 2000), purchasing a product: a degree. . . In a consumerist environment, though, student evaluations are not ‘good’ data. They measure how easy the instructor is, how fun, and sometimes, as in the case of the Rate My Professor website, how sexy he or she is. Such data should not be used by students or organisations to evaluate an instructor’s ability to teach. (pp. 61–62)

In this way, “traditional” SETs may actually *discourage* good teaching. Additionally, Hattie’s (2009) meta-analysis of over 800 meta-analyses revealed that teacher variables account for only about 30% of the variance in student learning, but student variables account for almost 50%. “Traditional” SETs ask few if any questions about student variables related to learning, leaving both faculty and evaluators of faculty blind to significant potential influences on student learning.

Relationship between SET scores and teaching quality. “Traditional” SETs also do not help institutions define “teaching quality” based on student learning objectives [SLOs], do not help faculty members improve their teaching by identifying areas in need of professional development, and do not help evaluators determine a faculty member’s strengths and weaknesses as a teacher (American Educational Research Association [AERA], 2013). In fact, as AERA noted, “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 “traditional” SETs as a measure of teaching quality is questionable at best (Abrami, d’Apollonia, & Rosenfield, 1997; Kulik, 2001; Wachtel, 1998), and “There is strong evidence that student responses to questions of ‘effectiveness’ do not measure teaching effectiveness.” (Stark & Freishtat, 2014, p. 2). Further, 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) the relevance of course materials (Seldin, 2006).

Additionally, as Titus (2008) noted,

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. (pp. 401–402)

There is also a large and diverse literature documenting multiple variables that have nothing to do with teaching effectiveness that exert biasing influences on “traditional” SET scores. The most well-known example is instructor gender. Boring, Ottoboni, and Stark (2016) reported multiple problems: a) traditional SETs are biased against female instructors by an amount that is both statistically significant and large, b) it is not possible to adjust for this bias statistically, because it depends on so many factors, c) traditional SETs are more sensitive to students' gender bias and grade expectations than they are to *teaching effectiveness*, and d) these gender biases can be large enough to cause more effective female instructors to get lower SETs than less effective male instructors.

In their study of 654 Chief Academic Officers [CAOs] at American higher education institutions, Jaschik and Lederman (2017) reported that 70% of CAOs at public doctoral institutions were either “somewhat” or “very” concerned that SETs may not be accurate measures of teaching quality, and 83% of those CAOs were either “somewhat” or “very” concerned specifically about bias in SETs against female or minority faculty members. Yet, so entrenched are “traditional” SETs as a method for evaluating faculty that only 55% of CAOs at public doctoral institutions reported that their institutions were reconsidering how they used SETs to evaluate faculty, even with those serious concerns.

Problems with Using SETs Inappropriately in Evaluating Teaching

The last section documented the failure of “traditional” SETs to accurately reflect either student learning or teaching quality. There are also multiple issues with SETs being used in inappropriate ways to make judgements about faculty teaching. Penny (2003) documented that many administrators who use SETs to evaluate faculty teaching are insufficiently aware of the vast research literature on SETs, are not sufficiently trained to know how to appropriately interpret such data, and may not even be aware of their own ignorance about these issues.

In this section, we outline the three most important issues with respect to using SETs inappropriately to evaluate faculty teaching: a) using SETs as a replacement for meaningful evaluation of faculty teaching, b) using quantitative data from SETs in statistically unjustifiable ways, and c) failing to interpret qualitative data from SETs in context.

Using SETs as a replacement for meaningful evaluation of faculty teaching. As Stark and Freishtat (2014, p. 9) have observed, “We don’t measure teaching effectiveness. We measure what students say, and pretend it’s the same thing. We calculate statistics, report numbers, and call it a day.” Although this approach is quick, cheap, and easy, it is not an *evaluation* of faculty teaching and does not tell us anything meaningful about faculty teaching. This is echoed by Bain (2004, p. 171), who cautioned against using scores from SETs as *de facto* evaluations of faculty, stating, “If we ask students the right question, their answers can help

evaluators make judgments about the quality of teaching, but student ratings are not, by themselves, evaluations.”

Instead, it is critical that any data from SETs be *interpreted* by knowledgeable and informed evaluators as part of a holistic process of faculty evaluation. Again, Bain (2004) stated,

Any good process [of teaching evaluation] should rely on appropriate sources of data, which are then *compiled* and *interpreted* by an evaluator or evaluative committee. Student remarks and ratings, in other words, are not evaluations; they are one set of data that an evaluator can take into consideration. . . an evaluation is an informed attempt to answer important questions, but it requires difficult decisions and can't be reduced to a formula. Professors and their evaluators should focus on the qualities of learning objectives and the efforts to help students achieve them rather than on numbers. (pp. 167–168)

Of specific relevance to Georgia Southern University, the University System of Georgia [USG] (2017) Board of Regents [BOR] [Policy Manual Section 8.3.5.1](#), paragraph 1, stated “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. Using uninterpreted data from SETs as the sole, primary, or even major component of summative faculty teaching evaluations is not just contrary to their purpose, it violates BOR policy.

Using quantitative data from SETs in statistically unjustifiable ways. It has also been well documented in the literature that administrators and others who evaluate faculty teaching by means of quantitative data from SETs do so in ways that are not statistically valid. As Spooen, Brockx, and Mortelmans (2013) have observed,

Administrators prefer aggregated and overall measures of student satisfaction, often failing to consider both basic statistical and methodological matters (e.g., response rate, score distribution, sample size) when interpreting SET (Gray & Bergmann, 2003; Menges, 2000) and making spurious inferences based on these data. For example, Franklin (2001) reported that about half of the SET administrators involved in the study were unable to provide sound answers to several basic statistical questions. (p. 622)

In addition, Boysen, Kelly, Raesly, and Casener (2014) reported,

Differences in means small enough to be within the margin of error significantly impacted. . . department heads' evaluation of teaching techniques. . . The results suggest that. . . administrators do not apply appropriate statistical principles when evaluating teaching evaluations and instead use a general heuristic that higher evaluations are better. (p. 641)

Further, even administrators *who received an explicit warning* against overinterpreting small and statistically non-significant differences in student evaluation means still judged instructional

methods more negatively if the methods were associated with small and statistically non-significant reductions in evaluation means (Boysen, 2014).

Even when statistically significant differences between means exist, because the amount of student learning that can be predicted by SETs is small at best, “interpreting fine distinctions in teaching evaluations is tantamount to making decisions about teachers based on factors extraneous to their primary job responsibility or, even worse, random error in measurement” (Boysen et al., 2014, p. 643).

Comparing mean SET scores also presumes that such comparisons are statistically valid. They are not (Stark & Freishtat, 2014):

Personnel reviews routinely compare instructors’ average scores to departmental averages. Such comparisons make no sense, as a matter of Statistics. They presume that the difference between 3 and 4 means the same thing as the difference between 6 and 7. They presume that the difference between 3 and 4 means the same thing to different students. They presume that 5 means the same thing to different students and to students in different courses. They presume that a 3 “balances” a 7 to make two 5s. For teaching evaluations, there’s no reason any of those things should be true. (pp. 5–6)

Instead, Stark and Frieshtat (2014, p. 2) argued, “Student ratings of teaching are valuable when they ask the right questions, report response rates and score distributions, and are balanced by a variety of other sources and methods to evaluate teaching.”

Failing to interpret qualitative data from SETs in context. Although student comments on SETs can have value in both faculty evaluation and for the purpose of improving faculty teaching, they are also known to be vulnerable to inappropriate, discriminatory, bullying, and abusive remarks (Vasey & Carroll, 2016). It is critical to remember that comments are anonymous and thus written both with impunity by the student and with no independent way to verify their veracity. As Stark and Freishtat (2014) cautioned,

While some student comments are informative, one must be quite careful interpreting the comments: faculty and students use the same vocabulary quite differently, ascribing quite different meanings to words such as “fair,” “professional,” “organized,” “challenging,” and “respectful” (Lauer, 2012). Moreover, it is not easy to compare comments across disciplines (Cashin, 1990; Cashin & Clegg, 1987; Cranton & Smith, 1986; Feldman, 1978), because the depth and quality of students’ comments vary widely by discipline. (p. 8)

If comments are to be used, any comments not about teaching effectiveness should be disregarded and those that do address teaching effectiveness should be interpreted in context and with caution.

Best Practice Recommendations

Based on the literature reviewed above, and other sources from the SET literature, the following list of (condensed) best practices is presented:

Do not

1. Use “omnibus items about ‘overall teaching effectiveness’ and ‘value of the course’” (Stark & Freishtat, 2014, p. 20). Such items are misleading and especially prone to being influenced by irrelevant factors (Stark & Freishtat, 2014).
2. Compute *means* for SET scores or compare means of SET scores (Boysen et al., 2014), whether between instructors or to cut-off scores (University of Wisconsin-La Crosse, 2007). “Such averages do not make sense statistically” (Stark & Freishtat, 2014, p. 20). Further, “Averages can emerge from a variety of distributions of ratings. . . Each distribution might suggest something quite different about the success of the teaching. . . What kind of teacher does the department want? What can help each one improve?” (Bain, 2004, p. 171). Means cannot inform these questions.
3. Compare teaching “in courses of different types, levels, sizes, functions, or disciplines.” (Stark & Freishtat, 2014, p. 20).
4. Generalize from responders to the SET to the whole class if response rates are low; such generalizations are unreliable (Stark & Freishtat, 2014).

Do

1. Ask SET questions that focus on what students learn (AERA, 2013). Asking about the SLOs for the course gets to the question, “Did the course deliver what it promised to deliver?” 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). It is also consistent with Georgia Southern University Faculty Handbook Section 205.01’s description of superior teaching as “focused on student learning outcomes.”
2. Ask SET questions that focus on student self-motivation to learn. “Students’ motivation determines, directs, and sustains what they do to learn. As students enter college and gain greater autonomy over what, when, and how they study and learn, motivation plays a critical role in guiding the direction, intensity, persistence, and quality of the learning behaviors in which they engage” (Ambrose et al., 2010, p. 5).
3. Ask SET questions about student effort and attendance. They “indicate the interest and motivation of students in a particular course and are at least partly dependent upon the organization of and the teaching in that course” (Spooren et al., 2013, p. 609).
4. Report the *distribution* of scores for individual items, the number of responders, and the response rate (Stark & Freishtat, 2014).
5. Treat SETs as a faculty development opportunity, focusing on formative rather than summative evaluation (Vasey & Carroll, 2016). This is also mandated by BOR policy, as noted above.
6. Interpret SETs in the context of the course, and taking into account research on SETs and biasing influences (Vasey & Carroll, 2016).
7. “Pay attention to student comments—but understand their limitations. Students typically are not well situated to evaluate pedagogy” (Stark & Freishtat, 2014, p. 20). In evaluations, written comments from student ratings of instruction that are not about teaching effectiveness should be disregarded. Comments that are about teaching effectiveness should be evaluated cautiously in the context of the course.
8. Continuously train administrators in how to appropriately use SET data. “The proper collection and interpretation of SET data depend upon administrators

having sound methodological training and regular briefing on the major findings and trends in the research field” (Spooren et al., 2013, p. 622).

Section III: Design of the Pilot Student Ratings of Instruction Instrument

In designing the pilot SRI instrument, the ad hoc Committee

1. was officially charged to develop an instrument that
 - a. would function primarily as a *formative* assessment *to inform and improve teaching effectiveness* (per BOR policy), and
 - b. was consistent with the recommendations of the 2014 ad hoc Committee on SRIs, namely
 - i. that it incorporate best practices from the research literature on SRIs (reviewed in Section II),
 - ii. that it focus on student learning, learning behaviors, and formative feedback, and
 - iii. that it provide opportunities for students to specify in writing how the instructor promoted student learning.
2. reviewed the report of the 2014 ad hoc Committee on SRIs,
3. reviewed existing measures for which we were given permission to use (Angelo, 2015; Skowronek, Friesen, & Masonjones, 2011),
4. reviewed the available information about the development and pilot testing of the SRI currently in use at Georgia Southern University (Georgia Southern University Faculty Senate, 2001, p. 4; Griffin, 2001), and
5. followed best practices in questionnaire design (Artino et al., 2014), including
 - a. synthesizing the results of a literature review with other data and considerations (e.g., BOR policy, logistical requirements, the results of the 2014 ad hoc Committee on SRIs’ findings from faculty and chair questionnaires, etc.),
 - b. developing items (DeVellis, 2016; Fowler, 2013; McCoach et al., 2013; Sedlmeier, 2006),
 - c. assessing response process validity (e.g., asking students how they interpret the items and response options),
 - d. pilot-testing the measure.

With respect to point #4, a comparison between the development and pilot testing of the SRI currently in use and the development and pilot testing efforts of the ad hoc Committee may be illustrative. Very little is documented about the pilot testing of the current SRI measure and nothing is documented about its development. See Table 1.

Table 1. SRI Committee Comparison

Element	2000 Provost Committee	2014-2017 ad hoc Committee
Type of committee	Provost’s “Student Ratings Committee”	ad hoc Senate Committee
Committee membership selected by	Provost Vandergrift	Senate Executive Committee

Committee charge and parameters for SRI instrument	Not documented	Appears in Section I of this Report.
Work officially documented through	1. One paragraph in February 12, 2001 Faculty Senate minutes under “Announcements: Vice Presidents: “[The Provost] distributed copies of the new Student Ratings of Instruction Instrument. He described the steps taken to develop this revised instrument. It will be used beginning with Spring 2001 classes and will be subject to periodic review.” <i>No other mentions in Faculty Senate minutes, Senate Librarian’s Report, or SEC minutes exist.</i>	1. 8 meetings’ minutes submitted to Librarian’s Report 2. 2 presentations to the Faculty Senate (2/15/16 & 3/6/17) 3. Draft Report submitted to Faculty Senate (1/31/17) 4. Final Report submitted to the Faculty Senate and Librarian’s Report (2/17/17)
Work unofficially documented through	1. Report on pilot study available on the faculty author’s webpage (Griffin, 2001)	1. 7 email progress updates to Senate Moderator 2. 1 meeting with Senate Moderator 3. 1 meeting with Provost
<i>Information Provided in Analysis</i>		
A. Questions on Pilot SRI	Neither exact items nor exact response scale provided, only brief descriptors (e.g., “Instructor Impartial”)	Copy of pilot measure
B. Sampling Procedure	Convenience	Stratified Random; strata 1. College 2. Graduate/undergraduate level 3. Small/medium/large class size 4. Online/face-to-face
C. Sample Size	2–4 instructors/college for a total of 24; population size not provided	132 instructors balanced across college; population size provided
D. Sample Characteristics	None provided	1. Compliance rate 2. Student response rate 3. College 4. Graduate/undergraduate 5. Class size 6. Online/face-to-face

D. Inclusion of online courses	N/A	35 online courses included
E. Pilot faculty feedback	None provided	Analysis of faculty responses to two closed-ended and three open-ended questions; response rate
F. Pilot department chair feedback	None provided	Analysis of chair responses to two closed-ended and three open-ended questions; response rate

See Appendix A for the complete pilot Student Ratings of Instruction measure.

Section IV: Methodology of the Pilot Study and Feedback Survey of Faculty and Chairs

Pilot Study Selection Criteria and Strata

Pilot testing for both online and face-to-face courses was originally planned for Spring 2016. However, because of administrative issues beyond the control of the ad hoc Committee, pilot testing for face-to-face courses had to be delayed until Fall 2016.

Courses were selected for the pilot test using a stratified random sampling procedure. First, to be included in the sample, the course had to normally administer SRIs per Georgia Southern University Policy, which excluded

- Courses with only 1 student enrolled
- Independent study, directed study, student teaching, internship/practicum, thesis, labs, and dissertation
- Courses taught by non-Georgia Southern faculty
- Courses with the subject code WBIT and WMBA
- Team taught courses

Next, courses were sorted by four strata: 1. Whether the course was taught online (as determined by being evaluated via CourseEval) or face-to-face [online]; 2. Whether the course was taught at the graduate or undergraduate level [level]; 3. Whether the course enrollment was Small (2-29), Medium (30-47), or Large (48+), based on a tripartite division of all Spring 2016 courses by enrollment [size]; and 4. The college in which the course was housed (for the purposes of this sampling, courses offered through the Vice-President for Academic Affairs were treated as belonging to the VPAA as their “college”) [college].

Course selection was balanced across strata as much as possible to meet target numbers, and courses were randomly selected from within each stratum. For example, approximately equal numbers of Small, Medium, and Large courses were selected, whereas three times as many face-to-face courses were selected as online courses. Once a faculty member’s course had been selected, that faculty member’s other courses were removed from the sample to prevent duplication of course instructors.⁵ Additionally, in the Fall 2016 face-to-face pilot test, any faculty who were selected for the Spring 2016 online pilot test were removed from the sample for the same reason.

The process yielded a population of 350 eligible online courses in Spring 2016 and 2,130 eligible face-to-face courses in Fall 2016, for a total population of 2,480. From this population, 35 courses were selected for the Spring 2016 online pilot test and 100 courses were selected for the Fall 2016 pilot test⁶.

⁵ Some courses at Georgia Southern University are taught by department chairs. Those courses were included in the population from which the sample was drawn because individuals may have different perspectives on the pilot SRI as a faculty member than as a department chair. As a result, some individuals who were selected for the sample were invited to complete a feedback questionnaire as both a faculty member and a department chair and may have completed both.

⁶ Three courses selected for inclusion in the Fall 2016 face-to-face pilot test were discovered to be inappropriate choices after selection, one for administrative reasons and two because the courses were half-term courses that had already ended by the time of selection. This reduced the Fall 2016 face-to-face sample to 97 courses.

Pilot Study Recruitment

On April 4th, 2016, the Committee notified the faculty whose courses had been selected for the Spring online pilot test of their selection via email. This notification included information about which course had been selected, instructions for how to administer the pilot test⁷, and a copy of the pilot SRI instrument. Additionally, faculty were informed that after the conclusion of the online pilot test, they would be invited to complete a brief online questionnaire about their perceptions of the pilot SRI instrument. See Appendix B for a copy of the faculty questionnaire. Finally, the notification also included the following text:

It is important to remember that the SRI undergoing pilot testing has not been approved as an official instrument for evaluating faculty by the Faculty Senate or President Bartels (beyond approval to conduct the pilot test). As a consequence, the individual results from this pilot test should not be used in any form of faculty evaluation (e.g., annual evaluations, tenure and promotion, pre/post-tenure review, etc.), nor should faculty be penalized for having been selected to participate in the pilot test. Our committee recommends simply noting in any materials where the faculty member's teaching is being reviewed, "Professor NAME's PREFIX:NUMBER: TITLE course was selected for inclusion in the pilot testing for a new SRI form in Spring 2016. As a result, SRI data for that course are not available."

Also on April 4th, the Committee notified the department chairs for each selected faculty member and provided them with the same information via email. See Appendix C for a copy of the department chair questionnaire.

On October 18th, 2016, the Committee notified the faculty whose courses had been selected for the Fall face-to-face pilot test of their selection via email. This notification included information about which course had been selected, that their department would be responsible for customizing the form to the SLOs for their course, and a copy of the pilot SRI instrument. Additionally, faculty were informed that after the conclusion of the pilot test, they would be invited to complete a brief online questionnaire about their perceptions of the pilot SRI instrument. Finally, the notification included the same caution about using the pilot data for faculty evaluation that was given with the Spring pilot test, updated as appropriate for the Fall semester.

Also on October 18th, the Committee notified the department chairs for each selected faculty member and provided them with information about which course(s) in their department had been selected, instructions for preparing the pilot forms for each course, a copy of the pilot SRI instrument, information about the online feedback questionnaire, and the cautionary note about using pilot data for faculty evaluation. This email notification also informed department chairs that the numerical data result files for the selected course(s) were due back to the ad hoc Committee chair no later than the first day of the Spring 2017 semester, January 9th, 2017, and

⁷ Online courses are typically evaluated in CourseEval. However, the CourseEval software could not give different versions of SRIs (i.e., one to courses selected for the pilot test with the pilot SRI, and one to the other courses with the existing SRI). As a result, online evaluations were collected via a Qualtrics questionnaire. Faculty were provided with a template where they would fill in the CRN, SLOs for the course, etc., and asked to post it to Folio along with the link to the Qualtrics questionnaire and instructions directing the students to the link.

that the chairs were to review the data and share it with the selected faculty members by that date so they would both be prepared to complete the online feedback questionnaire.

