

Redesigning the Course and Teacher Ratings: Methods, Outcomes, and Lessons Learned

S. Stavros Valenti, Hofstra University

Kevin Patrick Nolan, Hofstra University

Dr. Lynn A. Albers, Hofstra University

Dr. Lynn Albers is an Assistant Professor in Mechanical Engineering of the Fred DeMatteis School of Engineering and Applied Science at Hofstra University. Her previous academic contribution was as one of the founding five faculty/staff at Campbell University, helping the newly formed School of Engineering grow and establish roots in the community. A proponent of Hands-On Activities in the classroom and during out-of-school time programs, she believes that they complement any teaching style thereby reaching all learning styles. She earned her doctorate in Mechanical Engineering from North Carolina State University specializing in thermal sciences where her dissertation research spanned three colleges and focused on Engineering Education. Her passions include but are not limited to Engineering Education, Energy Engineering and Conservation, and K-20 STEM Outreach. Prior to matriculating at NCSU, she worked at the North Carolina Solar Center developing a passion for wind and solar energy research while learning renewable energy policy. She combined these passions with K-20 STEM Outreach while a National Science Foundation Fellow with the GK-12 Outreach Program at NCSU where she began Energy Clubs, an out-of-school-time program for third, fourth and fifth graders to introduce them to renewable energy.

Redesigning the Course & Teacher Ratings: Method, Outcomes, and Lessons Learned

Less than one year into the COVID-19 pandemic, the provost and faculty union leadership at Hofstra University, a midsized private university in Hempstead, New York, agreed that the time was right for a reevaluation of the student evaluation of teaching (SET) process and policy, which included a Course and Teacher Ratings (CTR) system and Peer Observation of Teaching policies. For example, there was agreement that some of the fixed-choice items on the SET instrument, created more than 20 years ago, were not well suited for a growing number of online, hybrid, and team-taught courses. Important aspects of the learning experience, such as the course climate, were not included, whereas other dimensions, such as “difficulty,” were covered by multiple questions. In addition, the university had just begun to pilot an online SET process—a change that, argued by some members of the faculty, may distort the results of the course and teacher rating process. For these reasons, a thorough review of the course evaluation process was commenced.

Student evaluations of teaching, viewed from an organization perspective, are key components of a performance management system for teaching. Performance management systems serve to ensure that a set of activities and outputs meets an organization’s goals in an effective and efficient manner. The feedback afforded by these systems is used to facilitate employee development and inform personnel decisions. Using performance management as a lens for understanding SETs, a university-wide Blue Ribbon Committee (“Committee”) was formed to evaluate simultaneously the SET and the peer observation processes and policies and to make recommendations as appropriate. The authors of this paper served on the Committee. No deadline for the Committee’s work was set in advance, although it was hoped that recommendations could be forwarded to the full faculty for review by the end of the Spring 2022 term. This paper describes the method, outcomes, and lessons learned in our attempts to improve a SET instrument using a performance management mindset.

From Paper-and-Pencil to an Online Process

Prior to Spring 2020, “Our University” (OU) was using a paper-and-pencil system for SET data collection. COVID-19 necessitated the suspension of SETs for the Spring 2020 term, and the adoption of a stop-gap online SET distribution and data collection system for the Fall 2020 onward. To make the online SET process conform closely to existing SET policy for in-class, paper-and-pencil administration, a temporary system was designed to permit faculty to schedule a specific start date, with a 48-hour response window, for the administration of the online SET for their course sections. The faculty union and University administration entered a memorandum of agreement (MOA) wherein the Committee was tasked with examining the congruence between fall 2020 “online” SET scores and the fall 2019 “paper-and-pencil” SET scores. Within the Course and Teaching Ratings MOA was the clause, “For the 2020-2021 academic year, [SETs] are expected, but not required, to be administered for all courses with enrollment greater than five students.” It is possible that lifting the firm requirement of SET administration for the Fall 2020 term may have impacted the results of subsequent analyses.

