

# Student Opinion Survey – Task Force Whitepaper

## Background

In 2018-2019, the Provost asked Faculty Senate to form a Task Force to make a series of recommendations regarding the administration of Student Opinion Surveys (SOS) on the three UW Oshkosh campuses. The method historically utilized by UW Oshkosh will be discontinued after Spring Semester of 2020. An integrated system for SOS collection and administration must be functional by summer of 2020.

## Purpose

This document lays out a set of recommendations providing guidance to the Provost on:

- (A) A university wide SOS instrument
- (B) Survey delivery method
- (C) SOS stewardship
- (D) Analysis and presentation of SOS results
- (E) Use of SOS data in personnel matters

## Policy guiding our work

UW Oshkosh policy is described in [FAC 6.4.](#) and [FAC 6.6.](#)

[Regent Policy Document 20-2, Student Evaluation of Instruction.](#)

UW System Administrative Policy 1254 (formerly HR 5) specifies that faculty and instructional academic staff must participate in an annual performance management process that incorporates the use of student evaluation of instruction in accordance with [Regent Policy Document 20-2, Student Evaluation of Instruction.](#)

## Contents

- 1) Recommendations
- 2) Committee membership
- 3) Summary of outreach and feedback activities
- 4) Resources consulted
- 5) Appendices

\*\*\*

## 1) RECOMMENDATIONS

### (A) A university-wide SOS instrument

**Background:** The Task Force was conscious of the dual purposes of Student Opinion Surveys. On the one hand, student opinion survey results are an evaluation of instruction used in personnel decisions. On the other hand, student opinions can provide insightful feedback that may inform instructional change and improvement for an instructor, course or program.

- **Recommended universal survey instrument:**

- ★ A single survey should be administered with a set of common questions (see below)
- ★ All students should have an opportunity to evaluate instruction in every course
- ★ Departments, programs and/or colleges should have the option to add questions, including comments, about the course and/or program
  - Comment prompts should be directive (examples provided below), not open-ended
  - Additional questions should be reviewed by an external body, such as the Improvement of Instruction Committee
    - This committee should maintain a bank of approved questions that could be made available as comment options

- **Recommended common questions:**

**Background:** Given that SOS results are utilized (as one tool among several) to evaluate teaching of individual instructors, the recommended questions (below) were designed to focus on the instructor’s teaching, organization, and assessment of learning. Questions were revised to be answered on a single Likert scale. The Task Force felt it was important to reduce the number of common questions to allow space for departments, programs and colleges to tailor additional questions that may impact course or program evaluation goals.

Common questions were vetted in four open forums in early November and a follow up month-long online survey to all faculty and academic staff. Student input was also solicited. Students requested the question below about “encouraging an inclusive environment.”

1. The materials (e.g., texts, readings, videos, etc.) helped me learn.
2. The activities (e.g., lectures, discussions, peer interactions, assignments, etc.) helped me learn.
3. The instructor was well prepared for class.
4. The instructor clearly explained concepts related to course material.
5. The instructor encouraged an inclusive learning environment.
6. The instructor clearly communicated expectations for tests, assignments, and projects.
7. The instructor was responsive to requests for help.
8. The instructor followed the policies in the syllabus.
9. The instructor’s feedback helped me learn.

- **Recommended Likert Scale:**

**Background:** UWO and UWC used different Likert scales. Moreover, some individual departmental forms at Oshkosh have varied the scale by question. The Task Force recommends a single common five-point scale and the addition of a “does not apply” option. This scale was vetted in four open forums in early November and a follow up month-long online survey to all faculty and academic staff.

Hardly Ever – Sometimes – About half the time – Most of the time - Almost always

Does not apply

- **Recommended Comment Prompts**

**Background:** Many departments currently solicit comments on instruction. The practices, however, vary. Open ended comments can be problematic. The Access Campus practice previously included directed comment prompts. We recommend that departments retain the option to solicit comments but provide students with more direction.

Following the revised practices at the University of Oregon, we recommend providing departments the option of adding two questions that solicit feedback about instruction in line with the common questions. The comment question provides a drop-down menu, a “none of these” option, and a mandatory comment space.

**Sample language:**

Please select one element that has been most helpful to your learning, and then provide a detailed written comment about what worked well and why.