Feedback Survey Selection Criteria and Recruitment

On May 9th, 2016, the Committee emailed all selected faculty and chairs from the Spring 2016 online pilot for whom they had received usable pilot data. This email contained an invitation to provide feedback about the data provided by the pilot form and a link to a Qualtrics questionnaire. On May 31st, 2016, these faculty and chairs were sent a reminder email that again invited them to provide feedback via the Qualtrics questionnaire if they had not already done so. The questionnaire remained open until June 3rd, 2016.

On January 10th, 2017, the Committee emailed all selected faculty and chairs from the Fall 2016 face-to-face pilot from whom they had received usable pilot data by the January 9th deadline. This email contained an invitation to provide feedback about the data provided by the pilot form and a link to a Qualtrics questionnaire. The questionnaire remained open until January 23rd, 2017.

Section V: Results of the Pilot Study

General Information

The 132 courses/faculty sampled reflects just over 5% of the 2,480 eligible courses and 14% of the 938 eligible faculty at Georgia Southern University.⁸ A total of 100 of the 132 selected courses returned usable data, reflecting a compliance rate of almost 76%.⁹ A total of 36 department chairs had courses in their unit selected for the pilot test, reflecting 82% of the 44 eligible departments including FYE and Honors.¹⁰ Of those 36 chairs, 28 had faculty with course data/returned data by the due date.

The number of student respondents varied from one to 136 across the courses. A total of 2,513 student responses (318 online, 2,195 face-to-face) were obtained from a total enrollment of 4,045 students (1,086 online, 2,959 face-to-face) in the 100 courses, reflecting an overall response rate of 62.13%, an online response rate of 29.28%, and a face-to-face response rate of 74.18%. Typical response rates for Student Evaluations of Teaching are 30-50%, which is considered “low” (Spooren et al., 2013).

Sample by Stratum

Table 2. Sample by delivery method

Delivery Method	Sampled	Returned
Online	35	30
Face-to-face	97	70

Table 3. Sample by course level

Course Level	Sampled	Returned
Undergraduate	93	71
Graduate	39	29

Table 4. Sample by course size

Course Size	Sampled	Returned
Small (2-29)	46	33
Medium (30-47)	41	30
Large (48+)	45	37

⁸ Based on data from the 2015-2016 Georgia Southern University Factbook.

⁹ “Returned” means that usable data were received by the ad hoc Committee by the deadline. In the case of software processing errors that rendered the data unreadable/unusable, such data were treated as “unreturned” for the purposes of these numbers, as they could not be used in the Committee’s analyses.

¹⁰ Some individuals chaired more than one department, and some departments experienced a change in chair from the Spring 2016 online pilot to Fall 2016 face-to-face pilot, so this percentage may not be precise.

Table 5. College in which the course was housed

College	Sampled	Returned
COBA	16	12
COE	18	12
CEIT	16	13
CHHS	17	14
CLASS	18	12
JPHCOPH	16	15
COSM	17	10
VPAA	14	12

Compliance Rate Analyses

To examine if the courses for which no usable data were received differed on the selection strata from the courses for which usable data were received, a series of planned comparisons was executed. Chi-square tests revealed no significant differences between the groups for

1. The ratio of online to face-to-face courses, $\chi^2 (1, N = 132) = 2.57, p = .11$;
2. The ratio of undergraduate to graduate courses, $\chi^2 (1, N = 132) = 0.059, p = .81$;
3. The ratio of small to medium to large courses, $\chi^2 (1, N = 132) = 1.58, p = .45$.

Additionally, an independent *t*-test comparing the average enrollment between the two groups revealed no significant difference, $t (38.08) = 0.37, p = .71$, equal variances not assumed.

Because it was not expected to have at least five unusable courses in each college (nor was it observed for the majority of colleges), statistical comparisons of missing courses by college were not possible. Instead, visual inspection of distributions was used. Compliance rates varied from a low of 59% to a high of 94%, but all reflected a clear majority of courses sampled.

These results give the Committee confidence that the final sample accurately reflects the strata from which it was drawn.

Strata Analyses¹¹

Within each of the three strata (online, level, size), comparisons of the distributions of responses for each of the 11 non-SLO variables (student, course, instructor) were conducted with nonparametric tests. The attendance variable was not included in these analyses, as it did not apply to online courses. A Bonferroni correction to original alpha level of $p < .01$ for 33 comparisons yielded an adjusted $p < 0.0005$ for each test.

Online vs. Face-to-Face. Three significant differences emerged via Mann-Whitney U-tests:

- Student: Expected grade: $U = 220,375.5, p < .0005$. Students enrolled in online courses reported higher expected grades than students in face-to-face courses.
- Instructor: Presentations/Explanations: $U = 270,174, p < .0005$. Students enrolled in face-to-face courses were more likely to rank instructors' presentations/explanations as

¹¹ "College" was chosen as a stratum to ensure equal distribution of the sample across the Colleges, not for planned comparisons in the data analysis.

having “A large amount” of influence on their learning than students enrolled in online courses.

- Instructor: Enthusiasm: $U = 242,952, p < .0005$. Students enrolled in face-to-face courses were more likely to rank instructors’ enthusiasm for the subject as having “A large amount” of influence on their learning than students enrolled in online courses.

Given popular student expectations for online courses, and logistical differences between the two delivery methods, these differences are not surprising. For the other eight variables, no significant differences between online and face-to-face courses emerged.

Undergraduate vs. Graduate. Six significant differences emerged via Mann-Whitney U-tests:

- Student: Self-motivated to learn: $U = 450,261.5, p < .0005$. Students enrolled in graduate courses reported higher levels of self-motivation to learn than students enrolled in undergraduate courses.
- Student: Seek professor’s assistance: $U = 447,563, p < .0005$. Students enrolled in graduate courses reported higher levels of seeking assistance from their professor than students enrolled in undergraduate courses.
- Student: Study time: $U = 508,626, p < .0005$. Students enrolled in graduate courses reported higher levels of study hours than students enrolled in undergraduate courses.
- Student: Expected grade: $U = 247,024, p < .0005$. Students enrolled in graduate courses reported higher expected grades than students enrolled in undergraduate courses.
- Course: Interesting: $U = 333,830.5, p < .0005$. Students enrolled in graduate courses reported higher levels of interest in the course than students enrolled in undergraduate courses.
- Instructor: Connected SLOs: $U = 442,253.5, p < .0005$. Students enrolled in graduate courses reported higher levels of influence on their learning from the instructor connecting the SLOs to the class materials.

Given typical differences between undergraduate and graduate students, the first five differences would be expected. The sixth difference may be a product of different SLOs for graduate courses. For the remaining five variables, no significant differences between undergraduate and graduate courses emerged.

Small vs. Medium vs. Large Class Size. Four differences emerged via Kruskal-Wallis tests:

- Student: Seek professor’s assistance: $H(2) = 47.66, p < .0005$. Students enrolled in Large courses reported lower levels of seeking assistance from their professor than students enrolled in Small or Medium courses.
- Student: Study time: $H(2) = 18.47, p < .0005$. Students enrolled in Small courses reported higher levels of study hours than students enrolled in Medium or Large courses.
- Student: Expected grade: $H(2) = 26.21, p < .0005$. Students enrolled in Large courses reported lower expected grades than students enrolled in Small or Medium courses.
- Course: Difficulty: $H(2) = 62.19, p < .0005$. Students enrolled in Medium courses reported lower levels of difficulty with the subject matter of the course than students enrolled in Small or Large courses.

Given typical differences between Small, Medium, and Large courses (e.g., small courses are more likely to be upper division or graduate level and in a student's major), the first three differences are not surprising. The fourth difference may be an artifact of course sequencing: introductory "gateway" courses may prevent students who have the most difficulty with the material from advancing to upper division courses; such introductory courses are often Large, whereas mid-level courses may be more common in the Medium size. For the other seven variables, no significant differences between Small, Medium, and Large courses emerged.

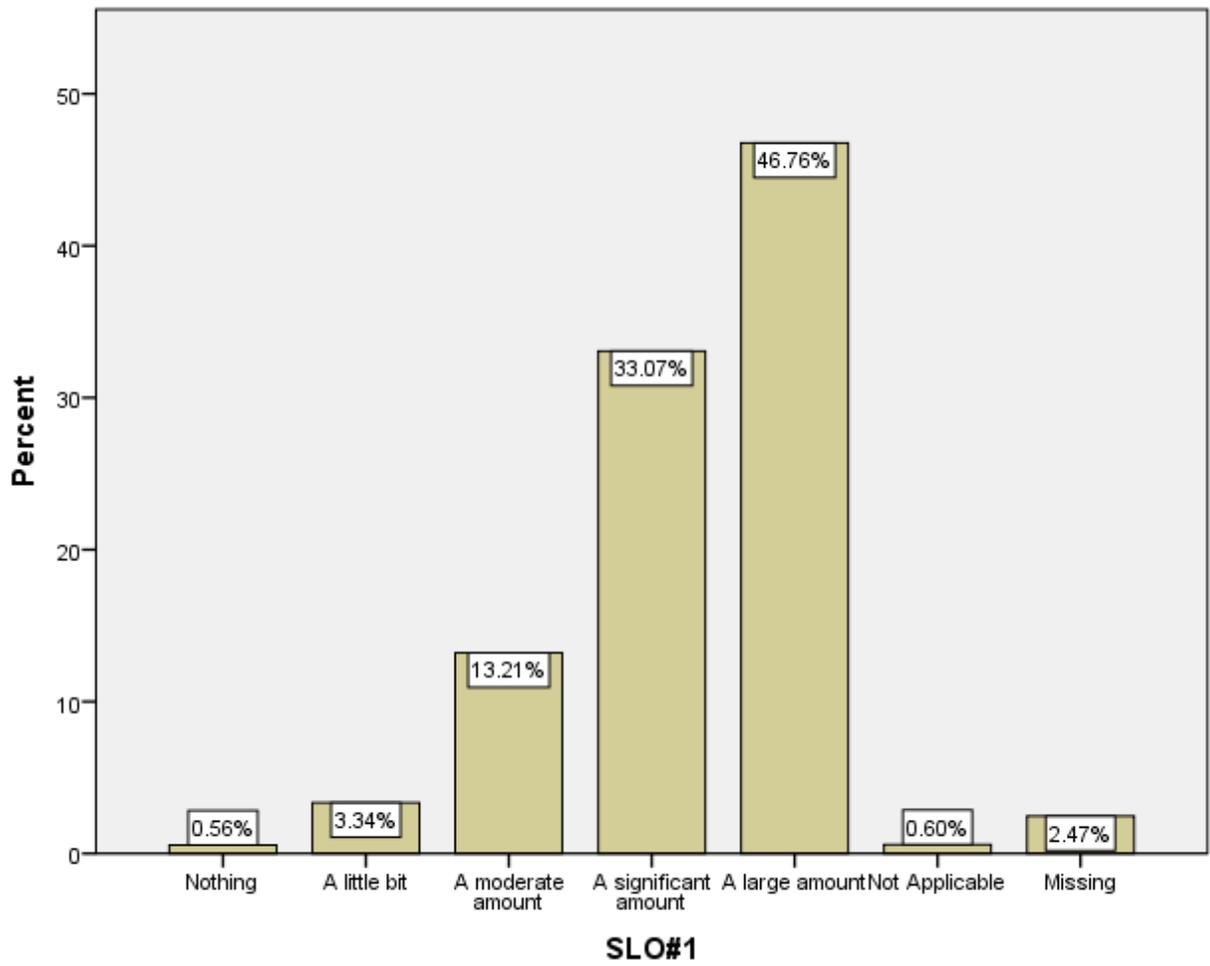
Frequency Distributions for Pilot SRI Items

Below, frequency distributions for each item on the pilot SRI are presented, organized by the first four major sections of the SRI: Student Learning Objectives, The Student, The Course, and The Instructor. The fifth section, Comments on the Learning Experience, requested open-ended qualitative responses from students and as such was not analyzed for this report.

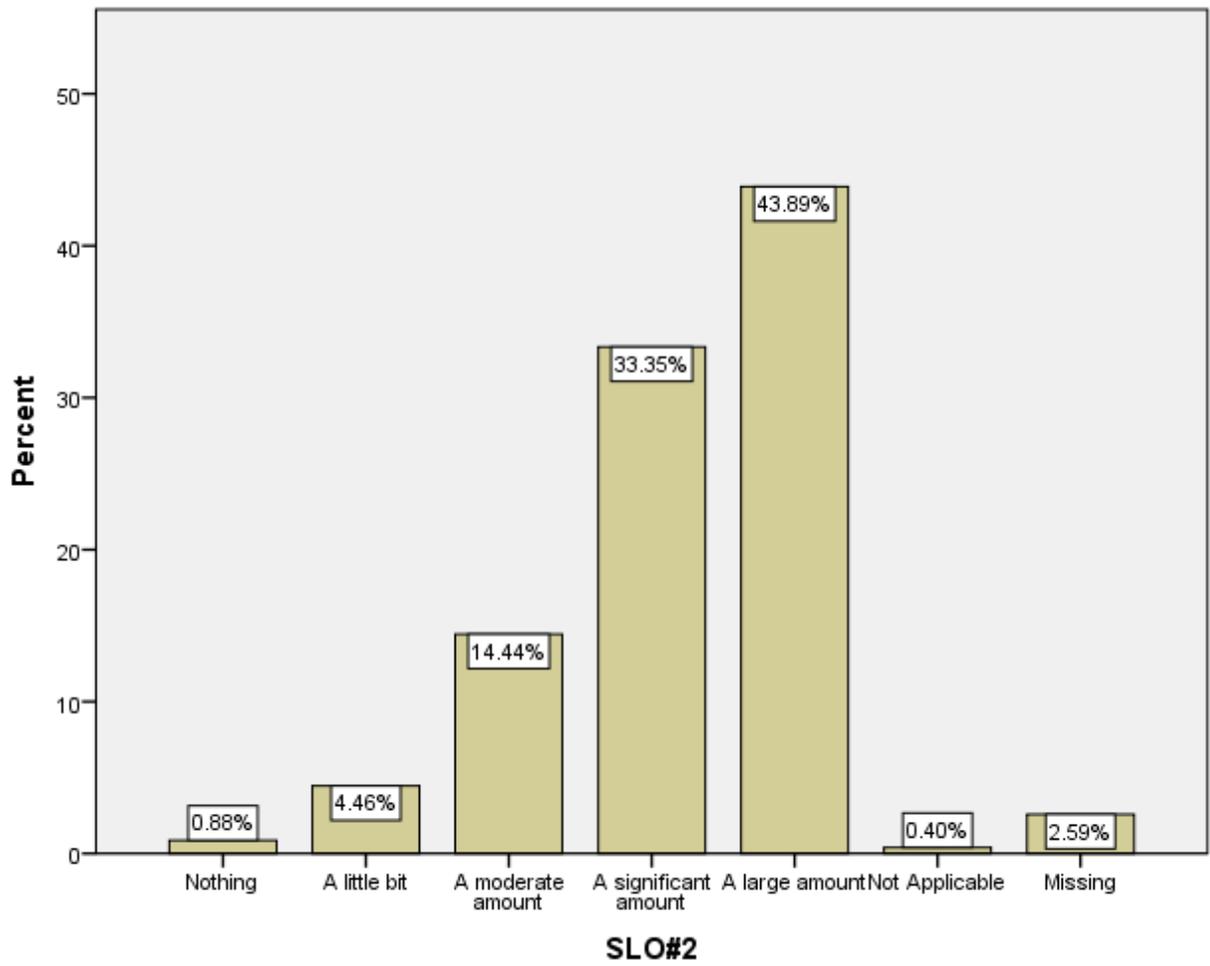
Section I: Student learning objectives. This section contained 10 items. The instructions read, "Indicate how much you learned on each student learning objective for this course with the following scale." Response options were 1 = *Nothing*, 2 = *A little bit*, 3 = *A moderate amount*, 4 = *A significant amount*, 5 = *A large amount*, N = *Not Applicable*. The sixth scale option was labeled "N" instead of "6" to clearly indicate it was not a value on the ordinal scale. Items 1-10 were the Student Learning Objectives [SLOs] for that course. In the event that a course had fewer than 10 SLOs, students were instructed "Select 'N' for this question." Note: This is the only section where missing data is included in the graphs, because a significant number of respondents chose to leave items blank instead of selecting "N."

Note: N = 2,513 for SLO#1-10 because missing data are included as a separate bar in the graphs.

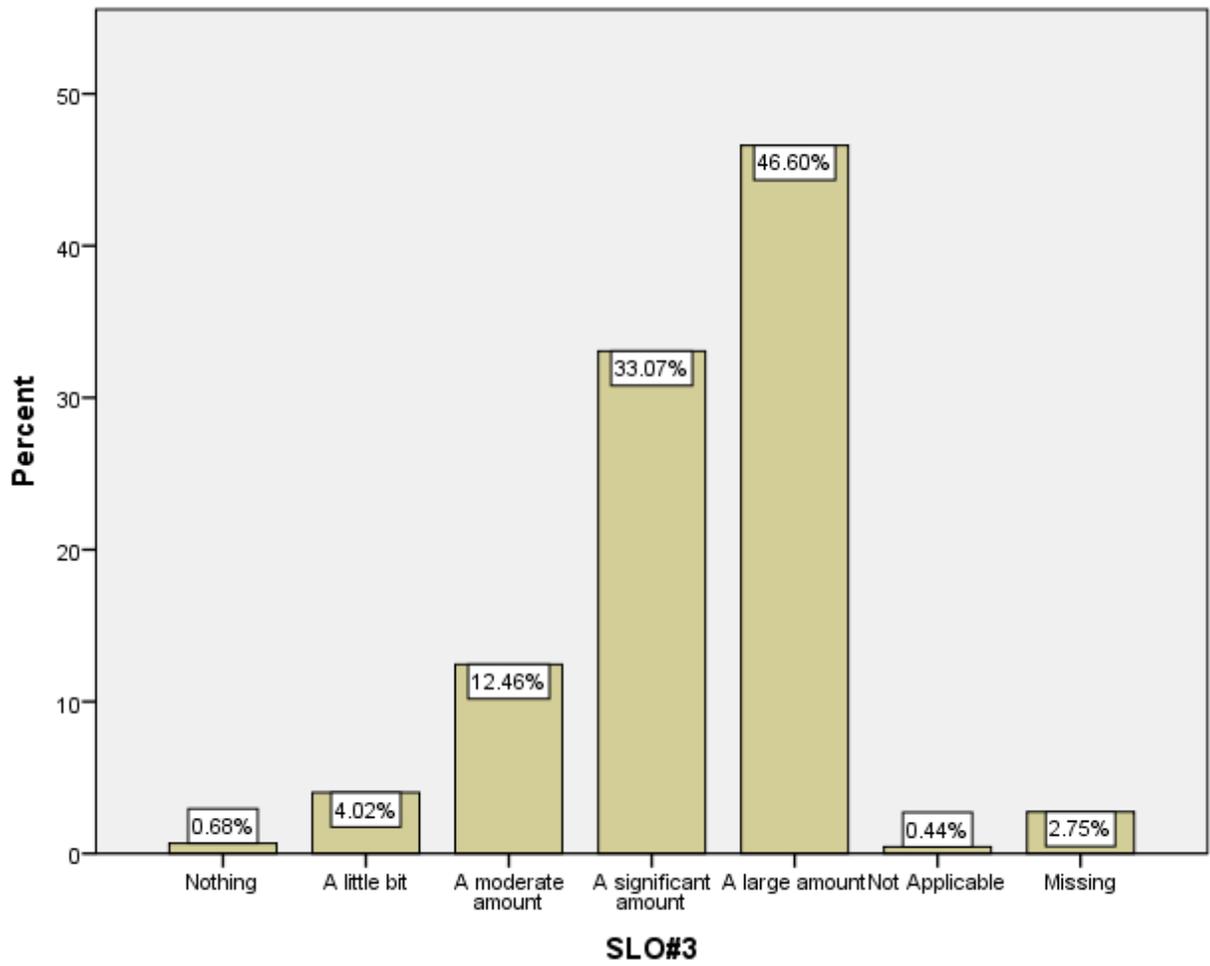
Q1: SLO#1.