Of the 2,327 eligible sections for SET administration in fall 2020, 52% (N=1,217) were scheduled for the distribution of an online SET. Of those sections, 84% (N=1,024) received at least one response, and 49% (N=598) received 5 or more completed forms from students and, following university policy, were eligible for consideration in personnel actions. These results suggest an overall response rate of approximately 41% based on the percentage of student responses within sections having at least one response. This response rate was lower than those for fall 2017 (84%), fall 2018 (84%), and fall 2019 (83%). However, our 41% response rate was not far off the 50% average response rate that has been found across universities that use online student evaluations [2]. Disaggregating fall 2020 response rates by instruction method revealed the highest rates in face-to-face (45%) and online synchronous (44%) courses, and the lowest rates in online asynchronous courses (27%).

Online SETs are now the norm at institutions of higher education due to a lower environmental impact, convenience, lower costs, and timely feedback for instructors. Because SETs were optional during fall 2020 and completed under novel circumstances, the Committee recommended retaining the online method with continued efforts on improving student response rates (e.g., encouraging instructors to give students time in class to complete the SET, explaining the value of SETs for the continuous improvement of instruction, sending reminders to students before and during the SET response interval).

Guiding Principles

Our diverse Committee agreed on three guiding principles (a) Update the SET instrument to make it a more useful instrument for faculty development; (b) Include items that capture student perceptions of class climate; (c) Broaden the scope of teaching behaviors assessed to reflect the broad range of course structures and effective teaching styles of our faculty.

The Committee researched and compared the SET standards and processes at OU to peer and aspirant institutions. As part of this research, we examined best practices for preventing bias in responses from students. We note that none of OUs peer and aspirant schools are using paper-and-pencil SET data collection processes. Benchmarking information and best-practice insights were gleaned for both the solicitation of instructional feedback from students and peer observations of teaching from the following academic institutions via their online resources:

Algonquin College	Georgia Institute of Technology	St. John’s University	University of North Carolina at Charlotte
Augsburg University	Indiana University of Pennsylvania	The Pennsylvania State University	University of South Carolina
Baruch College - CUNY	Iowa State University	The University of Maine	University of Toronto
Boise State University	Loyola Marymount University	University of California, Berkeley	University of Vermont
Clemson University	McKendree University	University of Colorado	University of Wisconsin-Madison
Colorado State University	Queens College - CUNY	University of Kentucky	Vanderbilt University

DePaul University	Rutgers University	University of Michigan	Western Michigan University
-------------------	--------------------	------------------------	-----------------------------

Assessing the Current SET Instrument

1. Developing a competency model

Since no formal competency model for teaching had been developed at the University at the time, the Committee began with a content analysis of SET items and the dimensions that peer observations of teaching should assess according to the existing faculty policy series. (Table 1.) The core competencies fell into four key areas: Learning/Development, Class Climate, Instruction, and Assignments/Assessments

Table 1. General Competencies Evaluated via CTRs (Course Teacher Ratings) and Peer Observation of Teaching

CTR Items	General Competency	Peer Observation of Teaching Dimensions (FPS#49)
In your opinion, the instructor demonstrates a mastery of the subject that is (Outstanding – Poor).	<i>Demonstrating Subject Mastery</i>	Mastery of the material presented in light of the current state of knowledge in the discipline.
The instructor’s presentation of the subject matter is (Always clear – Never clear).	<i>Clear & Effective Instruction</i>	Clarity of presentation and effectiveness of communication skills.
		The effectiveness of the form of presentation (e.g., lectures, discussions, use of technology).
		The appropriateness of the material presented in light of the stated purpose of the course.
The instructor encourages student participation (A great deal – Not at all).	<i>Facilitating Student Participation</i>	Encouragement and management of classroom or online participation.
The instructor encourages meaningful interactions from/among students (A great deal – Not at all).		
The instructor’s responses to your questions are (Always clear – Never clear).	<i>Responding to Student Questions/Comments</i>	Responsiveness of instructor to students (questions and comments and other contributions).