- Course materials (texts, readings, videos, etc.)
- Course activities (e.g., lectures, discussions, assignments, etc.).
- Interactions between students
- The organization of this course
- The instructor’s preparation for class
- The way the instructor explained material
- Clear communication of expectations for assignments
- The inclusive environment
- Support from the instructor
- Feedback from the instructor
- None of the elements are helpful for my learning

*Conditional display:* based on the selection above, a specific prompt occurs such as: “What specifically about the [element selected] helped you learn?” Or “Please say more about how none of the elements above are helpful to your learning”.  
Response required.

Please select one element that could most use some improvement to help you learn, and then provide a detailed written comment about what specific changes you suggest.

- Course materials (texts, readings, videos, etc.)
- Course activities (e.g., lectures, discussions, assignments, etc.).
- Interactions between students
- The organization of this course
- The instructor's preparation for class
- The way the instructor explained material
- Clear communication of expectations for assignments
- The inclusive environment
- Support from the instructor
- Feedback from the instructor
- None of the elements above need improvement

*Conditional display:* based on the selection above, a specific prompt occurs such as “What specific change in the [element selected] would help you learn?” Or “Please say more about how none of the elements above need improvement to help you learn” Response required.

Additionally, departments and programs may want to survey students about course attributes or program qualities (for example, related to lab, clinical, etc.). Best practices to consider for departments and programs that want to add questions include keeping the number of additional questions to a minimum (less than 5), focusing on soliciting meaningful and actionable information, and avoiding the use of open-ended comment questions. We have recommend (see above) that additional questions be reviewed by an external body such as the Improvement of Instruction Committee, and that this committee maintain a bank of approved questions that could be made available as comment options.

- **Recommended “Other” Questions**

**Background:** The current SOS university form asks students about their class standing (freshman, sophomore, etc.), if they are taking the course for a grade (they usually are), what they think their grade is, and the number of hours per week that they studied outside of class. Instead, we recommend asking students about *how they supported their own learning* by attending class, studying, and meeting with their instructor.

- ★ Recommend that appropriate “how did you support your learning” questions be developed for each mode of instruction – face to face, self-paced online, synchronous online, and point to point / distance learning.

**Sample questions for face-to-face classes:**

**How Did You Support Your Own Learning?**

How often did you attend class?

- I attended all classes
- I attended most classes
- I attended about half the time
- I attended less than half of the classes

How many hours per week did you spend on this course studying outside of class?

- More than 8 hours each week
- 6-8 hours each week
- 4-6 hours each week
- 2-4 hours each week
- Less than 2 hours each week

Approximately how many times did you interact with the instructor outside of class (e.g. by email, office hours, question immediately before or after class)?

- 1-3 times in total
- 4-6 times in total
- 7-10 times in total
- More than 10 times in total
- I did not interact with the instructor outside of class (in person or electronically).

### **(B) Survey delivery method**

**Overview and background.** The current method of SOS delivery (to students) and collection (by an SOS coordinator) is no longer tenable. The Task Force considered three options:

Option 1) (re)construct a paper SOS survey process. This would require a new scanner, the development of a new software program to replace the current process, FTE to administer the SOS across the three campuses (including printing, distribution, collection, and analysis, and dissemination of results). The Task Force is also cognizant of the environmental impact of paper copies and their transportation to and from access campuses.

Option 2) extend our current contract with Qualtrics to encompass Qualtrics Classroom. We had an online presentation with the Qualtrics staff that answered several questions but raised others. Qualtrics Classroom is already going forward for online SOS delivery and collection. This option builds on that existing capacity. We anticipate that this option will require dedicated staff for a design and implementation phase, and in the long term to maintain the process.

Option 3) contract with a vendor that provides course evaluation processes – examples include [Campus Labs](#) or [Evaluation Kit](#). These options have not been evaluated separately or comparatively with respect to cost or functionality.

- **Recommended Delivery Method is Online**

We recommend sending a request for proposal with requirements to these vendors in order to compare the minimum costs of service (not including new investments in infrastructure and/or personnel). An RFP process would help us to clarify the expected requirements for the provider.

With respect to options 2 and 3, online surveys of instruction are among several integrated products that these companies provide. The costs associated with option 1 are unknown at this point but estimated to exceed what we currently spend on our paper process. Qualtrics (option 2) has a cost advantage as we have an existing contract that will encompass this functionality for all courses.