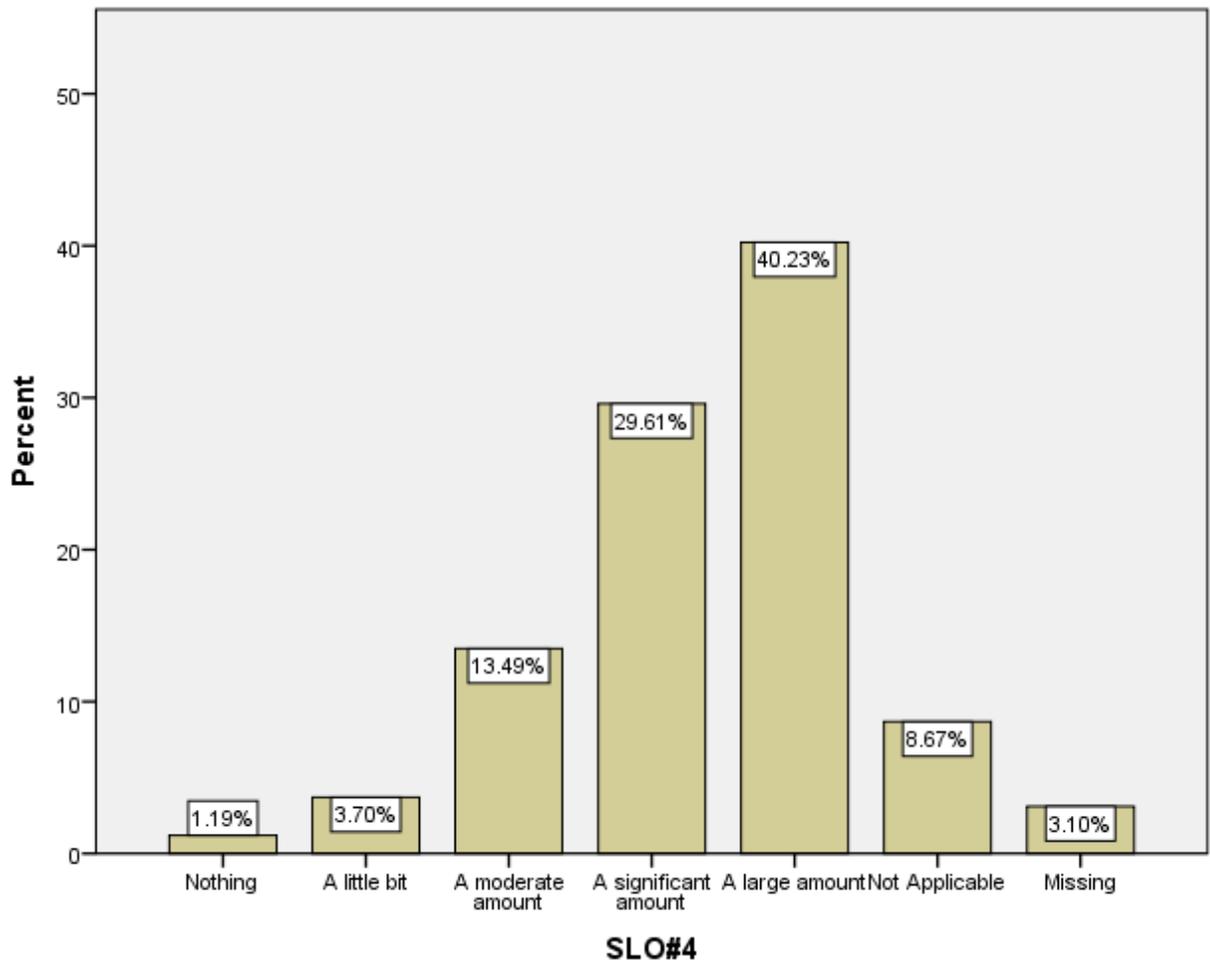
Q2: SLO#2.



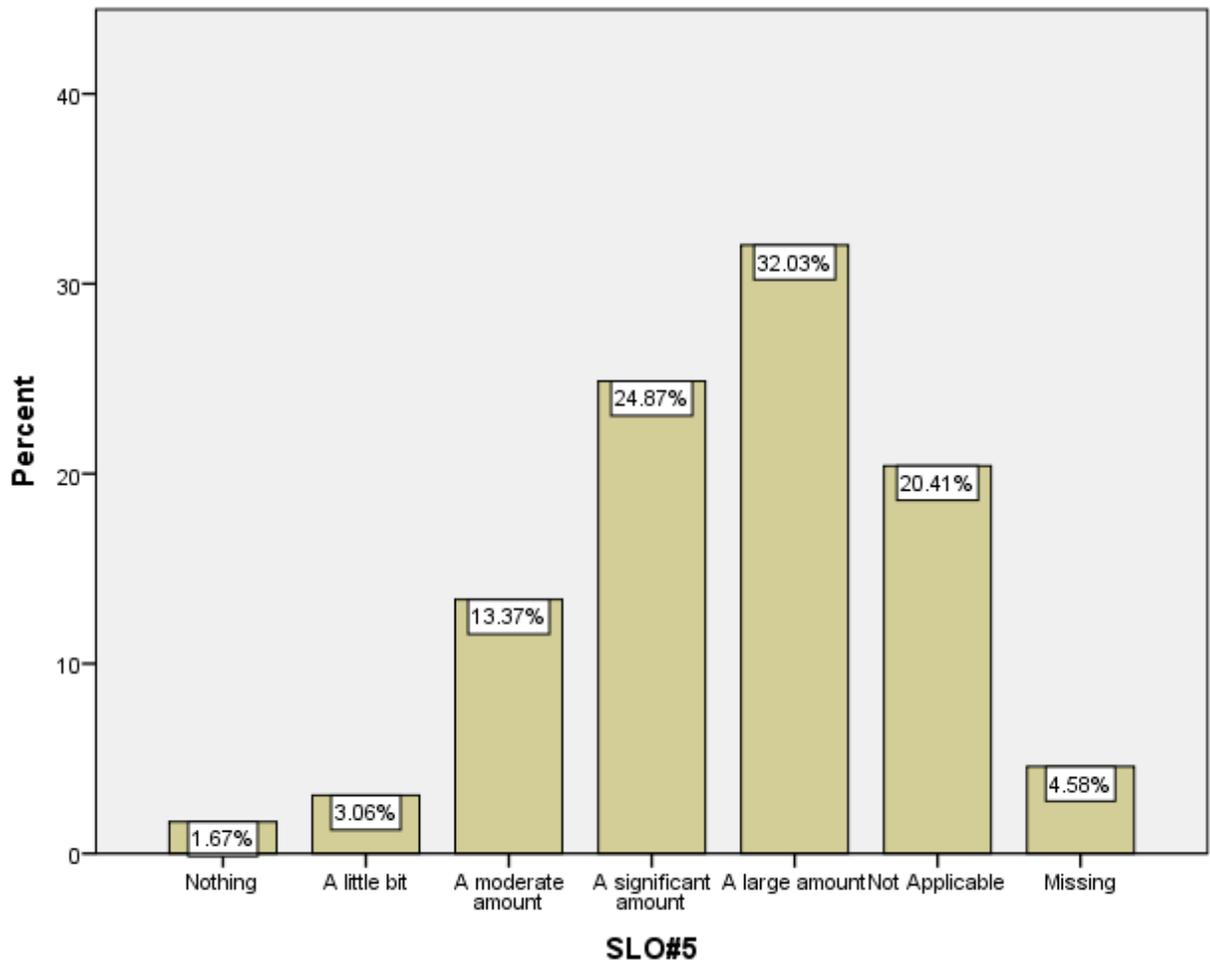
Q3: SLO#3.



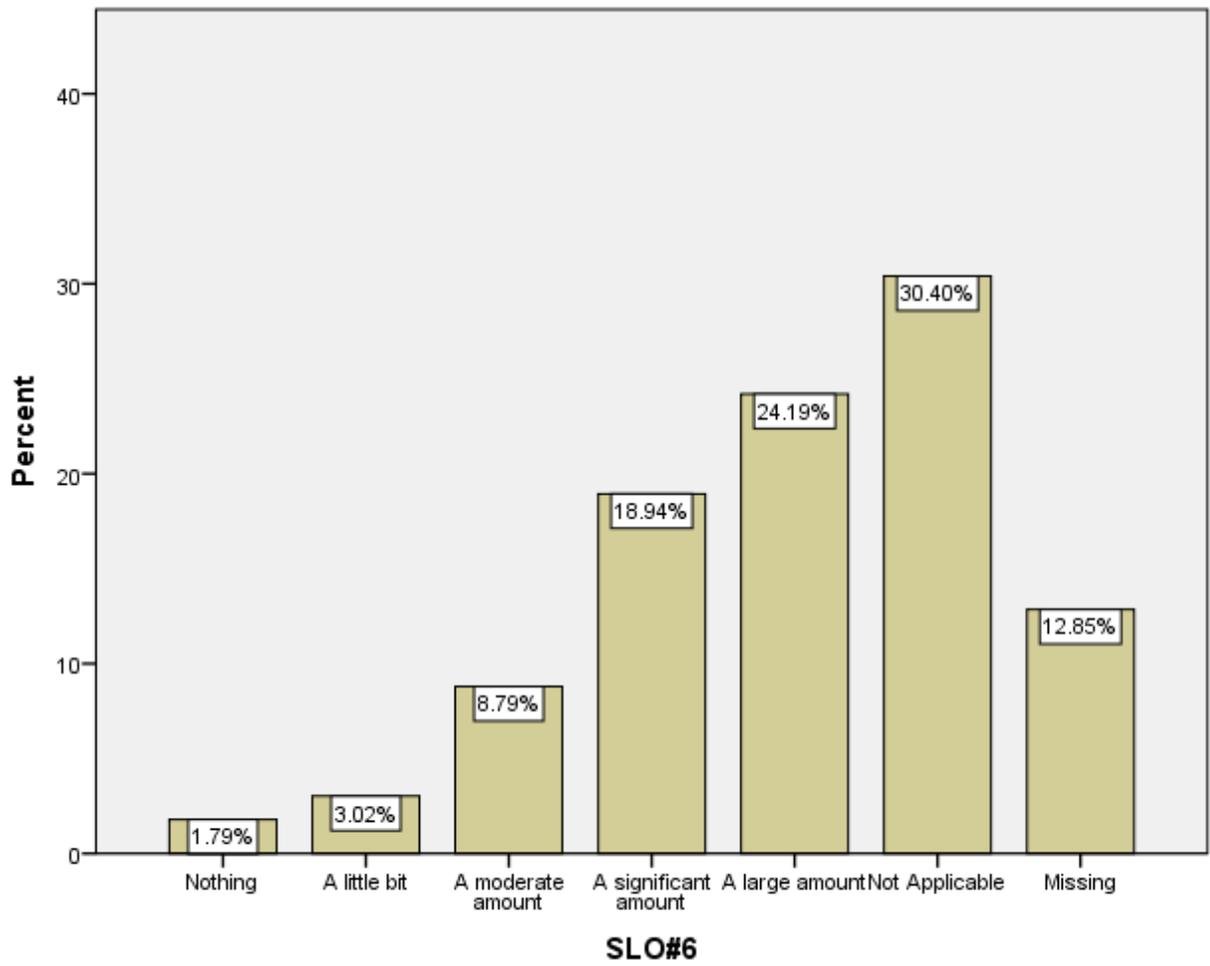
Q4: SLO#4.



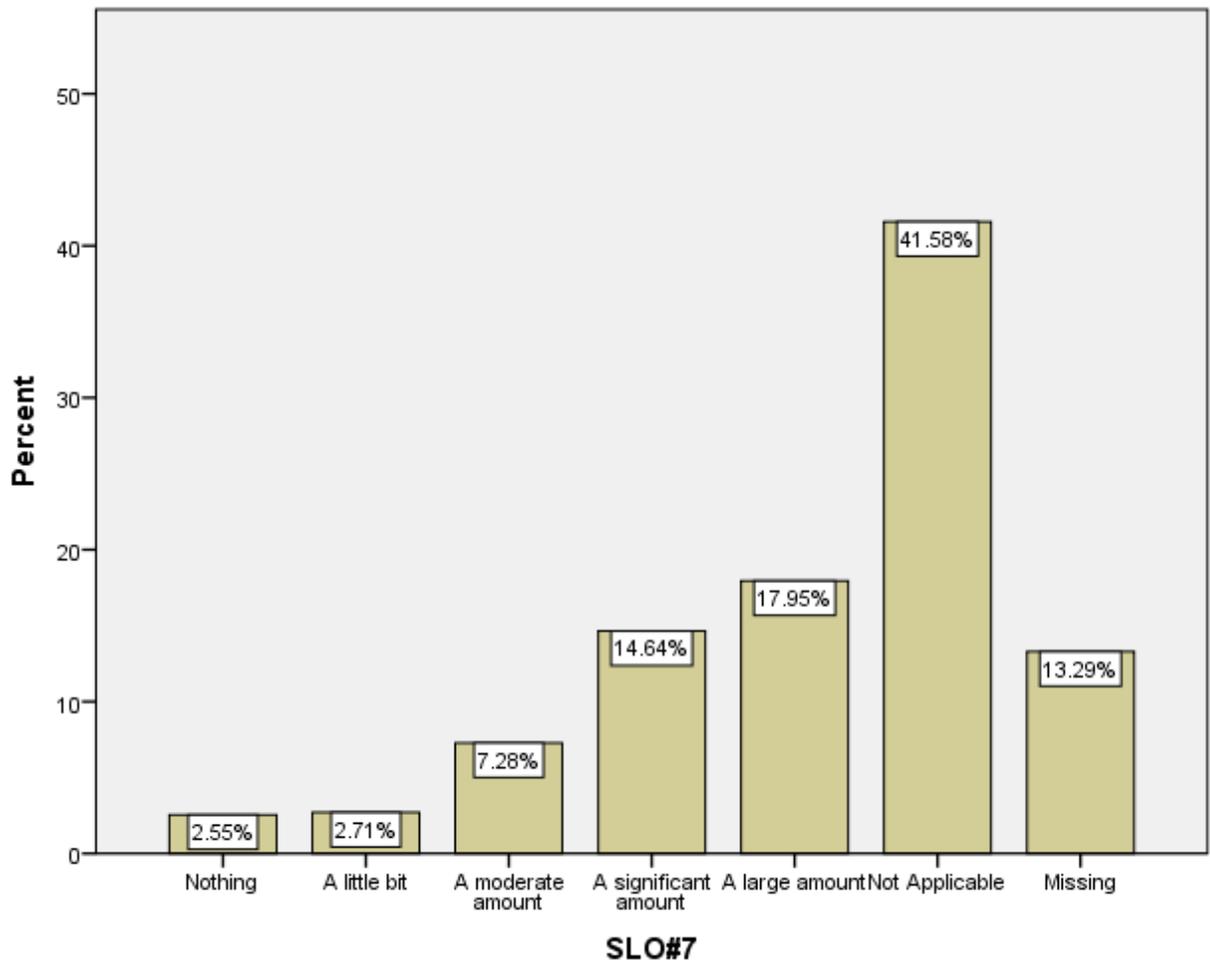
Q5: SLO#5.



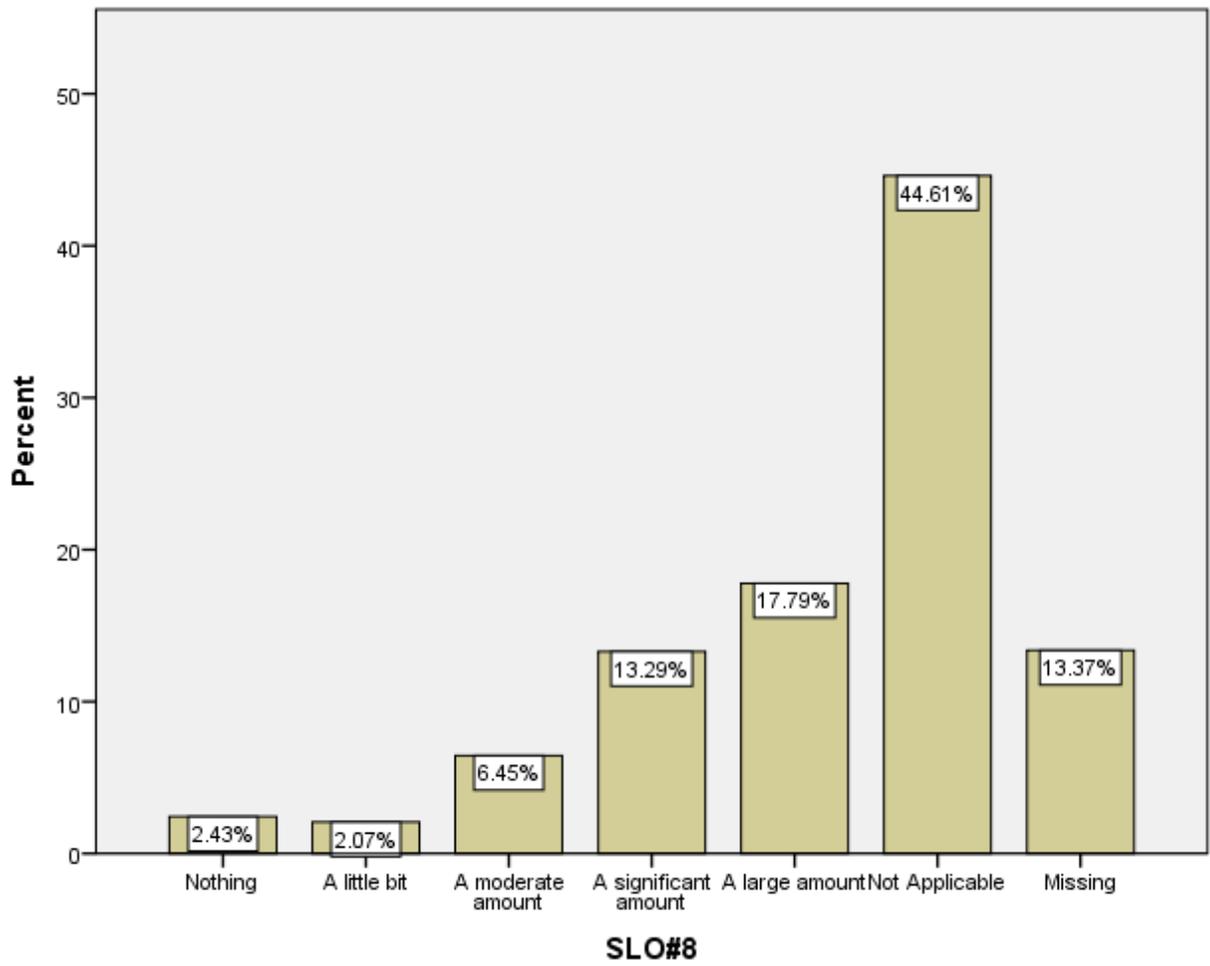
Q6: SLO#6.



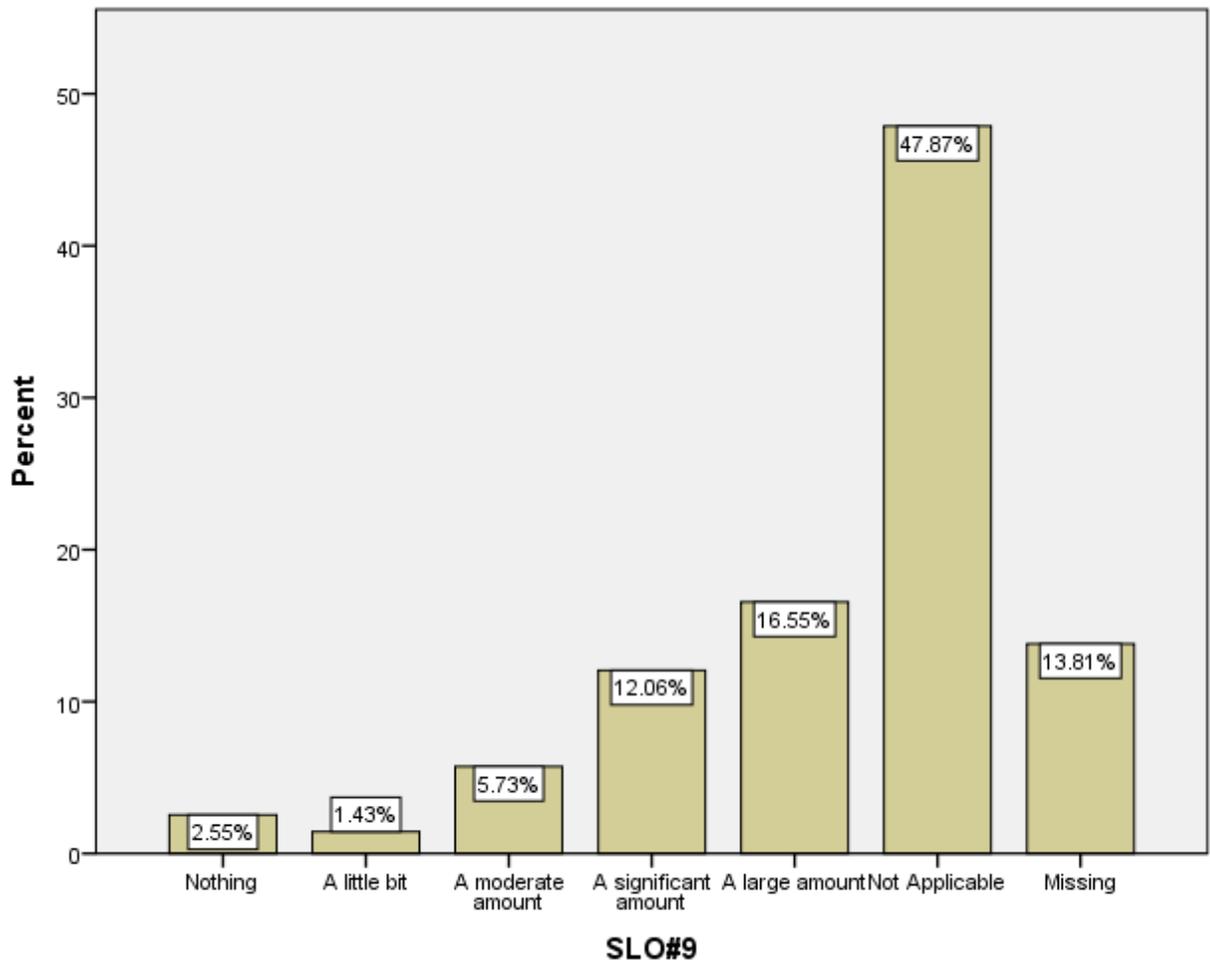
Q7: SLO#7.