The instructor presents a grading policy that is (Very clear – Very unclear).	<i>Fairness & Transparency in Grading</i>	
The instructor follows a grading policy that is (Fair – Unfair).		
Exams are based on materials covered in class and/or assigned during the course (Always – Never).	<i>Effective Assignments & Assessments</i>	
Assignments contribute to the learning experience in this course (A great deal – Very little).		
The quality of the feedback the instructor gives concerning tests or other assignments is (Very useful – Not useful).		
Considering the level of difficulty of this course, the course is paced (Very fast – Very slow).	<i>Academic Rigor</i>	
Compared to other courses you have taken at this University, the level of difficulty of this course is (Very difficult – Very easy).		
The examinations or graded assignments are (Very difficult – Very easy).		
Texts and other required reading materials for the course are (Very difficult – Very easy).		
As a result of this course, your knowledge in this area of study (Increased greatly – Remained the same).	<i>Student Learning</i>	
	<i>Classroom Management</i>	Classroom management (e.g., prompt start time, classroom control)
How would you rate the instructor's effectiveness as a teacher (Outstanding – Very poor)?	<i>Overall Teaching Effectiveness</i>	Overall assessment of the class or online module.

2. Feedback from the Department Chairs

With the general competencies established, the next step involved soliciting feedback from the Department Chairs on their beliefs about the usefulness of the SET instrument and peer observation of teaching procedure. Next, an online survey was developed and administered to the Department Chairs that included 4 Likert-style survey items soliciting their beliefs about the usefulness of the SET instrument and peer observation of teaching procedure for formative and summative purposes, and 2 open-ended items asking for suggestions concerning how these practices might be improved. Twenty-five chairs responded, completing both sections of the survey. Results (Tables 2 and 3) suggest that chairs generally agree that the peer observation of teaching procedure is useful for both facilitating faculty development (92%) and making personnel decisions 88%). Chairs also generally agree that the SET system was useful for making personnel decisions (72%) but disagreed that the system was useful for facilitating faculty development (48%). A content analysis of the open-ended feedback provided suggests that Chairs most commonly believe that the SET system could be improved by (a) revising the question set to include more/different content – with less of a “customer service” focus, and (b) providing results to faculty more quickly. Chairs also believe that the peer observation of teaching procedure could be improved by (a) modifying the process to address leniency bias – overly positive reviews, and (b) creating a more structured system that addresses a broader range of teaching behaviors across observers, observation periods, and faculty being observed.

Table 2. Chairs Survey: Likert-style Percent Responding

Likert-Style Survey Items	Disagree	Neutral	Agree
To what extent do you agree or disagree that the current Course and Teacher Rating (CTR) system is useful for facilitating professional development for faculty?	48%	12%	40%
To what extent do you agree or disagree that the current Course and Teacher Rating (CTR) system is useful for making personnel decisions (e.g., repappointment, tenure, promotion)?	8%	20%	72%
To what extent do you agree or disagree that the current Peer Observation of Teaching system is useful for facilitating professional development for faculty?	8%	0%	92%
To what extent do you agree or disagree that the current Peer Observation of Teaching system is useful for making personnel decisions (e.g., reappointment, tenure, promotion)?	4%	8%	88%

Table 3. Chairs Survey: Content Analysis of Open-ended Responses

System	Content Analysis of Open-ended Responses	Count
CTR	Revise the question set to include more/different content (e.g., less customer service)	8
	Provide results to faculty more quickly	5
	Implement mechanisms to increase response rate (e.g., tie to grade disclosure)	4
	Adapt the form to account for online and hybrid classes	2
	Administer all CTRs online	2
	Include more, and more direct, open-ended questions	2
	Distribution of CTRs should not be tied to course times	2
	Ask fewer questions	2
	Create a mobile platform	1
	Faculty should have to report how feedback was addressed	1
	Replace new online system with in-class CTRs, as before	1
	Copy chairs on open-ended comments made on CTRs	1
	Peer Observation	Modify system to address leniency bias (e.g., overly positive reviews)
Create a more structured system that addresses more aspects of teaching effectiveness		3
Reduce how often they are conducted to account for high-volume requirements for adjunct faculty		2
Partner feedback received with training and development interventions		1
Identify ways to deal with demographic bias		1
Faculty should have to report how feedback was addressed		1
System needs to be modified to account for format of online/hybrid courses as well as teaching		1
Teachers should not be able to select the specific course for which they will be evaluated		1

3. Factor analysis of the current SET instrument

Archival data was used to test the psychometric properties of the current SET form to examine if it measures the various aspects of instruction that it purports to evaluate. Based on previous analyses, the university condensed the 16 items of the current SET instrument into 4 subscales: (1) Overall Evaluation of Instructor and Course, (2) Workload/ Difficulty, (3) Grading/ Feedback Quality, and (4) Interaction/ Encouragement. However, an exploratory factor analysis using orthogonal rotation revealed only 3 factors, and none of them corresponded to “Overall Evaluation of Instructor and Course” -- the only factor that is consistently used for summative teaching appraisals and personnel decisions.