- ★ *A two-year trial of Qualtrics Classroom (that builds on our existing contract)*
- ★ *Engage in an RFP process with Qualtrics to clarify expectations*

- **Recommended Implementation Actions**

**Background:** We recommend forming an implementation team to work with Qualtrics in the design of the survey, assignment of course and instructor level variables, parameters of data permissions, and presentation of results to users on a dashboard. The dashboard provides visual presentation of the data to the end user as well as access to tools for analyzing and downloading data. The implementation team also needs to design a method for evaluating how well the new system is working so that continuous improvement can occur in consultation with Qualtrics as the new online SOS process is launched, and so that the two-year trial may be assessed prior to the end of the trial period. Members of the implementation team should include individuals from the units needed to administer online SOS.

- ★ *An implementation team should be formed to work with Qualtrics to design and develop the online SOS survey, identify appropriate course and instructor level variables, and design the university “dashboard” (see D below)*
  - *This team should consist of:*
    - *A liaison to Qualtrics*
    - *An individual from institutional research*
    - *An individual from Canvas Administration*
    - *Individual(s) from the Improvement of Instruction Committee*
    - *An individual from OCE*
    - *Individual(s) from the SOS taskforce*
    - *An individual from the Provost’s office (ultimately responsible for the SOS process)*

- Additional individuals as needed to assure representation from all colleges and campuses.
    - The implementation team should additionally design an evidence-based evaluation process before the trial starts
  - ★ Prior to the end of the trial contract period with Qualtrics, an evaluation of the process and product should be conducted.
- **Recommended reporting structure for SOS delivery and oversight**

**Background:** The Provost’s Office ultimately is responsible for administration and supervision of the SOS process. The ongoing administration and operation of an online SOS process will differ from our current practice (which has a fulltime dedicated position) and will differ from the design and implementation phase of the project. The Provost’s office will need to establish a permanent position and reporting structure of support personnel needed to administer the online SOS on an ongoing basis. Administration, evaluation, and support of an online SOS process will be a cross-unit effort. A new SOS position may be part time with significant dedicated time as needed throughout different points in a semester. The individual occupying this position will need to work with, and have the support of the registrar, Canvas administrators, faculty and staff, deans, program directors and chairs, and the information technology office.

- ★ The Provost’s office should create a permanent part-time position and reporting structure of support personnel needed to administer the online SOS on an ongoing basis.

- **Recommended practices to support online delivery**

Many universities now use online processes with documented declines (often precipitous) in student response rates. There are several ways to mitigate the decline, but even following best practices, declines in student participation are inevitable. Best practice includes setting aside class time for students to complete the survey (as they do a paper survey) on their own device and rewarding the class as whole for achieving a high collective response rate. Noted problems include low response rates (well below 50%) and responses by students who normally do not attend class and tend to be disgruntled.

While a move to online administration may reduce response rates, research indicates that the distribution of results is unlikely to change; “distributions of most faculty members continue to cluster at the high end of the scale as do most aggregate departmental and college distributions” (Linse 2017).

- ★ Educate students on the importance of participating in the SOS process; and provide clear instructions for students as part of the survey.
- ★ Instructors should set aside class time to administer the SOS electronically in face-to-face classes.

### **(C) SOS Data Stewardship**

**Overview and background.** Concerns about data security, custodianship, and access motivated discussion about the relationship between a third-party provider of survey services, UW System, and the institution in licensing agreements; about how individual personnel data is stored, for how long and what happens to the that data when they separate from the university; and about who controls access to an individual’s SOS data. These were difficult questions. The recommendations below reflect existing UWO practice as well as need for consistency across campuses.

- **Recommended Policy Clarification**

- ★ The Improvement of Instruction Committee should propose a university-wide SOS data use policy that provides assurances and direction regarding the protection of personnel data and access by supervisors and employees to their data. Such a policy should address these key points:
  - Clarify existing policies regarding the storage and access of personnel data.
  - It is the responsibility of faculty and instructional academic staff to download, contextualize, and present their own SOS data for personnel decisions.
  - Chairs (or equivalent in a supervisory role) should have access to instructor-level data for those who report to them.
  - Instructors, Chairs, Program directors, and Deans should have access to aggregate program-level data.
  - Provide a mechanism for employees to access their data in the event they leave the university (and no longer have a Net ID). [clarify current policy and ensure that we don’t violate any other policy statements]