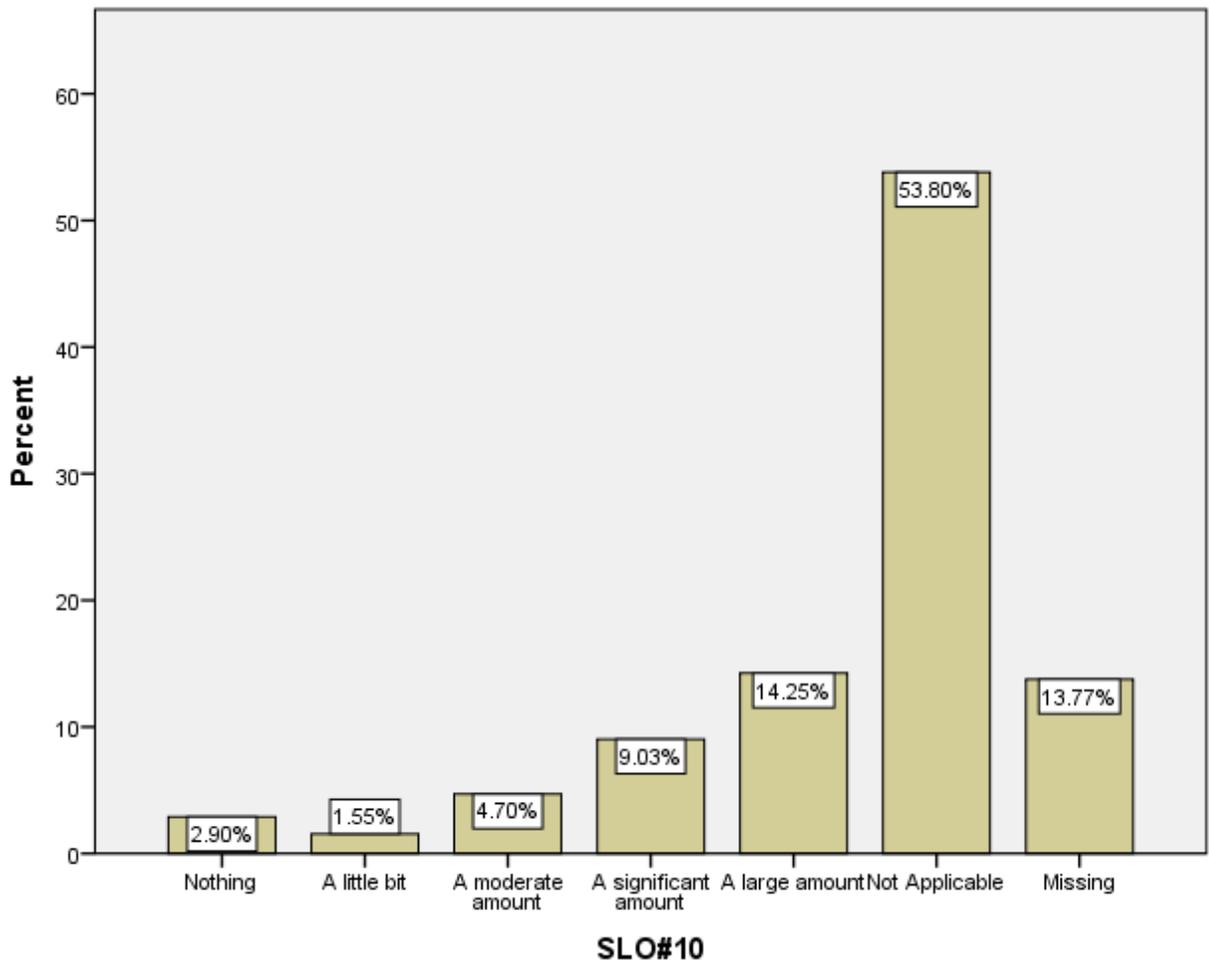
Q8: SLO#8.



Q9: SLO#9.



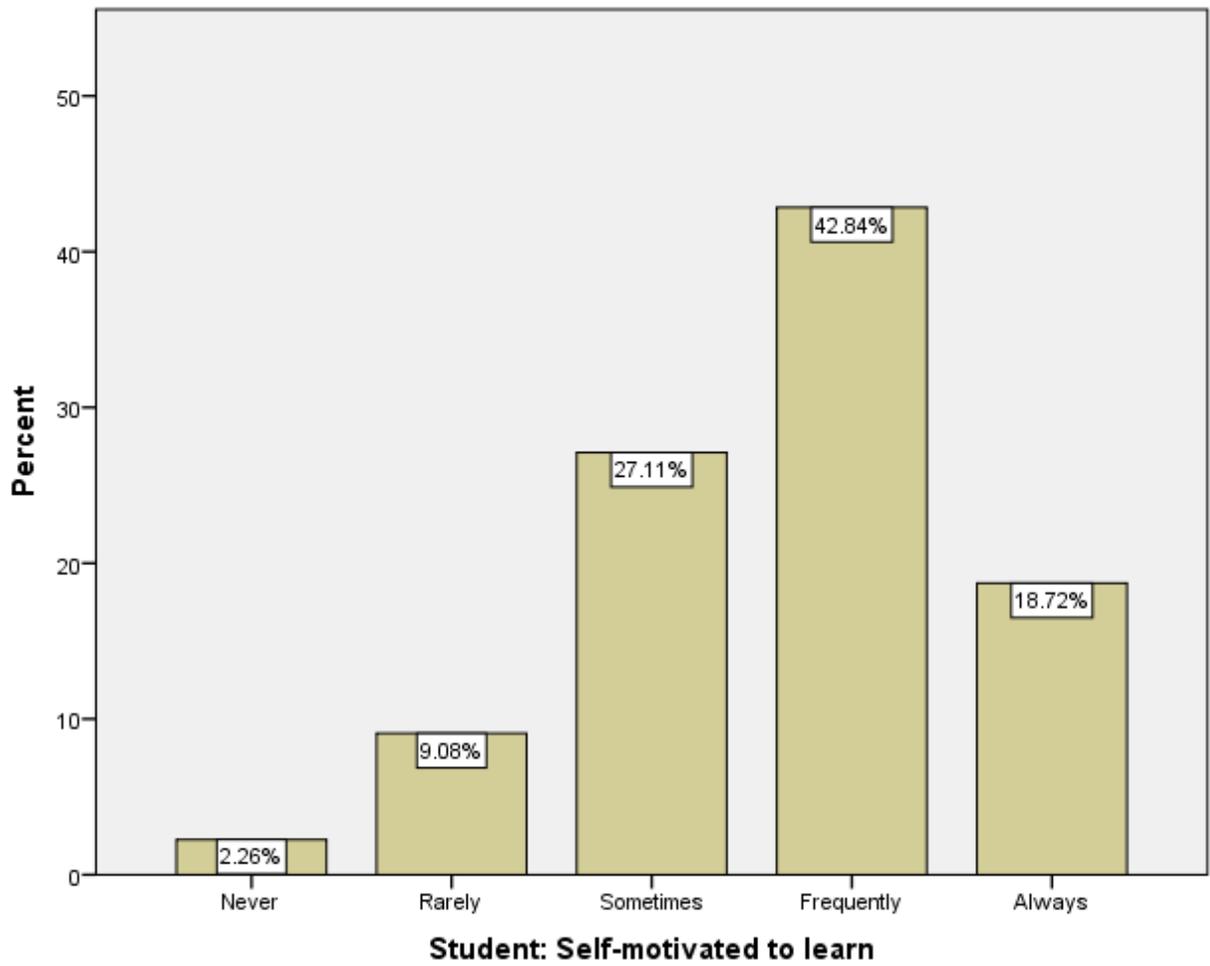
Q10: SLO#10.



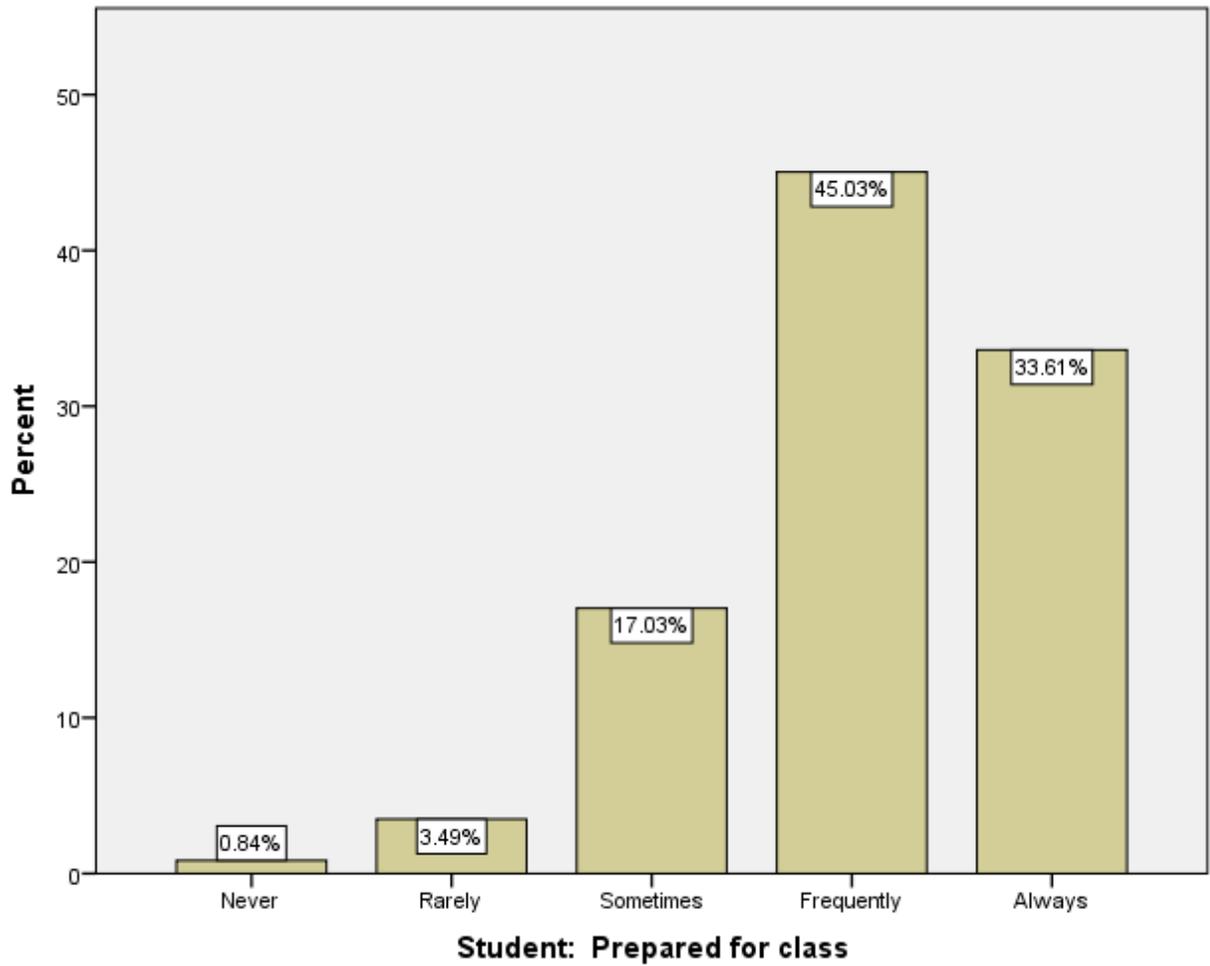
As would be expected, the number of missing or Not Applicable responses increased with each question, reflecting the greater probability of a course not having that many SLOs as the number of SLOs increased. Additionally, the greater frequency of responses in the “A significant amount” and “A large amount” categories is consistent with the overall expectation of effective teaching at Georgia Southern University.

Section II: The student. This section contained six items. The instructions read, “Student learning depends in large part on what the student does. In this section, we will ask you questions about you that are connected to your learning in this course.” Response options for the first three questions were 1 = *Never*, 2 = *Rarely*, 3 = *Sometimes*, 4 = *Frequently*, 5 = *Always*.

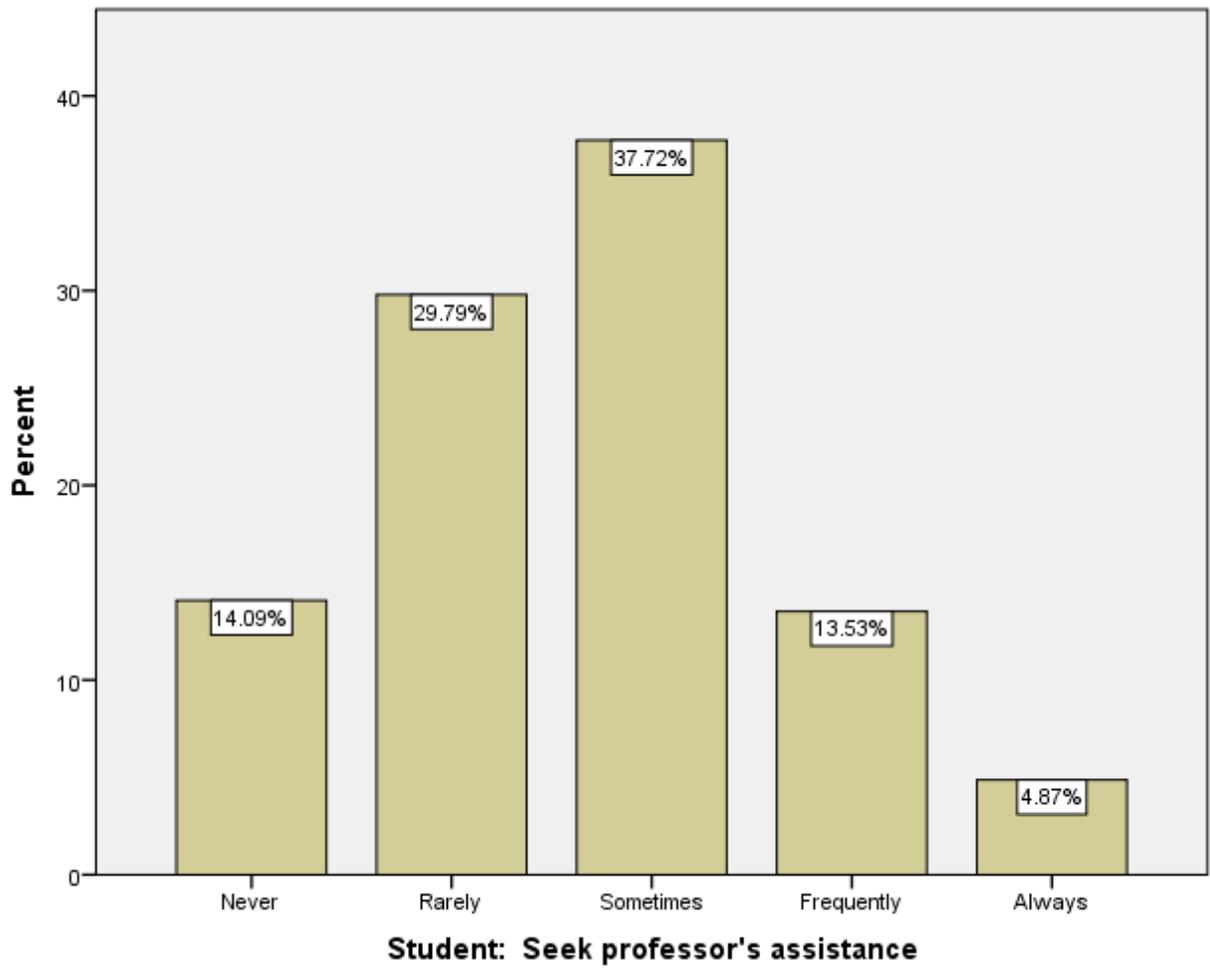
Q11: How often were you self-motivated to learn this course material? (N = 2,479).



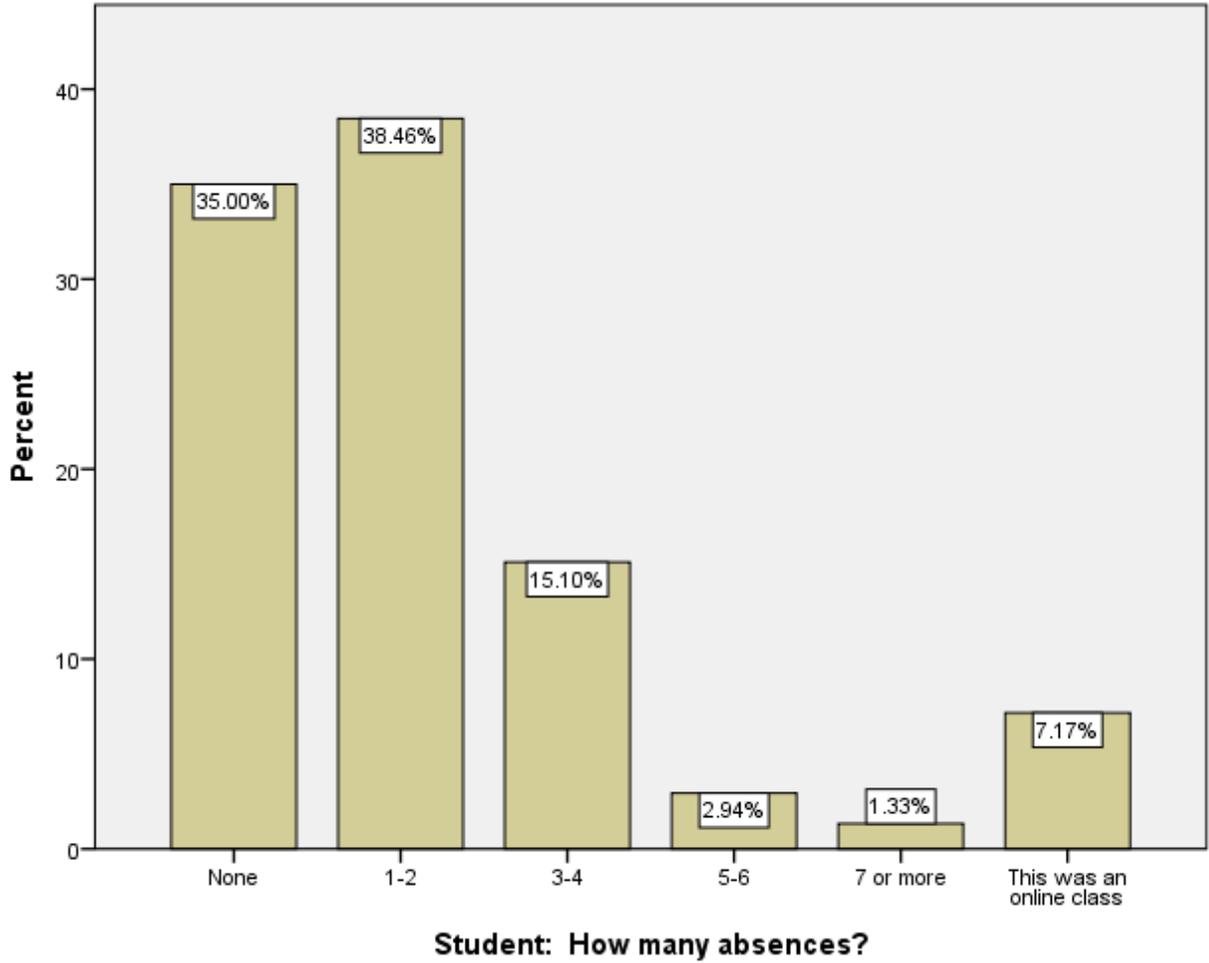
Q12: Based on the professor's stated expectations, how often were you fully prepared for class? (N = 2,496).