4. Assessment of the influence of instructor race/ethnicity and gender

Archival data was used to examine if SET subscale scores differed by instructors’ race/ethnicity and/or gender to investigate potential systematic discrimination that might exist in the evaluation system. Research on best practices for identifying and minimizing bias in student evaluations was also conducted.

What follows are preliminary and exploratory analyses of possible associations of SET results with gender and race/ethnicity. Concerning the “Overall Evaluation of Instructor and Course” subscale of the current SET instrument, results suggest that scores between female and male instructors in our sample did not significantly differ. Some of the previous studies on instructor gender effects on SETs revealed no empirical evidence of an overall bias in ratings related to the gender of the instructor [4], whereas others report statistically significant gender effects [5]. Research has also revealed interactions of modest magnitude such as a tendency of students to rate more favorably professors that share their gender [3].

When examining instructor ethnicity, White-Not Hispanic Origin instructors on average received slightly better overall scores than instructors of other races/ethnicities. Analyses involving other subscales suggest that Black-Not Hispanic Origin instructors received the slightly less positive scores on “Grading/ Feedback Quality,” Asian instructors received the slightly less positive scores on “Interaction/ Encouragement,” and female instructors received slightly better scores on “Interaction/ Encouragement” than male instructors. Our findings are in line with previous research that note small effect sizes for the overall influence of instructor race/ethnicity on SET scores [1], [7].

Based on these preliminary findings, the Committee suggests that additional, more complete analyses be conducted on archival data collected via the current SET instrument to examine the influence of gender and race/ethnicity on SET results. Efforts should be made to address sampling bias (e.g., participation was voluntary in 2020) and identify potential confounds (e.g., types of courses taught) and methodological artifacts (e.g., questionable psychometric properties of the SET subscales) that may have influenced the direction and magnitude of observed effects. Additionally, given the importance of the issue, the Committee suggests that further monitoring of race/ethnicity and gender group differences be regularly conducted with use of student feedback forms regardless of whether the form remains that in current use or a modified version of the instrument is adopted.

Revising the SET Instrument

Our research on the psychometric properties of the current SET instrument form raises concerns about the construct validity of the measure. Furthermore, only 72% of Department Chairs who responded to our survey agreed that SET scores are useful for making personnel decisions – with only 48% of respondents agreeing that they are useful for facilitating faculty development. The most frequently given feedback from Department Chairs about the SET system is that it should assess a greater variety of instructional dimensions.

In examining the informal competency model measured by OUs implicit performance management system (Table 1), SET items were determined to assess constructs relating to:

- Demonstrating Subject Mastery (1 item),
- Clear & Effective Instruction (1 item),
- Facilitating Student Participation (2 items),
- Responding to Student Questions/Comments (1 item),
- Fairness & Transparency in Grading (2 items),
- Effective Assignments & Assessments (3 items),
- Academic Rigor (4 items),
- Student Learning (1 item), and
- Overall Teaching Effectiveness (1 item)

Next, the Committee examined SET items from peer and aspirant institutions and compiled a list of more than 200 items. These items were then classified according to the competencies/constructs they were perceived to assess. A review of this content, both across and within the academic

institutions from which it was derived, identified multiple points of parity as well as multiple points of differentiation from the content assessed by OU’s SET form. Notably, the form used by OU dedicates more items to the assessment of students’ perceptions of academic rigor than other institutions, whereas other institutions tend to focus more attention on assessing multiple aspects of student learning (e.g., This course challenged me intellectually), the development of communication and critical thinking skills (e.g., This course increased my ability to think critically), having clear learning goals and objectives (e.g., Course goals and learning objectives were clearly communicated), and facilitation of a welcoming instructional climate wherein values relating to diversity and inclusion are emphasized (e.g., The instructor created a welcoming and inclusive learning environment).