### **(D) Analysis and presentation of SOS results**

**Overview and background.** The key question here is what measures are produced by the survey program and how the results are presented. In addition to the raw scores produced by the questions above, the SOS instrument is capable of automatically capturing additional data related to the course and instructor. There are several factors that may affect (and can bias) SOS scores that are out of an instructor’s control. Such variables include instructor’s gender, perceived attractiveness, ethnicity, nationality, and race (Boring 2017; Mengel, Sauermann and Zolitz 2018; Spooren, Brockx, and Mortelmans 2013; Uttl, White & Gonzalez 2017). Factors such as whether the course is a requirement, the level and size of the class, or the mode of instruction can also affect SOS scores. At minimum, collected variables, beyond the raw data, should include available instructor and course specific attributes.

Evidence suggests that typically there are issues with the statistical validity of SOS scores related to the non-standard distances between ratings on the Likert scale used. “Most student ratings distributions are skewed, i.e., not normally distributed, with the

peak of the distribution above the midpoint of the scale. The mean misrepresents the ratings in a skewed distribution because a few low ratings in the tail of the distribution can pull the mean down” (Linse 2017). This makes interpreting results – especially when comparing means – problematic. Many departments use the mean or median of SOS scores. However, the use of a mean is not recommended because the numbers produced by the Likert scale are not numbers as such, but a way of ranking responses.

Better statistical choices include the mode (the most common response), the median (middle response), the range and interquartile range (to show variability) and / or a bar chart or frequency table.

If an online method of delivery is approved, there will be options of how the data is presented to end users – these decisions will need to be made with the vendor to construct the UWO dashboard from which individual instructors would be able to analyze, interpret and download their data.

- **Recommended Presentation of Data**

- ★ **Dashboard reporting options should include:**
  - Distributions, sample sizes, and response rates for each question on the instrument.
  - Mean and median for each question.
  - Query options that allow the instructor to compare SOS responses for the same course across time.
  - Query options that allow the instructor to generate and display trends by question.
  - Query options that allow the instructor to generate and display responses by course type and mode of instruction (or other control variables).
  - Query options that allow the instructor to download and present data in graphic, tabular or text (for comments) form.
  - Query options that allow the instructor (chair, dean, etc.) to view and download aggregate data. [box plots?]
- ★ **Cross-listed and multiple sections can be aggregated and disaggregated by instructor level variables as appropriate**

### **(E) Use of SOS data in personnel matters**

**Overview and background.** UW System mandates in [Regent Policy Document 20-2](#) that we collect information from student surveys of instruction to be used to improve instruction. These data results are also only one of several instruments used in decisions for retention, promotion, tenure, and merit of faculty and instructional staff. How student surveys of instruction are to be administered and used is up to each institution. UW Oshkosh’s current policies are described in [FAC 6.4.](#) and [FAC 6.6.](#) of the Faculty and Academic Staff Handbook.

It is important to keep in mind that student opinion surveys measure student perception about their experience in a course taught by a faculty member at one point in time. Student perception should not be confused with teaching evaluation.

UWO evaluates teaching holistically, reflecting a teacher-scholar model and consisting of peer evaluations, SOS scores accompanied by a narrative, and a statement of teaching philosophy. This multidimensional method of evaluation is sometimes termed a Teaching Quality Framework. Despite long use of this holistic model, there is a good deal of apprehension about the use of SOS data, at UW Oshkosh and nationally.

Concerns that are commonly raised include the following (Linse 2017):

- Student ratings may be used as the sole measure of teaching
- Faculty may manipulate students to achieve higher ratings
- Students may be biased against certain faculty members (and no one will notice)
- Ratings may not reflect use of effective teaching methods
- Correlations with other variables make the ratings invalid or unreliable
- Online response rates may be too low to be representative
- Students may not take the ratings seriously, may lie, or may be overly critical

Existing research highlights the following points regarding the use of SOS data in personnel decisions (Barre 2015, 2018; Linse 2017):