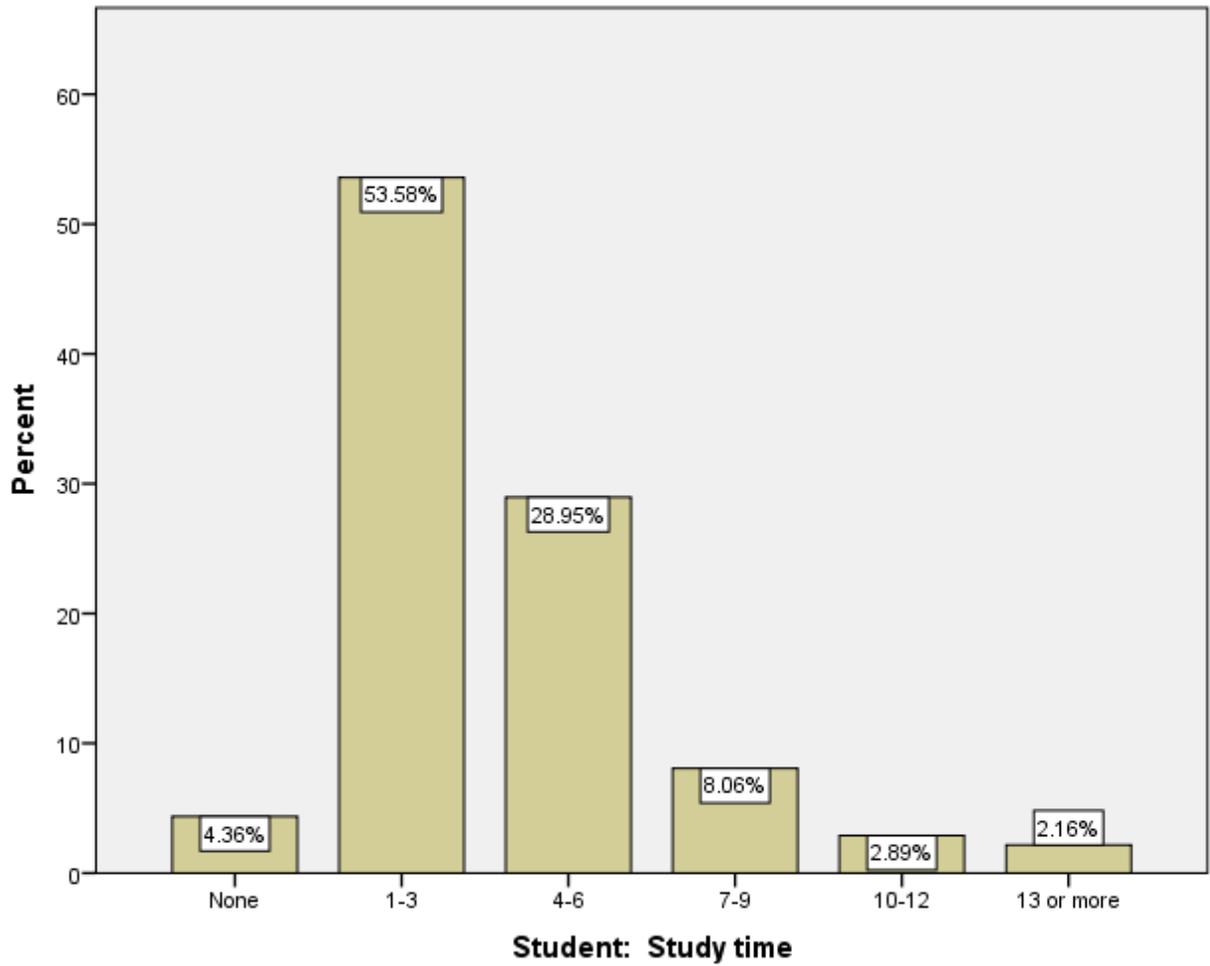
Q13: How often did you seek the professor's assistance? (N = 2,484).



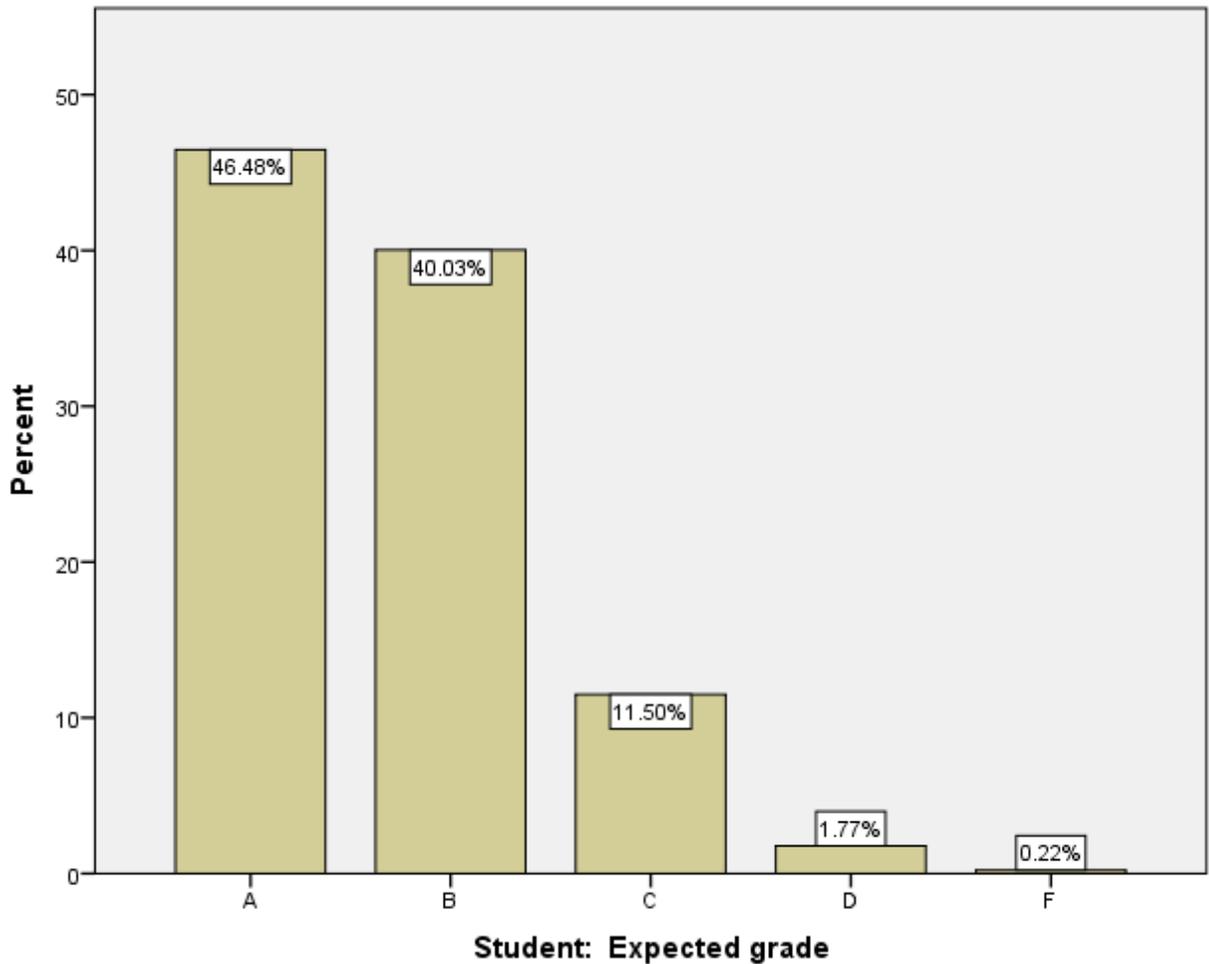
Q14: How many class meetings did you miss in this course? (N =2,483). Question 14 asked about attendance. Response options were A = None, B = 1-2, C = 3-4, D = 5-6, E = 7 or more, F = This was an online class. Option F was included so that students enrolled in fully online courses could still respond to this question.



Q15: Approximately how many hours per week did you spend on this course outside of scheduled class time? (N = 2,456). Question 15 asked about study time. Response options were A = None, B = 1-3, C = 4-6, D = 7-9, E = 10-12, and F = 13 or more.



Q16: What grade do you think you will earn in this course? (N = 2,313). Question 16 asked about expected grade. Response options were A, B, C, D, and F.



Analysis. For the student preparation question (Q12), data from Georgia Southern University’s 2015 National Survey of Student Engagement [NSSE] are available for comparison. One NSSE item asks, “During the current school year, about how often have you done the following? Come to class without completing the readings or assignments?” Response options are *Very often*, *Often*, *Sometimes*, and *Never*. Among Georgia Southern University Seniors, 23% selected either *Very often* or *Often*. These numbers map very closely onto the percentage of students who answered that they were *Rarely* or *Sometimes* prepared for class on the pilot SRI.

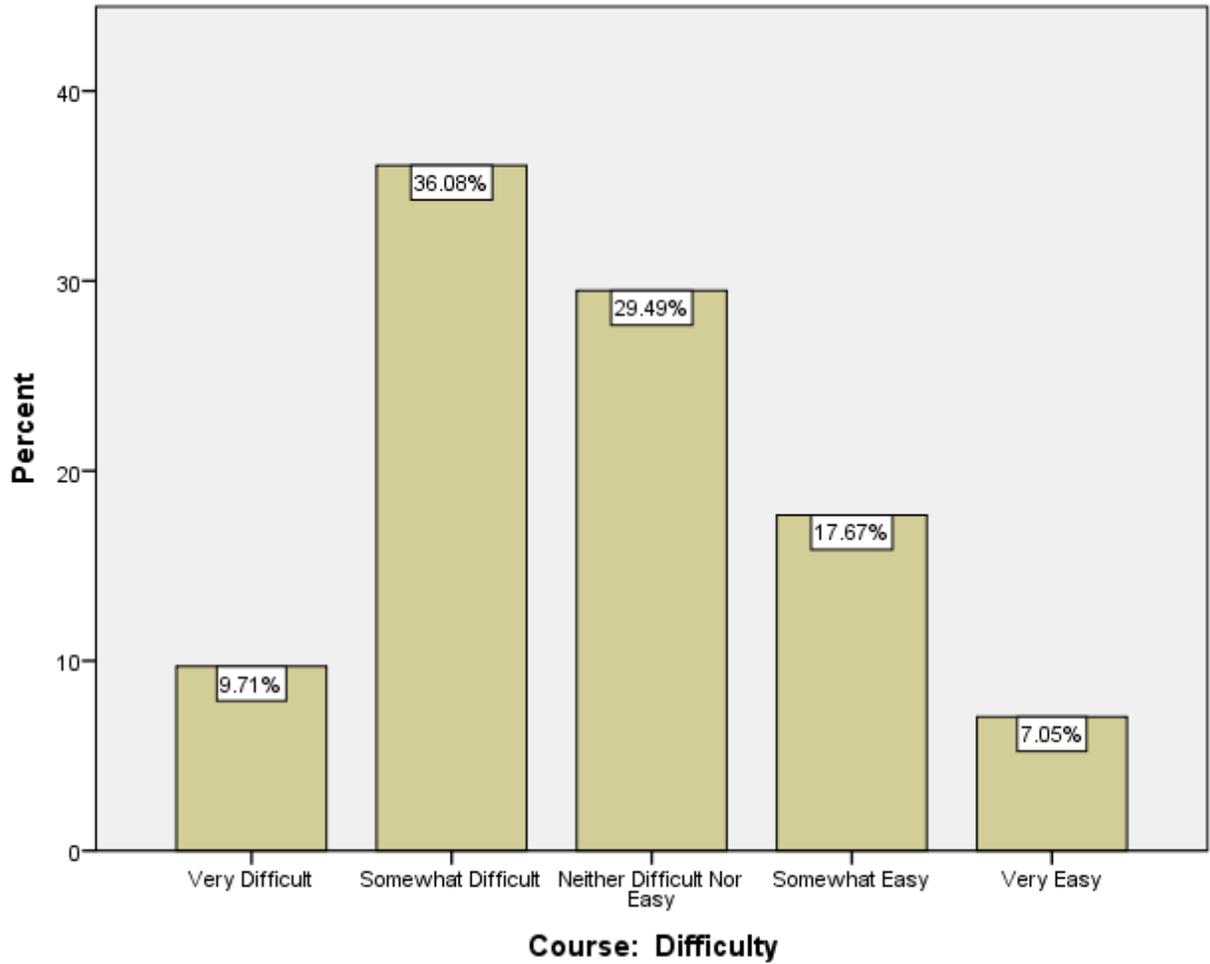
The different distribution of answers for Q13 (seeking assistance) from Q11 (self-motivation) and Q12 (preparation) suggests students were closely reading the questions rather than simply locating the column on the form that reflected their overall level of enjoyment of the course and then marking all questions with that score (Titus, 2008).

For the class attendance question (Q14), over 70% of student respondents reported two or fewer absences. Given the 62% response rate to the SRI, this seems fairly accurate, as students who more frequently were absent were likely also absent when the SRIs were administered.

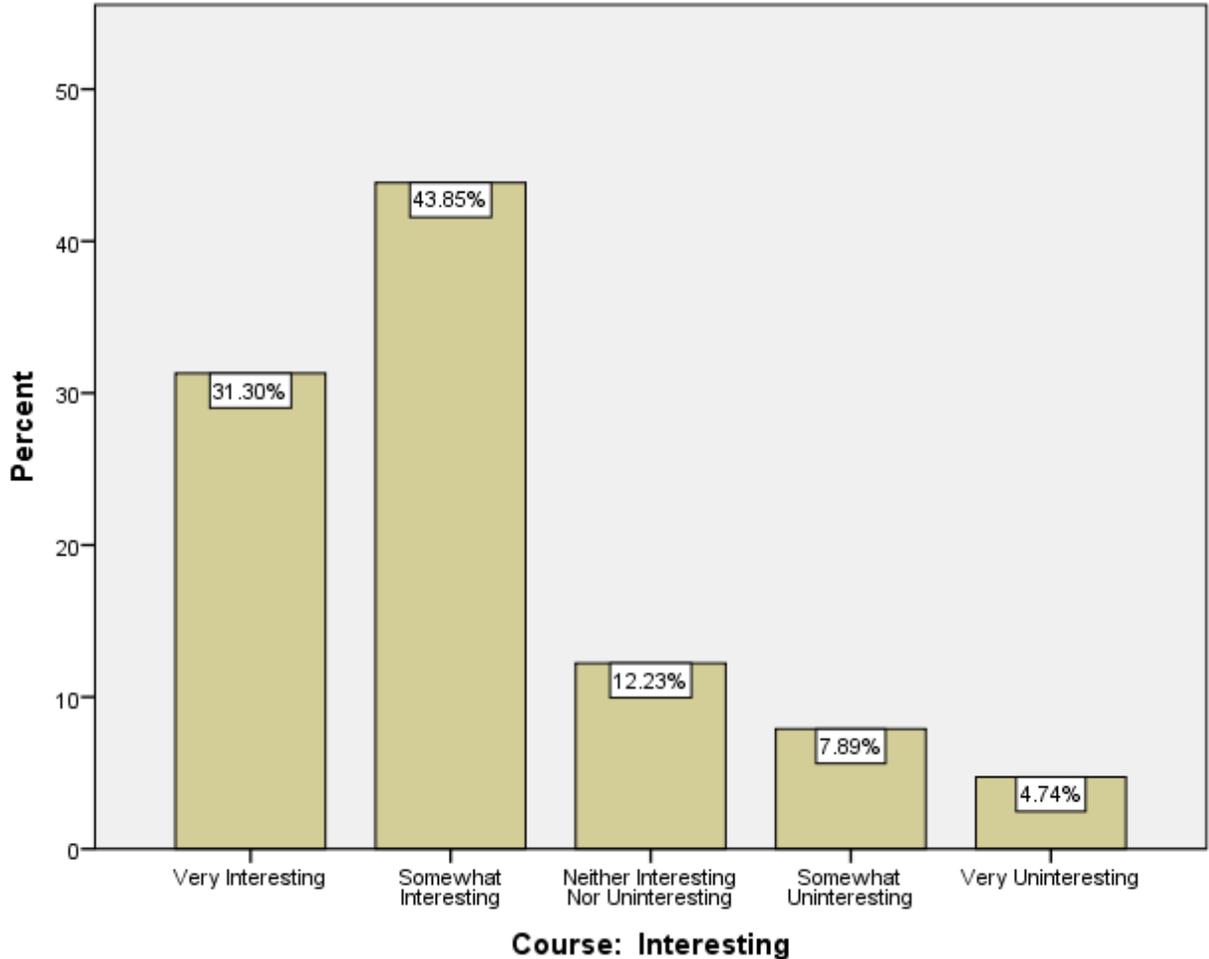
For the study time question (Q15), again NSSE data are available for comparison. One NSSE item asks, “About how many hours do you spend in a typical 7-day week doing the following? Preparing for class (studying, reading, writing, doing homework or lab work, analyzing data, rehearsing, and other academic activities).” Response options are *0 hrs*, *1-5 hrs*, *6-10 hrs*, *11-15 hrs*, *16-20 hrs*, *21-25 hrs*, *26-30 hrs*, and *More than 30 hrs*. Although the NSSE item combines study time for all courses, it is possible to approximate an appropriate comparison by multiplying the response options on the pilot SRI by five (reflecting five times the study time for five similar courses in which the student might be simultaneously enrolled). In this case, the response option, *1-3* would become 5-15 hours, and the next response option, *4-6* would become 20-30 hours. On the NSSE survey, among Georgia Southern University Seniors, 50% reported studying between 6-15 hours/week, and 28% reported studying between 16-30 hours/week. These numbers are virtually identical to the data obtained on the pilot SRI.

Section III: The course. This section contained two items.

Q17: How difficult was the subject matter of this course? (N = 2,411). Question 17 asked about the difficulty of the subject matter of the course. Response options were A = *Very difficult*, B = *Somewhat difficult*, C = *Neither difficult nor easy*, D = *Somewhat easy*, and E = *Very easy*.