To begin the process of narrowing down the list of potential content to include in a revised SET form, 10 members of the Committee were surveyed about their beliefs concerning the importance of assessing 40 different competencies/constructs nested within 8 unique competency domains via the CTR system. Based on the survey findings and a review of the student feedback forms used at other academic institutions, the Committee determined that the target length for the University CTR form should be between 10 and 15 items, with two of those items dedicated to evaluating students’ overall rating of the course and overall rating of the instructor. The feedback provided from Committee members via the internal survey was then used to create an abbreviated version of the survey for the community to solicit feedback from students and faculty concerning their beliefs about what content is most important to include on a revised CTR form.

For the Community Survey, 14 items were written to collect insights about what content the University students and faculty believe is most important to include in a revised SET form. The survey was drafted following a review of more than 200 CTR items compiled from benchmarking research. Faculty and Student participants were asked to rank order the 14 items in terms of their importance. The community survey demographics are in Table 4.

Table 4. Community Survey Demographics

	Student	Faculty
Sample Size (n)	700	227
Gender		
Female	67.60%	46.12%
Male	28.10%	49.14%
Other/Prefer Not to Answer	4.3%	4.74%
Race/Ethnicity		
Asian	12.28%	8.89%
Black-Not Hispanic Origin	9.14%	2.02%
Hispanic	11.80%	4.87%
White-Not Hispanic Origin	64.62%	74.09%
Other	1.56%	5.67%
Prefer Not to Answer	.96%	4.46%
School		

Social Science	2.08%	8.26%
Education	9.08%	7.83%
Humanities/Arts	12.95%	18.70%
Science/Math	9.82%	18.26%
Health Science	17.11%	15.65%
Medicine	6.55%	1.74%
Business	16.07%	14.35%
Engineering	9.97%	6.96%
Communication	13.54%	8.26%
Undecided	2.83%	0.00%

The rankings assigned to each item were analyzed separately for students and faculty, and then results across the samples were compared to inform retention decisions. The relative perceived importance of each item was calculated by subtracting the frequency with which the item was identified as being of bottom-three importance from the frequency with which it was identified as being of top-three importance. Using these findings, a set of 11 items that assessed a range of constructs indicating teaching effectiveness were retained. These items and their relative rankings of importance by students and faculty are provided in Table 5.

Table 5. Relative Importance Rankings Assigned to Retained Items

Item	Student	Faculty
The instructor's presentation of course material was clear and effective.	1	2
The instructor treated all students with respect.	2	5
The instructor created a welcoming learning environment.	3	3
Through this course, my knowledge of the subject increased.	4	1
The instructor conducted class in an organized manner.	5	8
The instructor gave me constructive feedback on assignments and assessments.	6	10
The instructor clearly communicated course goals and requirements.	7	6
Assignments positively contributed to the learning experience in this course.	8	9
Graded assignments and assessments were a fair reflection of the material taught in this course.	9	11
This course advanced by professional development.	10	7
This course helped me develop intellectual and/or critical thinking skills.	11	4

Pilot Test of the Revised SET Form

A revised SET instrument was piloted in the fall 2021 term. The study was performed to investigate the psychometric properties of the instrument and examine issues pertaining to its use for performance management at the University. The Committee first drafted a revised version of the SET form based on feedback provided by the University Community Content Survey. The form included 13 Likert-style survey items that were written to evaluate a range of constructs indicative of teaching effectiveness and an open-ended question asking students to “Please comment on the course and instructor.” Like the current SET form, the revised form is designed to assess students’ beliefs about

- how well instructors present course material
- the extent to which course goals/requirements are well-understood,
- the extent to which assignments/assessments were a fair reflection of the material taught in the course and positively contributed to the learning experience
- the quality of feedback given by instructors
- the extent to which their knowledge of the subject increased
- the overall effectiveness of the instructor as an educator.