- Research consistently fails to find evidence of a compelling correlation between measures of student learning and ratings of teaching quality and effectiveness (Uttl, White and Gonzalez, 2017). As a result, SOS data should not be used as the only, or even primary, evidence of teaching effectiveness. Every effort should be made to reinforce that the SOS is part of a holistic assessment that includes peer observations, reviews of teaching materials, statements of teaching philosophy, and instructor narratives about teaching effectiveness.
- SOS data should not be used to compare individual faculty members or instructors to each other or to compare individual faculty members or instructors to a department or college average. Measures of central tendency (e.g., means) are inappropriate statistics to use with ordinal data (categories) (Stark & Freishtat, 2014). Berk (2006) suggests that comparisons, especially between subject areas, are inherently flawed and should be avoided. Instead, comparing one instructor's scores to their own previous scores would be more informative and would allow individuals to see how the instructor has improved or changed over time. SOS results are most appropriately used to document patterns over time.
- Faculty members or instructors should present a complete history of SOS data rather than a single composite score for personnel decisions (Linse 2017).
- As noted above, there are several factors that may potentially affect SOS scores that are out of an instructor's control. These include attribute of the instructor

such as their gender, perceived attractiveness, ethnicity, nationality, and race (Boring 2017; Mengel, Sauermann and Zolitz 2018; Spooren, Brockx, and Mortelmans 2013; Uttl, White & Gonzalez 2017). Other course specific factors such as the topic, whether the course is a requirement, the level and size of the class, or the mode of instruction may also affect SOS scores (Spooren, Brockx, and Mortelmans 2013; Uttl and Smibert 2017). These are all important variables that evaluators (e.g., chairs, deans, hiring committees, tenure and promotion committees) need to take into consideration when interpreting SOS data.

- Contextual data allow individuals to discuss and interpret their SOS scores within a broader context. Departments can provide opportunities for instructors to offer their own perspective on their results, and discuss any special circumstances that may be relevant to interpreting results (a new course or innovations in methods or materials, a shift from an elective to a required course, changes in departmental goals, etc.).
- Validity of quantitative data is linked to the number of respondents and sample size (Goodman, Anson, & Belcheir, 2015; Barre, 2015). Committees and administrators should carefully consider sample size when using quantitative data from student surveys of instruction.

At UW Oshkosh there is quite a bit of variation across colleges and departments in the practice of SOS reporting in personnel documents. Some of the current practices are ill-advised. Additionally, there are challenges presented by different types of programs and instruction. For example, SOS results from cross-listed courses (including point-to-point instruction) currently are not merged into a single course response. This practice reduces the number of responses and can produce skewed and often identifiable data. Collaborative online programs have been traditionally evaluated separately by UW System. Finally, cost-recovery programs and some graduate programs may be using different and/or additional instructor evaluation methods.

- **Recommended University-wide Practices**

- ★ Instructors should present a complete history of SOS data for personnel decisions for all courses taught in the relevant time period of evaluation for the personnel action.
- ★ SOS data should not be used as the only, or even primary, evidence of teaching effectiveness.
- ★ Individuals should not be compared to a departmental, college or university mean score (or set of scores) as a primary measure of performance.
- ★ Departments should provide instructors with opportunities to contextualize their SOS results in a narrative statement and supplement the minimum aggregate data with additional data analysis.
- ★ Departments or programs should develop instructor SOS reporting policies that facilitate personnel decision making and annual performance review
- ★ Colleges should adopt policies that provide guidance on how data are reported in personnel files to facilitate personnel actions.

- We recommend that, at a minimum, a summary by question over the time period of review be presented.

\* Acceptance of these recommendations will require revisions to **FAC 6.4.** and **FAC 6.6.**, and subsequently to college and departmental bylaws.

## 2. COMMITTEE MEMBERSHIP

- Stephen Bentivenga (COLS Faculty, Oshkosh Campus)
- Debbie Busse (SOS Coordinator, Oshkosh Campus)
- Carol Collien (COLS Faculty, Oshkosh Campus)
- Chris Edwards (COLS Faculty, Oshkosh Campus)
- Carmen Faymonville (Provost’s Office)
- Renee Gralewicz (COLS Faculty, Fox Cities Campus)
- Zach DeBroux (Student Representative, Oshkosh Campus)
- Gail Panske (COLS Faculty, Oshkosh Campus)
- Druscilla Scribner (COLS Faculty, Oshkosh Campus)
- Juergen Sidgman (COB Faculty, Oshkosh Campus)
- Stacey Skoning (CEHS Faculty, Oshkosh Campus)
- Jim Stevens (COLS Academic Staff, Oshkosh Campus)
- Julie Thyssen (IT and Qualtrics Support, Oshkosh Campus)
- George Waller (COLS Faculty, Fox Cities Campus)-Chair