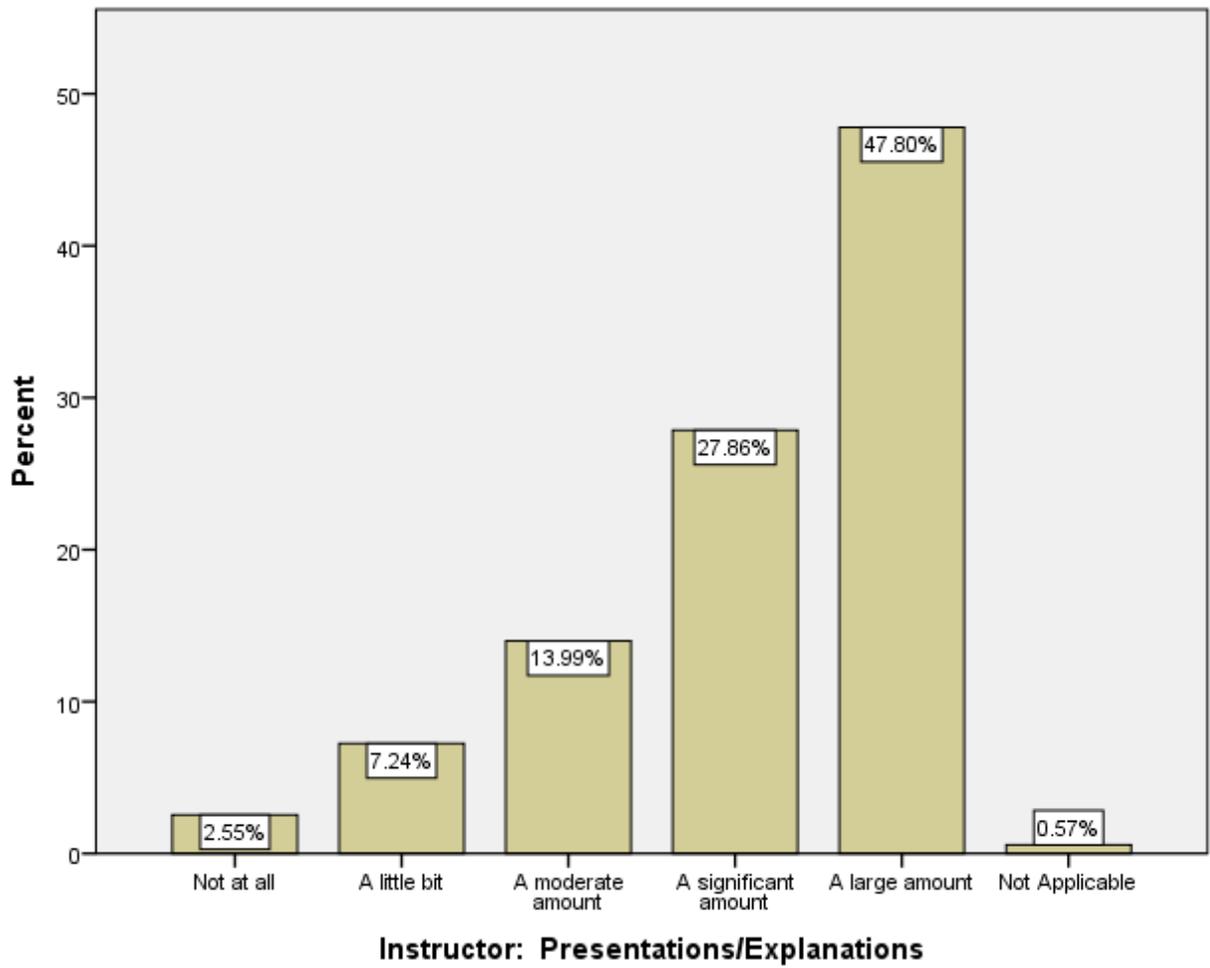


Q18: How interesting was the subject matter of this course? (N = 2,470). Question 18 asked about student interest in the course. Response options were A = *Very interesting*, B = *Somewhat interesting*, C = *Neither interesting nor uninteresting*, D = *Somewhat uninteresting*, and E = *Very uninteresting*. This item was reverse-scored.

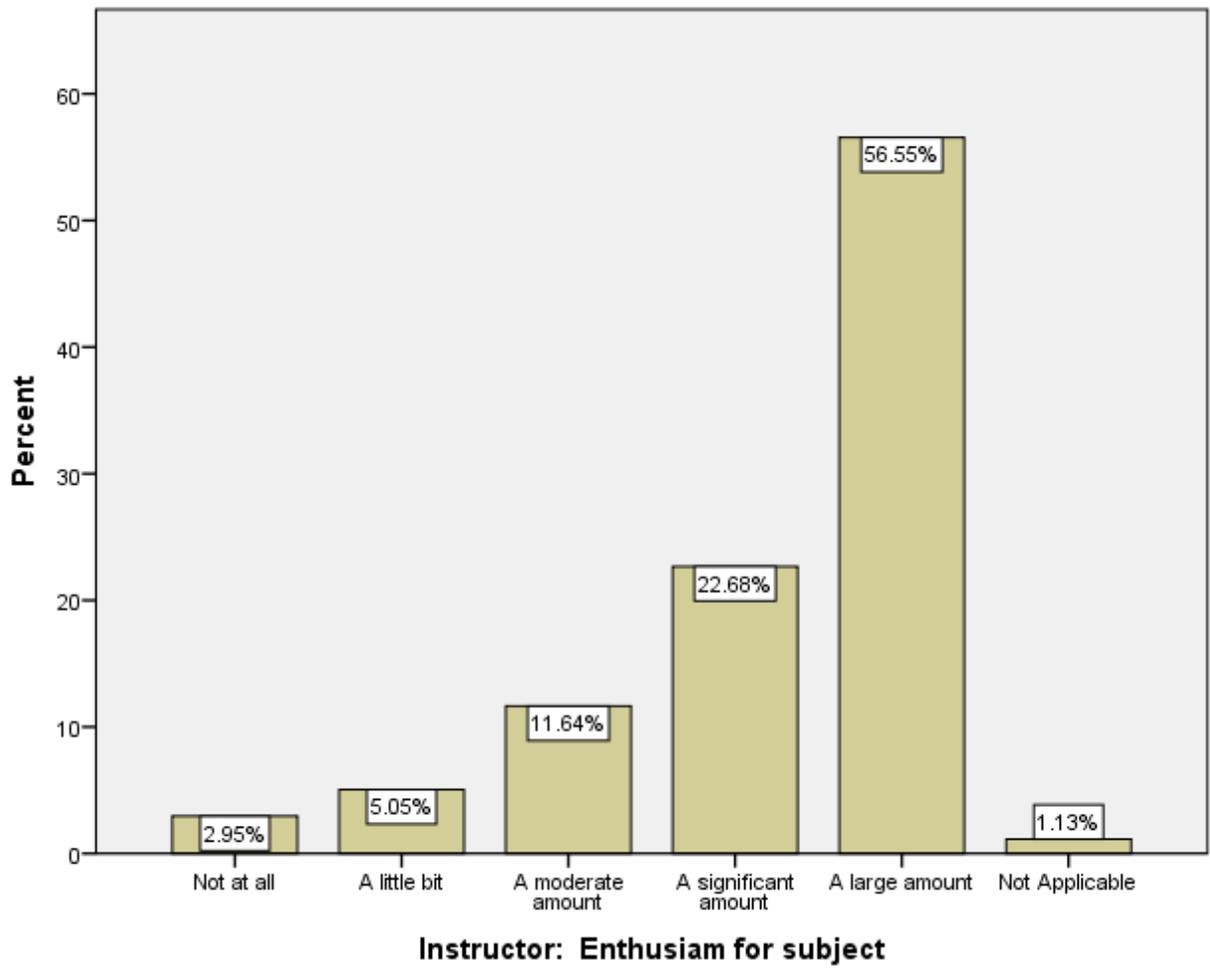


Section IV: The instructor. This section contained four items. The instructions read, “In this section, we will ask you about specific behaviors of your professor that are connected to student learning. Using the following scale, indicate how much each aspect of the professor’s approach influenced your learning.” Response options were 1 = *Not at all*, 2 = *A little bit*, 3 = *A moderate amount*, 4 = *A significant amount*, 5 = *A large amount*, N = *Not Applicable*. The sixth scale option was labeled “N” instead of “6” to clearly indicate it was not a value on the ordinal scale.

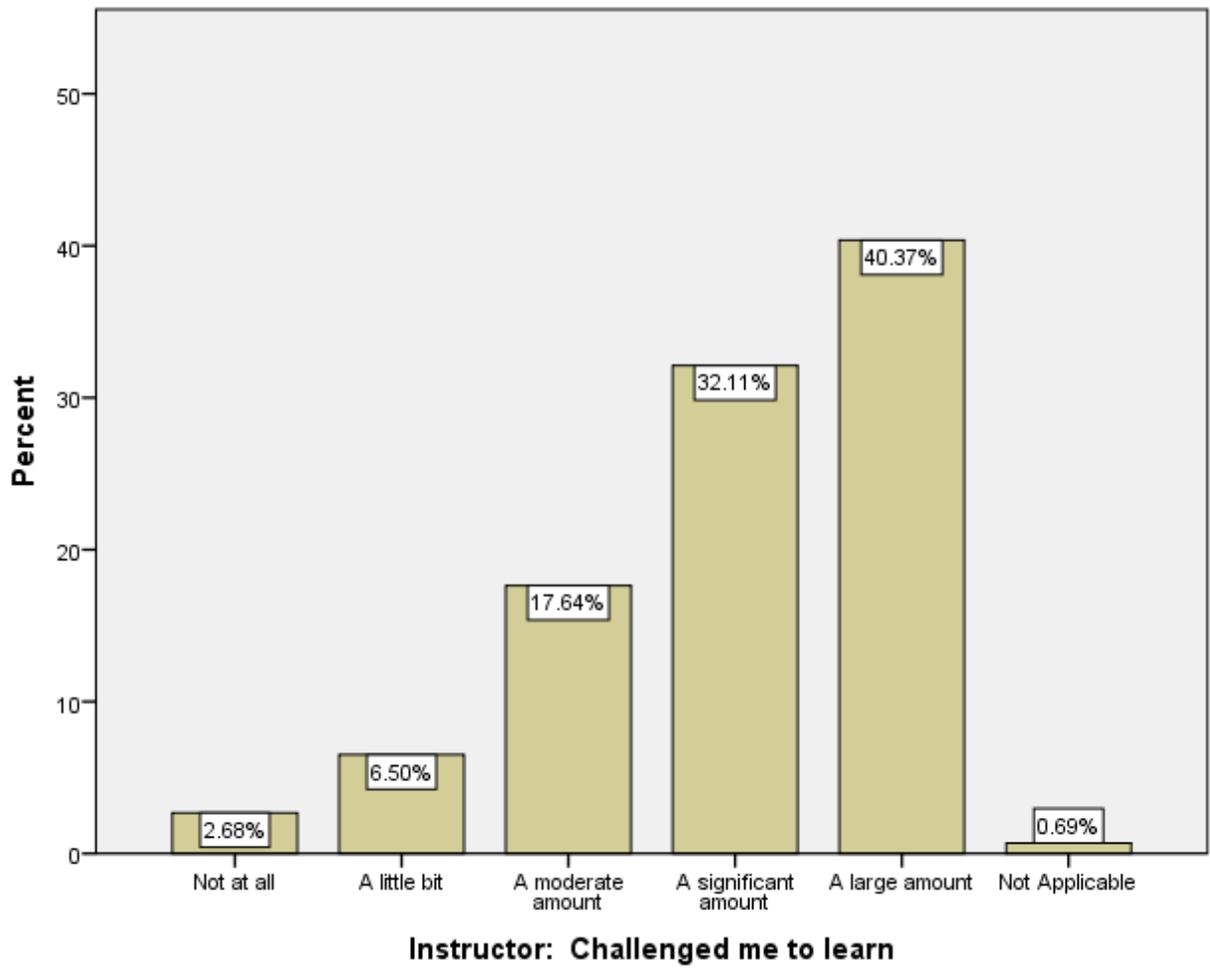
Q19: The professor's presentations/explanations: (N = 2,473).



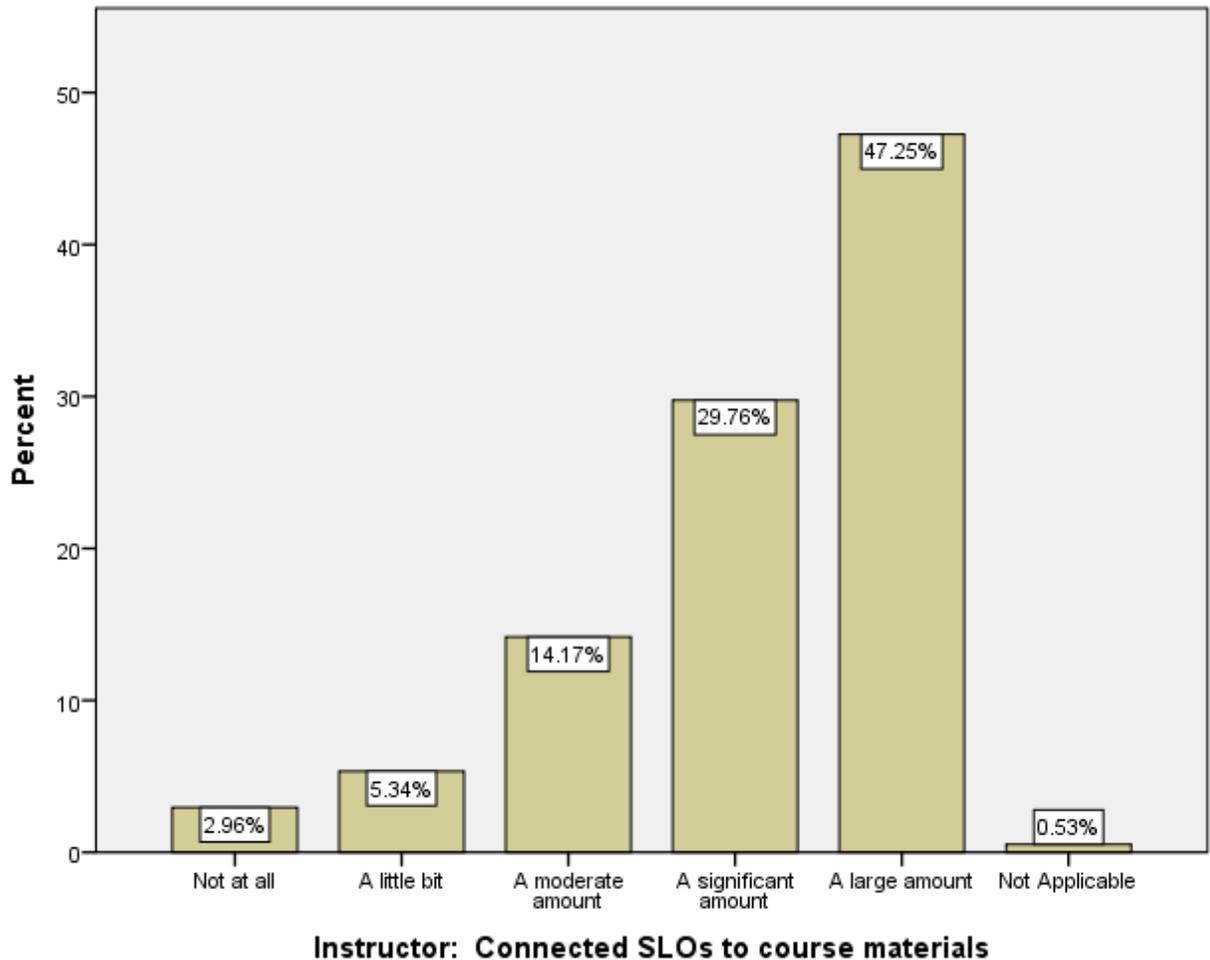
Q20: The professor's level of enthusiasm for the subject: (N = 2,474).



Q21: The level at which this professor challenged me to learn: (N = 2,460).



Q22: The extent to which this professor connected the student learning objectives to the class activities, assignments, and assessments: (N = 2,470).



Analysis. For the professor challenged me to learn question (Q21), NSSE data are again available for comparison. One NSSE item asks, “During the current school year, to what extent have your courses challenged you to do your best work?” Response options are 1 = *Not at all* to 7 = *Very much*. Although the wording of the questions is slightly different, and the NSSE item asks about all courses combined, both questions tap the same underlying concept. Among Georgia Southern University Seniors, 60% selected responses 6 or 7, which is quite similar to the percentage of students who selected *A significant amount* or *A large amount* on the pilot SRI.

Section VI: Results of the Feedback Survey of Faculty and Department Chairs

Overview

The feedback from both the Spring and Fall pilot testing was combined for faculty and chairs respectively, and results were analyzed for emergent themes for each group. This method of analysis is identical to the one used by the 2013-2014 ad hoc Committee on Student Ratings of Instruction to analyze the results of their survey. Each comment was independently coded by at least two members of the Committee. In the event of disagreement between coders, which was rare, the Committee Chair made the final decision.

Information on the sample and response rate is presented first. Results are presented by questionnaire version; first, the faculty version, next, the department chair version. Within each section, a general summary of results is presented first, followed by themes and exemplar quotations for each open-ended question in the questionnaire. Because some responses contained multiple ideas, some quotations below reflect only part of a respondent's answer to a particular question. Each response was independently coded for each theme, so a response could have been coded with more than one theme.

Both questionnaires provided brief introductory information about the pilot test and the over-arching goals of the pilot SRI instrument. For the Spring 2016 online pilot test, both questionnaires contained three open-ended questions (Questions #3, #4, and #5). After reviewing the feedback from the Spring 2016 online pilot test, the Committee added two additional yes/no questions to the questionnaire for the Fall 2016 face-to-face pilot test (Questions #1 and #2). See Appendices B and C.

In reviewing the responses below, it is important to remember that the data from the Spring 2016 online pilot SRI was sent by the Committee to the corresponding faculty and department chairs on May 9th, 2016, just after the conclusion of the Spring semester, so that they could review the data before the Qualtrics questionnaire closed on June 3rd, 2016. In contrast, data from the Fall 2016 face-to-face pilot test was processed by each department separately. Selected department chairs were informed that the data needed to be returned to the Committee, and the affected faculty members, no later than the first day of the Spring 2017 semester, Monday, January 9th, 2017, so that the selected faculty could review the data in time to respond to the Qualtrics questionnaire before it closed on Friday, January 20th, 2017.

Sample and Response Rate

Table 6. Sample and Response Rate by Collection Wave

Collection Wave	Total Sample	Number with Usable SRI Data	Responses	Response Rate
<u>Spring 2016: Online Courses</u>				
Faculty	35	30	11	37%
Chairs	20	18	7 ¹²	39%

¹² Two additional chairs wrote only that they could not answer the questions, so were dropped from the sample and not included in this number.

<u>Fall 2016: Face-to-Face Courses</u>				
Faculty	97	70	15 ¹³	21%
Chairs	32	21	8 ¹⁴	38%
<u>Combined</u>				
Faculty	132	100	26	26%
Chairs	36 ¹⁵	28	8-15 ¹⁶	29-54%

Given that this pool of responses for both faculty and chairs is very small and represents less than half of all individuals selected for the pilot test for both faculty and chairs, the Committee cautions the reader to **take care not to extrapolate or generalize from the summary of feedback below to the broader population of faculty and chairs** (i.e., the feedback is from only a tiny fraction of faculty and chairs and cannot be interpreted to represent the attitudes or opinions of other faculty and chairs).

Faculty Feedback

General summary. Results from the 26 faculty respondents provided mixed reviews of the proposed SRI form. Most faculty were able to identify one or more specific elements on the proposed form that they thought would be helpful in improving/enhancing student learning. Some faculty reported that the feedback from the proposed form would be more helpful in improving/enhancing student learning than the current SRI; other faculty reported that they preferred the current SRI, though mostly for reasons that were unrelated to improving/enhancing student learning (e.g., the lack of an omnibus instructor rating, the need to be able to compare faculty means). Unfortunately, many of the respondents did not answer *why* they felt the way they did or provide explanations or examples to clarify their answers. This limited the utility of the feedback to the Committee, especially given the low response rate.

Q1: Has your department chair shared the pilot SRI data from your selected course with you?

For the Fall 2016 respondents, 100% answered “Yes” to this question.

Q2: Have you examined the pilot SRI data for your selected course?

For the Fall 2016 respondents, 100% answered “Yes” to this question.

¹³ Seven faculty answered only the two yes/no questions, but did not provide any written responses, and thus had no data to include in the analyses, so were dropped from the sample and not included in this number. Another two faculty wrote answers indicating they had not reviewed the instrument or gave similar answers and were also dropped from the sample.

¹⁴ Three chairs answered only the two yes/no questions, but did not provide any written responses, and thus had no data to include in the analyses, so were dropped from the sample and not included in this number. Another three chairs wrote answers indicating they had not reviewed the instrument or gave similar answers and were also dropped from the sample.

¹⁵ Most department chairs had courses selected for both the online and face-to-face pilot testing, which is why the first two numbers in this row are not the sum of the numbers for each collection wave.

¹⁶ Because responses were anonymous, there is no way to know how many of the Spring and Fall respondents were the same individuals. Thus, the number of unique respondents could be as few as eight or as many as 15.

Q3: To what extent does the proposed new SRI form provide useful information for improving/enhancing student learning in your course? Which items did you find most valuable and why?

The following four general themes emerged from the questionnaires. Only two of 26 responses (8%) could not be classified into at least one of the four themes.

Theme 1 (N = 11, 42%). The focus on the SLOs on the form was valuable for improving/enhancing student learning.

“Having the SLO on the SRI provides more detail on student success and gives me feedback regarding their ability to grasp and learn the concepts. Low scores on the SLOs indicates that I may need to change my teaching style and assessment to ensure expectations are clear and content is covered.”

“I found the new element Section 1: Student learning objectives very helpful. It enables students to assess a range of learning objectives in a specific course, which consequently facilitates their reflection on their learning experience. Meanwhile, the instructor becomes more aware of the important role of the learning objectives and tends to make the connection between objectives and class activities/assessment more explicit for students.”

“Section 1 tied back directly to the learning objectives and answered whether or not I accomplished the course objectives.”

Theme 2 (N = 5, 19%). The focus on the Instructor items on the form was valuable for improving/enhancing student learning.

“I did find the section providing feedback on the teaching performance of the professor to be extremely helpful.”

“I thought the student answers to the question of how much I challenged them to learn was interesting and valuable information.”