## 3. SUMMARY OF OUTREACH AND FEEDBACK ACTIVITIES

The Task Force held two rounds of forums on all three campuses in the months of November and December 2019. The first round was to inform instructional staff about the Provost’s charge to the task force and to solicit comments and feedback on the set of common SOS questions the task force would recommend be included on the all-university student opinion survey forms. Task force representatives attended two initial forums on the Oshkosh campus (November 5 & 6); and one each on the access campuses (November 4-Fox Cities; November 8-Fond Du Lac). Following those forums, instructional staff on all three campuses were invited to participate in an online Qualtrics survey to provide additional feedback and comments on the set of 9 common questions developed by the task force. Approximately 130 responses were received.

Task Force representatives held a second round of forums on all three campuses to provide additional information, and to solicit feedback and commentary regarding potential Task Force recommendations to the Provost on SOS administration, data collection and storage, data stewardship, analysis and presentation of SOS results, and the appropriate use of SOS data in personnel matters. Open forums were held at the Oshkosh campus on December 4 and 10; on the Fox Cities campus on December 6; and on the Fond Du Lac campus on December 11. The dates, times, and locations of all forums and the Qualtrics question survey were announced, and invitations sent, via email and Outlook calendars, as well as postings on the UWO Faculty Senate website.

Questions

and concerns raised were discussed in the Task Force meetings and incorporated into the recommendations.

#### 4. RESOURCES CONSULTED

UW Colleges White Paper: “Student Survey of Instruction (SSI) Guiding Document”

[Statement on Student Evaluations of Teaching, American Sociological Association, September 2019.](#)

Barre, Elizabeth. 2015. “Student Ratings of Instruction: A Literature Review.” Reflections on Teaching and Learning: The CTE Blog. Rice University Center for Teaching Excellence. Retrieved from <https://cte.rice.edu/blogarchive/2015/02/01/studentratings>.

Barre, Elizabeth. 2018. “Research on Student Ratings Continues to Evolve. We Should, Too.” Reflections on Teaching and Learning: The CTE Blog. Rice University Center for Teaching Excellence. Retrieved from <https://cte.rice.edu/blogarchive/2018/2/20/studentratingsupdate>.

Berk, Ronald A. 2006. *Thirteen Strategies to Measure College Teaching*. Stylus: Sterling, Virginia.

Boring, Anne. 2017. Gender biases in student evaluations of teaching. *Journal of Public Economics*, 145, 27. DOI: 10.1016/j.jpubeco.2016.11.006

Goodman, James., Robert Anson and Marcia Belcheir. 2015. The Effect of Incentives and Other Instructor-Driven Strategies to Increase Online Student Evaluation Response Rates Assessment & Evaluation in Higher Education, 40(7): 958-970.

Linse, Angela R. 2017. “Interpreting and Using Student Ratings Data: Guidance for Faculty Serving as Administrators and on Evaluation Committees.” *Students in Educational Evaluation* 54:94-106.

Mengel, Friederike, Jan Sauermann, and Ulf Zolitz. 2018. “Gender Bias in Teaching Evaluations.” *Journal of the European Economic Association* 16.

Spooren, Pieter, Bert Brockx, and Dimitri Mortelmans. 2013. “On the Validity of Student Evaluation of Teaching: The State of the Art.” *Review of Educational Research* 83 (4): 598–642.

Stark, Philip B. and Richard Freishtat. 2014. “An evaluation of course evaluations.” *Science Open Research*: <http://www.specs-csn.qc.ca/site-com/qlp/2015-2016/2016-03-30/articles.pdf>

Uttl, Bob, and Dylan Smibert. 2017. “Student Evaluations of Teaching: Teaching Quantitative Courses Can Be Hazardous to One’s Career.” *PeerJ* 5 (May): e3299. <https://doi.org/10.7717/peerj.3299>

Uttl, Bob, Carmela A. White, and Daniela Wong Gonzalez. 2017. “Meta-Analysis of Faculty’s Teaching Effectiveness: Student Evaluation of Teaching Ratings and Student Learning Are Not Related.” *Studies in Educational Evaluation* 54: 22–42.