“Question 22 was probably the most helpful, since it gave me an indication of how well the course helps students achieve the learning objectives. Given the increased emphasis the university is placing on assessment, it's important that the learning objectives are achieved.”

Theme 3 (N = 2, 8%). The focus on the open-ended questions/comments section on the form was valuable for improving/enhancing student learning.

“I found the open-ended items very valuable.”

“Comments section most meaningful.”

Theme 4 (N = 8, 31%). Criticisms or suggestions for improvement.

“I think it provides little useful information in enhancing student learning”

“I don't think it is much different than the existing form.”

“I have to admit that I found this form to be neither appreciably more useful nor appreciably less useful to evaluation student learning and the quality of teaching. The major problem is where the SRI, as an instrument, sits in the wider ensemble

of evaluating effective teaching. It seems to me that, whether such changes emanate from the departmental, college, or university level, there should be more specific policies specifying best practices for how SRIs can be used in concert with other equally revealing measures of teaching. Such things include, but are not limited to internal faculty peer evaluation, development of new courses and pedagogical initiatives, participation or leadership in teaching panels or professional development, and even grading distributions. While SRIs are useful to a point, no form I have seen has yet to circumvent the problems uncovered by current pedagogical research suggesting that they are vulnerable to gender, racial, and other biases that have little to do with the extent of student learning or the quality and efficacy of faculty teaching. At best, SRIs are far better at measuring a student's subjective perception of what they learned, not the skills and content that they actually learned. In short, and within the spectrum of what SRIs are useful for, I find this form to be no better but not appreciably worse than the earlier forms.”

Q4: How does the feedback from this form compare to feedback from the existing SRI form? Which form do you think provides more useful information for improving/enhancing student learning in your course and why?

The following two general themes emerged from the questionnaires. Only eight of 26 responses (31%) could not be classified into at least one of the two themes, and the majority of those responses did not provide a direct answer (e.g., “See my previous response.”).

Theme 1 (N = 11, 42%). The feedback from the proposed SRI form provides **more** useful information for improving/enhancing student learning than the feedback from the current SRI form.

“Overall, I like the new form and find the information useful b/c the SLOs are now on the form. Understanding from the students perspective if they achieved excellence in the SLOs is what matters the most to me.”

“This instrument seems less influenced by how entertaining a class is.”

“I prefer the proposed new form, which is concise and neat. It goes beyond "assessment of learning," but aims for "assessment for learning.””

Theme 2 (N = 7, 23%). The feedback from the proposed SRI form provides **less** useful information for improving/enhancing student learning than the feedback from the current SRI form.

“I prefer the older form.”

“Ultimately, whether its right or wrong, these SRIs will be used to compare one faculty to another. Thus, we need a standardized measure.”

“I think a hybrid between the forms would be good. Instructors need some form of overall rating for them or the course.”

It should be noted that many of the responses classified into this theme did not actually answer the question asked. That is, criticism of the items on the proposed SRI form often did not appear to be based on the utility of those items for improving/enhancing *student learning*, but rather for other reasons (e.g., the lack of an omnibus instructor rating).

Q5: If you have any other feedback that you would like to provide the committee about the questions on the proposed new SRI form, please use the space below.

The following two general themes emerged from the questionnaires. Ten of 26 responses (38%) could not be classified into at least one of the two themes, and the majority of those responses either left the question blank or did not provide a direct answer (e.g., “I have none to offer.”).

Theme 1 (N = 14, 54%). Criticisms, suggestions for improvement, or issues for the Committee to consider in carrying out its charge.

"I definitely believe this new proposed form is better. However, to be frank, at the end of the meeting with my chair the discussion was still focused on mean scores. Mind you; mean scores off of Likert-type scale ordinal data. It's frustrating to hear, "Well let's talk about these weak areas. You received a 4.4 here. How might you improve?""

“I am unclear why adding the question regarding how many class meetings a student missed is important. While attendance for my class is mandatory, I treat students as adults and hold them accountable for missing class. If they still choose to miss, then that is a decision they made understanding the consequences. What value does this add to the SRI?”

“I am having difficulties with items that do not apply. This past semester students in another course rated me 3 (middle of the choices since there was no NA option) on things that were not even a part of the course and a 3 is NOT a good score to have listed on an annual evaluation. I am very concerned about that. Even on this document in pilot mode, one student evaluated 5 objectives that should have been NA. \Also, a faculty member should have the option of excluding any evaluation of any course that does not have at least a 50% response rate. This pilot of my course wasn't bad but only 6 of 37 students completed it. What can we do to entice the students to complete such an important evaluation?”

Theme 2 (N = 4, 15%). Support for the Committee’s work/proposed measure.
“I move that we adopt the new form across campus!”

“Thanks for your hard work putting the new form together.”

“Both the old form and the new SRI have very good questions but the new SRI is more direct.”

Department Chair Feedback

General summary. Results from the 8-15 department chair respondents also provided mixed reviews of the proposed SRI form. Approximately half of the chairs were able to identify one or more specific elements on the proposed form that they thought would be helpful in

improving/enhancing student learning. Criticisms of the proposed SRI form typically fell into two categories: a) complaints that specific items on the current SRI form were not on the proposed SRI form, and b) concerns about the logistical and administrative burden of individualizing each SRI for each course. With respect to the first category of criticisms, it should be noted that many of the items that these department chairs appeared to want on the SRI form were in direct conflict with the best practices for SRIs reviewed in Section II and would not contribute any information towards improving/enhancing student learning (Penny, 2003).

Q1: Have you examined the pilot SRI data for the course(s) selected in your department?

For the Fall 2016 respondents, 100% answered “Yes” to this question.

Q2: Have you shared the pilot SRI data with the instructor(s) for those course(s)?

For the Fall 2016 respondents, 75% answered “Yes” and 25% answered “No” to this question. This may explain why the faculty response rate for the Fall 2016 Pilot Feedback Questionnaire was lower than the Spring 2016 Pilot Feedback Questionnaire (where the ad hoc Committee shared the data with all course instructors on May 9th, 2016).

Q3: To what extent does the proposed new SRI form provide useful information for helping you to assist faculty with improving their teaching (and student learning)? Which items did you find most valuable and why?

The following three general themes emerged from the questionnaires. All responses could be coded into at least one theme.

Theme 1 (N = 4, 27%). The focus on the open-ended questions/comments section on the form was valuable for improving/enhancing student learning.

“I find student comments more valuable than any of the Likert scale questions.”

“that focus [on student learning] in the narrative comments is also helpful.”

“By far the most useful question was the first open-ended one.”

Theme 2 (N = 5, 33%). Other specific items/sections on the form were valuable for improving/enhancing student learning.

“I believe the proposed new SRI form provided useful information. I appreciated the varied formats used to relay the information.”

“The standard version of the SRI uses the first six questions to evaluate a course compared to other courses of similar credit value. I'm not sure there's much value in that. So specifically: question 11 on the pilot is a better question than number 1 on the standard one”

“I like that the SLOs seem to have their own question, and the explicit question about connecting SLOs to the assignments. SLOs are what the heart of course is about.”

Theme 3 (N = 12, 80%). Criticisms or suggestions for improvement.

“I did not find it valuable at all. The older form was more informative.”

“Much of It provides little useful information to assist faculty in improving their teaching. Students are not capable of providing meaningful responses to the learning outcome questions. Many of the remaining questions the instructor would already know the general answer to and would only be used for instructor evaluation, rather than self evaluation.”

“the administration of getting course objectives onto the form itself poses a logistical nightmare. What if a Department does not have specific objectives for multiple sections? Does each instructor put her/his own specific objectives onto the form? When? How? Similarly, the form does not seemingly take into account the specific milieu of online courses. Will a different form be used? And how do we get better response rate than the tiny response rate we get now?”

Q4: How does the feedback from this form compare to feedback from the existing SRI form in terms of helping you to assist faculty with improving their teaching (and student learning)? Which form do you think provides more useful information for helping you to assist faculty with improving their teaching (and student learning) and why?

The following two general themes emerged from the questionnaires. Only four of 15 responses (27%) could not be classified into at least one of the two themes, and the majority of those responses provided an ambiguous answer (e.g., “I don’t see much difference.”). Note: The references to the difficulty of reading and interpreting the feedback/output are solely found in the online sample and are an artifact of the method that was used by the Committee to return the data to the department chairs. Such issues would not persist if the form were officially adopted.

Theme 1 (N = 2, 13%). The feedback from the proposed SRI form provides **more** useful information for improving/enhancing student learning than the feedback from the current SRI form.

“I was able to provide more targeted feedback using the proposed form. The proposed items provided more of an opportunity to discuss instructor growth in teaching and to discuss best practices in the profession, given the inclusion of the SLO items. As such, I think the proposed form provides more useful information.”

“My sense is that the new form would be an improvement, although in this format, the output I received was difficult to read. I imagine this is already in the works, but some care needs to be put into creating output that is easily accessible for both faculty and chairs.”

Theme 2 (N = 9, 60%). The feedback from the proposed SRI form provides **less** useful information for improving/enhancing student learning than the feedback from the current SRI form.

“The older form gives more specific information about various facets of teaching. The new form seems to artificially inflate the scores. Since the major factor is the learning outcomes, I would think every teacher could do extremely well as long as their course is not poorly designed. Even an unskilled teacher can earn a high score for following their chosen learning outcomes. For questions 19-22, I find the difference between a significant amount and a large amount confusing as quantifiers. Added to this is the misery my ADA had to endure entering the learning outcomes. This would not be

logistically feasible for large numbers of faculty every semester. I prefer the older form which gave me more specific data that could help my faculty improve their teaching.”

“I prefer when student's provide an overall rating of the instructor. While this also tends to vary widely from one student to another, I find it gives more valuable information than most of the questions on the new form.”

“The feedback is difficult to read and to interpret. The questions do not provide useful information about an instructor's performance. In our currently accepted SRIs, what I find most useful are students written comments to a variety of questions about instructor performance. The absence of these questions and the narrative response makes the experimental instrument useless for me.”

Q5: If you have any other feedback that you would like to provide the committee about the questions on the proposed new SRI form, please use the space below.

One general theme emerged from the questionnaires. Four of 15 responses (27%) were blank and could not be classified into a code.

Theme 1 (N = 11, 73%). Criticisms, suggestions for improvement, or issues for the Committee to consider in carrying out its charge.

“The general question about "how would you rate overall this instructor" on the current assessment, while overly simplistic, was the most useful single measure of instructor performance that could be referred to in recommendation letters and annual evaluations, and should not have been removed.”

“The excessive burden that the new SRI forms put on both staff and faculty (staff in particular) is not acceptable. Our department will need additional funding for administrative assistants to make this SRI transition a success.”

“In terms of semantics for items 1-10 and 19-22, I tend toward "A significant amount" following "A large amount," rather than the other way around. In terms of confidentiality, I was concerned that instructors might be able to discern handwriting on the written response section of the form. In terms of working around this, departments could have the comments typed separately to then distribute to instructors.”

Section VII: Recommendations¹⁷

In developing its final recommendations, the ad hoc Committee on Student Ratings of Instruction considered many factors, including but not limited to the Committee's charge (see Section I), BOR policy, the recommendations of the 2014 ad hoc Committee on SRIs, the research literature on SRIs, institutional logistical requirements and resources, best practices in questionnaire design, the results of the data analyses of the pilot test data, official feedback from faculty and department chairs who had participated in the pilot test, unofficial feedback from faculty, department chairs, and administrative assistants who had been involved in the pilot test, and unofficial feedback from faculty who had not been involved in the pilot test. The Committee makes three independent recommendations.

Recommendation #1

The Committee recommends that the Faculty Senate adopt the proposed new Student Ratings of Instruction measure found in Appendix D to replace the version currently in use. The new measure would be approved for use beginning in the Fall 2017 semester and extending no later than the Summer 2027 term. Unless the Faculty Senate takes prior relevant action before the end of that ten year period, at the beginning of the Fall 2027 semester the Faculty Senate will revisit the issue and may either vote to extend approval of the measure to another date not more than 10 years in the future, or vote to temporarily extend approval of the measure and appoint a new ad hoc committee to investigate revising or replacing the measure.

Rationale. The proposed new SRI measure represents best practices in the published research literature on Student Ratings of Instruction (see Section II: Background), is consistent with Board of Regents guidelines that Student Ratings of Instruction should be used primarily to inform and improve teaching effectiveness, the end goal of which is student learning, and is focused on student learning, learning behaviors, and formative feedback, as the Committee was charged to create.

The proposed measure differs from the pilot measure in one major and three minor ways. The major difference is that Section I: Student Learning Objectives has been removed from the form and replaced in what is now Section IV: Comments on the Learning Experience by a new open-ended question. This was a difficult decision for the Committee because it meant deleting the section of the measure that provided the most course-specific feedback. As noted in Section II of this report, SET questions that focus on what students learn in the course are a recommended best practice in the literature and are consistent with Georgia Southern University Faculty Handbook Section 205.01's description of superior teaching as "focused on student learning outcomes." However, during pilot testing it became apparent that the Remark Office software that is used to process paper SRI forms for face-to-face courses is extremely sensitive to even the smallest user error. In customizing the SRI form for each course, as is required for Section I of the pilot form, even tiny deviations in the form could render the form unreadable by the software. In such cases, either the responses on each individual form would have to be manually entered to be retrievable, or the data would have to be treated as "missing" for that course. The Committee was aware of at least five courses from three different departments in

¹⁷ Initially, this section was to be written after consultation with the Faculty Senate at their February 7th, 2017 meeting. However, the Faculty Senate Executive Committee declined to place the ad hoc Committee on the agenda for that meeting and informed the Committee, "There is no need for feedback from the entire Senate until you have submitted the final report." As a result, the Committee had to develop its recommendations without Faculty Senate feedback.

the pilot sample that had unreadable data because of such errors. We find that error rate unacceptably high for an SRI form. Although the frequency of such user errors could be reduced by more thorough training on how to use the software for the departmental personnel who process the forms, given the number of departmental personnel institution-wide who would need such training, the turnover rate among those personnel, and the sheer volume of forms that each person would be responsible for processing, the Committee believes that the possibility for catastrophic error (e.g., all of a department's SRIs for a semester are unreadable) is too high to recommend any form which requires customization. As a result, we have removed the customizable Section I: Student Learning Objectives from the form.

The Committee also notes that Section I: Student Learning Objectives, was the source of the greatest amount of negative feedback, and especially legitimate negative feedback, from department chairs (both officially and unofficially). Multiple department chairs expressed concern about the amount of time that would be required to customize the pilot SRI measure for each course in their departments and stated that they would need additional funding and personnel to be able use the pilot SRI measure. The Committee defers to the judgment of these department chairs in matters of departmental budgets and resources and concludes that the institution does not currently have the resources to support a customizable SRI measure.

As a replacement for the Section I: Student Learning Objectives question, the Committee has added a third open-ended question to the beginning of what is now Section IV: Comments on the Learning Experience. This question explicitly asks students to reflect on what they have learned in the course and what their instructor did to facilitate their learning. Although this question is not directly focused on Student Learning Objectives, it does still directly ask about student learning in the course. The Committee believes the addition of this question represents the best possible compromise to inquiring about student learning in course-specific ways but without requiring a customizable SRI measure. The addition of a third open-ended question is also consistent with the feedback the Committee received (and which is echoed in the 2013-2014 ad hoc Committee on SRI's report) that many department chairs and faculty find student responses to open-ended questions to be the most valuable element of SRI measures.

The proposed measure also contains three minor revisions to what is now Section III: The Instructor. First, in response to a specific suggestion, the Committee has added the word "positively" to the instructions before the words, "influenced your learning" so that it is clear that the question is asking how the instructor *positively* influenced a student's learning. Second, the Committee has deleted the phrase, "Using the following scale," from the instructions. It was necessary to delete some text to offset the inclusion of the word "positively" to the instructions so that Section III would not spill over to the second page. Third, the Committee has relabeled the fourth and fifth response options from "A significant amount" and "A large amount" to "A large amount" and "A very large amount," respectively, in response to specific feedback about the response scale.

Recommendation #2

The Committee moves to amend the Faculty Handbook as follows:

1) Section 205.01 Criteria for All Types of Faculty Evaluation, second paragraph under Teaching, from:

Student ratings of instruction shall not be the sole measure of teaching effectiveness for any review, nor shall instructors be ranked according to student ratings for evaluation; rather, a

complete picture should be obtained through multiple sources. Documentation of teaching effectiveness is the responsibility of the faculty member.

To:

Student ratings of instruction shall not be the sole or primary measure of teaching effectiveness for any review at any level, nor shall instructors' ratings be compared to other instructors' (e.g., department means) or to specific "cut points" (e.g., 4.0 on a 5.0 scale) as part of their evaluation; rather, a complete picture should be obtained through multiple sources. The numerical data reported for each course shall consist of the frequency distribution of scores for each question, the number of responders, and the response rate for the course; measures of central tendency and variability (e.g., means and standard deviations) shall not be used to evaluate instructors at any level. Results should not be generalized beyond the students who responded, especially when the response rate is low. In evaluations, written comments from student ratings of instruction that are not about teaching effectiveness should be disregarded. Comments that are about teaching effectiveness should be evaluated cautiously in the context of the course. Research on student ratings of instruction and potentially biasing influences should also be taken into consideration in any use of student ratings of instruction data. Department chairs and other personnel who formally evaluate instructors' teaching effectiveness by means of student ratings of instruction data shall receive sound methodological training and regular briefing on the major findings in the research literature on how to appropriately use such data. Documentation of teaching effectiveness is the responsibility of the faculty member.

2) Section 205.06 Procedures for Faculty Evaluations, Section E.3, from:

Regents policy requires that a written system of student ratings of instruction be utilized in the annual evaluation of each faculty member (*Board of Regents Policy Manual*, § 8.3.5). Completed rating forms are kept on file in the department chair's office and are the property of the University.

To:

Regents policy requires that a written system of student ratings of instruction be utilized in the annual evaluation of each faculty member (*Board of Regents Policy Manual*, § 8.3.5). The use of student ratings of instruction data will be primarily formative, with the main goal of improving teaching effectiveness. Completed rating forms are kept on file in the department chair's office and are the property of the University.

3) Section 205.07 Student Ratings of Instruction, from:

Georgia Southern requires and conducts written or online student ratings of instruction each academic term (excluding summer) to provide information to faculty for their use in the improvement of teaching. Results are also used in faculty evaluation as mandated by Regents policy as a portion of an evaluation of teaching effectiveness. Department chairs return a summary of numerical results and students' written comments to faculty each academic term;

original responses are the property of the University. Courses shall be evaluated by students in the same manner as the course is conducted. Partially online courses whose content is offered 50% or more online are evaluated through CoursEval. As with any evaluation, faculty shall have the right to respond to student ratings regarding factors that might have influenced student ratings of instruction scores.

To:

Georgia Southern requires and conducts written or online student ratings of instruction each academic term (excluding summer) primarily to provide information to faculty for their use in the improvement of teaching. Results are also used in faculty evaluation as mandated by Regents policy as a portion of an evaluation of teaching effectiveness. Each academic term, department chairs return to faculty a numerical report on the frequency distribution of scores, the number of responders, and the response rate for each course, in addition to a typed copy of students' written comments for each course. At each annual review, department chairs discuss the results with each faculty member with the main goal of improving teaching effectiveness. Original responses are the property of the University. Courses shall be evaluated by students in the same manner as the course is conducted. Partially online courses whose content is offered 50% or more online are evaluated through CoursEval. As with any evaluation, faculty shall have the right to respond in writing to student ratings regarding factors that might have influenced student ratings of instruction scores. These responses shall be permanently appended to any future reports of that student ratings of instruction data.

Rationale. The Committee was charged to propose methods to make the evaluation of teaching effectiveness more equitable and consistently defined, assessed, and used across the university including 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. These changes reflect best practices in the use of Student Ratings of Instruction from the published research literature on the topic. For a more detailed rationale in support of these changes, see Section II: Background in this report.

Recommendation #3

The Committee moves that the Faculty Senate request that the Consolidation Operational Working Group that is tasked with the relevant faculty evaluation processes explore the possibility of centralizing the Student Ratings of Instruction process into one Office or Office per campus, including: printing, disseminating, administering, collecting, analyzing, and returning results to faculty and department chairs.

Rationale. During the course of the pilot testing of the proposed new Student Ratings of Instruction instrument, the Committee repeatedly received feedback from department chairs and administrative assistants that the existing SRI administration process was already burdensome enough to their departments, requiring significant amounts of time each semester. These individuals lamented having to spend so much time on the SRI process instead of other important departmental tasks. In addition, a decentralized approach results in both inefficiencies and inconsistencies, which would only be exacerbated by multiple campuses. A decentralized approach is inefficient because large numbers of individuals need to be continuously trained in how to process SRIs, and even the smallest change requires re-training of all those individuals. If a central office were responsible for the SRI process, a much smaller number of individuals would need to be trained. A decentralized approach is inconsistent because the process is open

to interpretation and error by every department chair and administrative assistant instead of a small number of individuals in a central office. For a university of Georgia Southern's size, a centralized SRI process is likely to be far more efficient and consistent than our current decentralized model.

Section VIII: References

- 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.
- Ambrose, S. A., Bridges, M. W., DiPietro, M., Lovett, M. C., & Norman, M. K. (2010). *How learning works: 7 research-based principles for smart teaching*. San Francisco, CA: Jossey-Bass.
- American Educational Research Association. (2013). *Rethinking faculty evaluation: AERA report and recommendations on evaluating education research, scholarship, and teaching in postsecondary education*. Washington, DC: Author. Retrieved from: <http://www.aera.net/Newsroom/NewsReleasesandStatements/AERAIssuesRecommendationsforRethinkingFacultyEvaluationforthe21stCentury/tabid/15254/Default.aspx>
- Angelo, T. (2015, October). *Doing assessment as if teaching and learning matter*. Presentation at the Assessment Institute in Indianapolis, Indianapolis, IN.
- Artino, A. R., Jr., La Rochelle, J. S., Dezee, K. J., & Gehlbach, H. (2014). Developing questionnaires for educational research: AMEE guide no. 87. *Medical Teacher*, *36*, 463–474. doi: 10.3109/0142159X.2014.889814
- Arum, R., & Roksa, J. (2011). Limited learning on college campuses. *Society*, *48*, 203–207. doi: 10.1007/s12115-011-9417-8
- Baker, P., & Copp, M. (1997). Gender matters most: The interaction of gendered expectations, feminist course content, and pregnancy in student course evaluations. *Teaching Sociology* *25*, 29–43. doi: 10.2307/1319109
- Bain, K. (2004). *What the best college teachers do*. Cambridge, MA: Harvard University Press.
- Boring, A., Ottoboni, K., & Stark, P. B. (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. *ScienceOpen Research*. doi: 10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1
- Boysen, G. A. (2014). Significant interpretation of small mean differences in student evaluations of teaching despite explicit warning to avoid overinterpretation. *Scholarship of Teaching and Learning in Psychology*, *1*, 150–162. doi: [10.1037/stl0000017](https://doi.org/10.1037/stl0000017)
- Boysen, G. A., Kelly, T. J., Raesly, H. N., & Casner, R. W. (2014) The (mis)interpretation of teaching evaluations by college faculty and administrators. *Assessment & Evaluation in Higher Education*, *39*, 641–656, doi: 10.1080/02602938.2013.860950
- Braga, M., Paccagnella, M., & Pellizzari, M. (2014). Evaluating students' evaluations of professors. *Economics of Education Review*, *41*, 71–88. doi: 10.1016/j.econedurev.2014.04.002
- 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. doi: 10.1086/653808
- 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.
- Cashin, W. E. (1990). Students do rate different academic fields differently. In M. Theall & J. Franklin (Eds.), *Student ratings of instruction: Issues for improving practice*. San Francisco, CA: Jossey-Bass Inc.

- Cashin, W. E., & Clegg, V. L. (1987). *Are student ratings of different academic fields different?* Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.
- Cohen, P. A. (1981). Student ratings of instruction and student achievements: A meta-analysis of multisection validity studies. *Review of Educational Research*, 51, 281–309. doi: 10.3102/00346543051003281
- Cranton, P. A., & Smith, R. A. (1986). A new look at the effect of course characteristics on student ratings of instruction. *American Educational Research Journal*, 23, 117–128. doi: 10.3102/00028312023001117
- d'Apollonia, S., & Abrami, P. C. (1997). Scaling the ivory tower, Part I: Collecting evidence of instructor effectiveness. *Psychology Teaching Review*, 6, 46–59.
- Davison, E., & Price, J. (2009). How do we rate? An evaluation of online student evaluations. *Assessment & Evaluation in Higher Education*, 34, 51–65. doi: 10.1080/02602930801895695
- Delucchi, M., & Smith, W. L. (1997a). A postmodern explanation of student consumerism in higher education. *Teaching Sociology*, 25, 322–327. doi: 10.2307/1319301
- Delucchi, M., & Smith, W. L. (1997b). Satisfied customers versus pedagogic responsibility: Further thoughts on student consumerism. *Teaching Sociology*, 25, 336–337. doi: 10.2307/1319304
- DeVellis, R. F. (2016). *Scale development: Theory and applications* (4th ed.). Newbury Park, CA: Sage.
- Feldman, K. A. (1978). Course characteristics and college students' ratings of their teachers: What we know and what we don't know. *Research in Higher Education*, 9, 199–242. doi: 10.1007/BF00976997
- Flaherty, C. (2015, June 10). Flawed evaluations. *Inside Higher Education*. Retrieved from <https://www.insidehighered.com/news/2015/06/10/aaup-committee-survey-data-raise-questions-effectiveness-student-teaching>
- Fowler, F. J. (2013). *Survey research methods* (5th ed.). Thousand Oaks, CA: Sage.
- Franklin, J. (2001). Interpreting the numbers: Using a narrative to help others read student evaluations of your teaching accurately. *New Directions for Teaching and Learning*, 87, 85–100. doi:10.1002/tl.10001
- Georgia Southern University. (n.d.). *NSSE 2015: Frequencies and statistical comparisons*. Retrieved from <http://em.georgiasouthern.edu/osra/nsse/>
- Georgia Southern University Faculty Senate. (2001, February 12th). *Georgia Southern University Faculty Senate Meeting*. Retrieved from: <http://digitalcommons.georgiasouthern.edu/faculty-senate-minutes/9/>
- Gray, M., & Bergmann, B. R. (2003). Student teaching evaluations: Inaccurate, demeaning, misused. *Academe*, 89(5), 44–46. doi: 10.2307/40253388
- Griffin, B. W. (January 24, 2001). *Pilot study of new student ratings of instruction instrument: Analysis of scores for evidence of validity*. Statesboro, GA: Georgia Southern University, Student Ratings Committee. Retrieved from: www.bwgriffin.com/gsu/courses/edur9131/content/PilotStudyStudentRatings.pdf
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London, United Kingdom: Routledge.

- 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, United Kingdom: Open University Press.
- Higher Education Research Institute. (2000). *The American freshman: National norms for fall 1999*. Los Angeles, CA: University of California, Los Angeles.
- 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.
- Jaschick, S., & Lederman, D., Eds. (2017). *2017 survey of college and university chief academic officers: A study by Inside Higher Ed and Gallup*. Retrieved from: <https://www.insidehighered.com/news/survey/2017-inside-higher-ed-survey-chief-academic-officers>
- 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. doi: 10.1080/03075079912331380128
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one’s own incompetence lead to inflated self-awareness. *Journal of Personality and Social Psychology*, 77, 1121–1134. doi: 10.1037/0022-3514.77.6.1121
- 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.
- Lauer, C. (2012). A comparison of faculty and student perspectives on course evaluation terminology. In J. Groccia & L. Cruz (Eds.), *To improve the academy: Resources for faculty, instructional, and organizational development* (pp. 195–212). San Francisco, CA: Wiley & Sons, Inc.
- McCoach, D. B., Gable, R. K., & Madura, J. P. (2013). *Instrument development in the affective domain: School and corporate applications* (3rd ed.). New York: Springer.
- McKeachie, W. (1997). Student ratings: The validity of use. *American Psychologist*, 52, 1218–1225. doi: 10.1037/0003-066X.52.11.1218
- 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.
- Menges, R. J. (2000). Shortcomings of research on evaluating and improving teaching in higher education. *New Directions for Teaching and Learning*, 83, 5–11. doi:10.1002/tl.8301
- Pascarella, E. T., Blaich, C., Martin, G. L., & Hanson, J. M. (2011). How robust are the findings of *Academically Adrift?* *Change*, 43, 20–24. doi: 10.1080/00091383.2011.568898
- Penny, A. R. (2003). Changing the agenda for research into students’ views about university teaching: Four shortcomings of SRT research. *Teaching in Higher Education*, 8, 399–411. doi:10.1080/13562510309396
- Pettit, E. (2016, June 16). How one professor is trying to paint a richer portrait of effective teaching. *The Chronicle of Higher Education*. Retrieved from <http://chronicle.com/article/How-One-Professor-Is-Trying-to/236827>
- Platt, M. (1993). What student evaluations teach. *Perspectives on Political Science*, 22, 29–40. doi:10.1080/10457097.1993.9944516
- Sedlmeier, P. (2006). The role of scales in student ratings. *Learning and Instruction*, 16, 401–415. doi:10.1016/j.learninstruc.2006.09.002

- Seldin, P. (2006). *Evaluating faculty performance: A practical guide to assessing teaching, research, and service*. Boston, MA: Anker Publishing Company, Inc.
- 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.
- Shepperd, J. W. (1997). Relevance and responsibility: A postmodern explanation of student consumerism in higher education. *Teaching Sociology*, 25, 333–335. doi: 10.2307/1319303
- 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>
- Smith, W. L. (2000). Teaching in a consumeristically charged environment. *Michigan Sociological Review*, 14 (Fall), 58–72.
- Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research*, 83, 598–642. doi: 10.3102/0034654313496870
- Stark, P. B., & Freishtat, R. (2014). An evaluation of course evaluations. *ScienceOpen*. doi: 10.14293/S2199-1006.1.SOR-EDU.AOFRQA.v1
- Titus, J. (2008). Student ratings in a consumerist academy: Leveraging pedagogical control and authority. *Sociological Perspectives*, 51, 397–422. doi: 10.1525/sop.2008.51.2.397
- 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>
- University System of Georgia. (2017). *Board of Regents policy manual: Official policies of the University System of Georgia*. Retrieved from: <http://www.usg.edu/policymanual>
- 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>
- Uttl, B., White, C. A., & Gonzalez, D. W. (in press). Meta-analysis of faculty’s teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*. <http://dx.doi.org/10.1016/j.stueduc.2016.08.007>
- Vasey, C., & Carroll, L. (2016). How do we evaluate teaching? *Academe*, 102(3), 34–39.
- Wachtel, H. K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment and Evaluation in Higher Education*, 23, 191–212. doi: 10.1080/0260293980230207
- Weinberg, B. A., Hashimoto, M., & Fleisher, B. M. (2009). Evaluating teaching in higher education. *Journal of Economic Education*, 40, 227–261. doi: 10.3200/JECE.40.3.227-261
- Wright, S. L., & Jenkins-Guarnieri, M. A. (2012). Student evaluations of teaching: Combining the meta-analyses and demonstrating further evidence for effective use. *Assessment & Evaluation in Higher Education*, 37, 683–699. doi: 10.1080/02602938.2011.563279

Yunker, P. J., & Yunker, J. A. (2003). Are student evaluations of teaching valid? Evidence from an analytical business core course. *Journal of Education for Business*, 78, 313–317. doi: 10.1080/08832320309598619

Appendix A: Pilot Student Ratings of Instruction Instrument

Instructor: _____

Course Prefix: _____ Number: _____ Section: _____

Course Reference #	<input type="text"/>									
	<input type="text"/>									
	<input type="text"/>									
	<input type="text"/>									
	<input type="text"/>									

IMPORTANT: Fill in Course Reference # to the left: enter number in the blocks and shade appropriate digits.

Section I: Student Learning Objectives

Indicate how much you learned on each student learning objective for this course with the following scale.

	Nothing	A little bit	A moderate amount	A significant amount	A large amount	Not Applicable
1.	<input type="radio"/>					
2.	<input type="radio"/>					
3.	<input type="radio"/>					
4.	<input type="radio"/>					
5.	<input type="radio"/>					
6.	<input type="radio"/>					
7.	<input type="radio"/>					
8.	<input type="radio"/>					
9.	<input type="radio"/>					
10.	<input type="radio"/>					

Section II: The Student

 Student learning depends in large part on **what the student does**. In this section, we will ask you questions about you that are connected to your learning in this course.

	Never	Rarely	Sometimes	Frequently	Always
11. How often were you self-motivated to learn this course material?	<input type="radio"/>				
12. Based on the professor's stated expectations, how often were you fully prepared for class?	<input type="radio"/>				
13. How often did you seek the professor's assistance?	<input type="radio"/>				

14. How many class meetings did you miss in this course?

- (A) None
- (B) 1-2
- (C) 3-4
- (D) 5-6
- (E) 7 or more
- (F) This was an online class

15. Approximately how many hours per week did you spend on this course outside of scheduled class time?

- (A) None
- (B) 1-3
- (C) 4-6
- (D) 7-9
- (E) 10-12
- (F) 13 or more

16. What grade do you think you will earn in this course?

- (A) (B) (C) (D) (E)

Section III: The Course

17. How difficult was the **subject matter** of this course?

- (A) Very difficult
- (B) Somewhat difficult
- (C) Neither difficult nor easy
- (D) Somewhat easy
- (E) Very easy

18. How interesting was the **subject matter** of this course?

- (A) Very interesting
- (B) Somewhat interesting
- (C) Neither interesting nor uninteresting
- (D) Somewhat uninteresting
- (E) Very uninteresting

Section IV: The Instructor

In this section, we will ask you about specific behaviors of your professor that are connected to student learning. Using the following scale, indicate how much each aspect of the professor's approach **influenced your learning**.

	Not at all	A little bit	A moderate amount	A significant amount	A large amount	Not Applicable
19. The professor's presentations/explanations:	(1)	(2)	(3)	(4)	(5)	(N)
20. The professor's level of enthusiasm for the subject:	(1)	(2)	(3)	(4)	(5)	(N)
21. The level at which this professor challenged me to learn: The extent to which this professor connected the	(1)	(2)	(3)	(4)	(5)	(N)
22. student learning objectives to the class activities, assignments, and assessments:	(1)	(2)	(3)	(4)	(5)	(N)

Section V: Comments on the Learning Experience

What aspect(s) of your class experience (course, professor, etc.) helped your learning most? **Provide specific, objective examples:**

What aspect(s) of your class experience (course, professor, etc.) could have been changed to help you learn more?: **Provide specific, objective examples:**

Appendix B: Faculty Questionnaire

Thank you for participating in the SRI Pilot Test. Because less than 5% of all courses were selected to participate in the pilot test, it is very important that we get feedback from the selected course instructors about the utility of the proposed SRI form and the information about teaching and learning that it generates. Please take a few minutes to give us your feedback.

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.” This BOR mandate, within which all USG institutions must operate, guided the development of the proposed new SRI form, which was designed to explicitly focus on collecting feedback from students to improve teaching effectiveness.

Q1 Has your department chair shared the pilot SRI data from your selected course with you?

- Yes (1)
- No (2)

Q2 Have you examined the pilot SRI data for your selected course?

- Yes (1)
- No (2)

Q3 To what extent does the proposed new SRI form provide useful information for improving/enhancing student learning in your course? Which items did you find most valuable and why?

Q4 How does the feedback from this form compare to feedback from the existing SRI form? Which form do you think provides more useful information for improving/enhancing student learning in your course and why?

Q5 If you have any other feedback that you would like to provide the committee about the questions on the proposed new SRI form, please use the space below.

Note: Questions #1 and #2 were included only on the Fall 2016 face-to-face courses questionnaire. The ad hoc Committee directly sent the data from the online pilot to affected instructors (rendering Q1 inappropriate for that sample). Q2 was added to the Fall 2016 questionnaire based on responses to the Spring 2016 questionnaire that had suggested some individuals had not reviewed the data before answering the questionnaire.

Appendix C: Department Chair Questionnaire

Thank you for participating in the SRI Pilot Test. Because less than 5% of all courses were selected to participate in the pilot test, it is very important that we get feedback from department chairs about the utility of the pilot SRI form and the information about teaching and learning that it generates. Please take a few minutes to give us your feedback.

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.” This BOR mandate, within which all USG institutions must operate, guided the development of the proposed new SRI form, which was designed to explicitly focus on collecting feedback from students to improve teaching effectiveness.

Q1 Have you examined the pilot SRI data for the course(s) selected in your department?

- Yes (1)
- No (2)

Q2 Have you shared the pilot SRI data with the instructor(s) for those course(s)?

- Yes (1)
- No (2)

Q3 To what extent does the proposed new SRI form provide useful information for helping you to assist faculty with improving their teaching (and student learning)? Which items did you find most valuable and why?

Q4 How does the feedback from this form compare to feedback from the existing SRI form in terms of helping you to assist faculty with improving their teaching (and student learning)? Which form do you think provides more useful information for helping you to assist faculty with improving their teaching (and student learning) and why?

Q5 If you have any other feedback that you would like to provide the committee about the questions on the proposed new SRI form, please use the space below.

Note: Questions #1 and #2 were included only on the Fall 2016 face-to-face courses questionnaire. The ad hoc Committee directly sent the data from the online pilot to affected instructors (rendering Q2 inappropriate for that sample). Q1 was added to the Fall 2016 questionnaire based on responses to the Spring 2016 questionnaire that had suggested some individuals had not reviewed the data before answering the questionnaire.

Appendix D: Recommended Student Ratings of Instruction Instrument

Instructor: _____

Course Prefix: _____ Number: _____ Section: _____

Course Reference #	0	1	2	3	4	5	6	7	8	9
	0	1	2	3	4	5	6	7	8	9
	0	1	2	3	4	5	6	7	8	9
	0	1	2	3	4	5	6	7	8	9
	0	1	2	3	4	5	6	7	8	9

IMPORTANT: Fill in Course Reference # to the left: enter number in the blocks and shade appropriate digits.

Section I: The Student

Student learning depends in large part on **what the student does**. In this section, we will ask you questions about you that are connected to your learning in this course.

- | | Never | Rarely | Sometimes | Frequently | Always |
|--|-------|--------|-----------|------------|--------|
| 1. How often were you self-motivated to learn this course material? | (1) | (2) | (3) | (4) | (5) |
| 2. Based on the professor's stated expectations, how often were you fully prepared for class? | (1) | (2) | (3) | (4) | (5) |
| 3. How often did you seek the professor's assistance? | (1) | (2) | (3) | (4) | (5) |
| 4. How many class meetings did you miss in this course? | | | | | |
| (A) None | | | | | |
| (B) 1-2 | | | | | |
| (C) 3-4 | | | | | |
| (D) 5-6 | | | | | |
| (E) 7 or more | | | | | |
| (F) This was an online class | | | | | |
| 5. Approximately how many hours per week did you spend on this course outside of scheduled class time? | | | | | |
| (A) None | | | | | |
| (B) 1-3 | | | | | |
| (C) 4-6 | | | | | |
| (D) 7-9 | | | | | |
| (E) 10-12 | | | | | |
| (F) 13 or more | | | | | |
| 6. What grade do you think you will earn in this course? | | | | | |
| | (A) | (B) | (C) | (D) | (E) |

Section II: The Course

- | | |
|---|--|
| 7. How difficult was the subject matter of this course? <ul style="list-style-type: none"> (A) Very difficult (B) Somewhat difficult (C) Neither difficult nor easy (D) Somewhat easy (E) Very easy | 8. How interesting was the subject matter of this course? <ul style="list-style-type: none"> (A) Very interesting (B) Somewhat interesting (C) Neither interesting nor uninteresting (D) Somewhat uninteresting (E) Very uninteresting |
|---|--|

Section III: The Instructor

In this section, we will ask you about specific behaviors of your professor that are connected to student learning. Indicate how much each aspect of the professor's approach **positively influenced your learning**.

- | | Not at all | A little bit | A moderate amount | A large amount | A very large amount | Not Applicable |
|---|------------|--------------|-------------------|----------------|---------------------|----------------|
| 9. The professor's presentations/explanations: | (1) | (2) | (3) | (4) | (5) | (N) |
| 10. The professor's level of enthusiasm for the subject: | (1) | (2) | (3) | (4) | (5) | (N) |
| 11. The level at which this professor challenged me to learn: | (1) | (2) | (3) | (4) | (5) | (N) |
| 12. The extent to which this professor connected the student learning objectives to the class activities, assignments, and assessments: | (1) | (2) | (3) | (4) | (5) | (N) |

Section IV: Comments on the Learning Experience

What is one important thing that you learned in this course? How did your professor contribute to your learning of it?

What aspect(s) of your class experience (course, professor, etc.) helped your learning most? **Provide specific, objective examples:**

What aspect(s) of your class experience (course, professor, etc.) could have been changed to help you learn more? **Provide specific, objective examples:**