Unlike the current CTR form, the revised form is also designed to provide feedback concerning students’ beliefs about

- their intellectual and professional development
- the organization of the course
- aspects of the learning environment
- overall quality of the course.

This extra content was included on the revised SET form based on best practices and the judgment of faculty and students. The revised form further differs from the current form in that feedback concerning students’ beliefs about instructors’ mastery of subject matter, appropriateness of student encouragement, course pacing, fairness in grading, and course difficulty is no longer solicited. This content was omitted on the revised form due to concerns about its usefulness for performance management purposes (e.g., perception of course difficulty relative to other courses at OU) and/or validity (e.g., the appropriateness of students rating instructors’ mastery).

The revised SET form also differs from the current form in its use of a consistent response scale across items. Whereas the current form utilizes a variety of response scales often unique to individual items (e.g., levels of magnitude, clarity, speed, difficulty), all items on the revised form are designed to be answered using the same 5-point Likert scale (Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree). Doing so aligns with best practices in cognitive survey methodology (c.f., Schwartz, 2007) as consistency tends to improve response accuracy by reducing cognitive demand and lowering likelihood of unintentional endorsements.

Using a consistent response scale also benefits analysis and interpretation of survey data in that it affords greater statistical validity in the comparison of response trends across items and/or subscales. The revised SET survey items and their corresponding response scales for the pilot study are presented in Table 6.

Table 6. CTR Revised from Items and Response Formats

5-pt Likert (Agreement) Response Format
1. The instructor clearly communicated course goals and requirements.
2. The instructor's presentation of course material was clear and effective.
3. The instructor conducted class in an organized manner.
4. The instructor created a welcoming learning environment.
5. The instructor treated all students with respect.
6. The instructor gave me constructive feedback on assignments and assessments.
7. Graded assignments and assessments (e.g., quizzes, exams, papers, projects, assigned problems, performances, presentations) were a fair reflection of the material taught in this course.
8. Assignments (e.g., readings, projects, assigned problems, performances, presentations) positively contributed to the learning experience in this course.
9. Through this course, my knowledge of the subject increased.
10. This course helped me develop intellectual and/or critical thinking skills.
11. This course advanced my professional development.
12. Overall, this instructor is an effective educator.
13. Overall, this course was a worthwhile experience.
Open-ended Response Format
14. Please comment on the course and instructor.

Useable data was collected from a sample of students (N = 850) enrolled in fall 2021 courses wherein their instructors volunteered to participate in the pilot study. Results indicated that the middle 50% of student completion times fell between 70 and 252 seconds. Item-level descriptive statistics suggest mean scores were typically favorable across items with considerable variance in participant responses. Bivariate correlations between items in the revised CTR pilot ranged in magnitude from $r = .40$ to $.78$ – with only 6 of the 78 relationships indicating potential redundancy among the items (i.e., $r > .70$). An examination of the content assessed by those items, however, indicates they evaluate unique, but related, aspects of teaching effectiveness.

Covariation in participants' responses to the items was further examined using exploratory factor analysis to investigate how the revised CTR items cluster together based on response patterns. Results identified four unique item clusters that map onto constructs relating to, "Learning/Development," "Class Climate," "Instruction," and "Assignments/Assessments." Student responses to the single item "Overall, this instructor was an effective educator" were most strongly influenced by beliefs relating to the "Class Climate" and "Instruction" clusters. Responses to the item "Overall, this course was a worthwhile experience" were most strongly influenced by beliefs relating to the "Learning/Development" cluster.

There was considerable variance in students' responses to each item despite the strong potential for sampling bias to produce leniency effects. Items appear to be assessing unique aspects of the educational experience, with responses to similar items (e.g., assignments/assessments items; climate items) clustering as expected.

Conclusion

Redesigning the Course & Teacher Ratings of a midsized, private University is no small feat. COVID-19 necessitated and expedited the transition from paper-and-pencil to online. Once online, then efforts were made to streamline the response rates. While a necessary change, this was not the primary initiative. A goal of the Committee was to redesign the Course & Teacher Ratings to remove bias and provide less subjective feedback. Student evaluations of teaching, viewed from an organization perspective, are key components of a performance management system for teaching. Performance management systems ensure that a set of activities and outputs meets an organization's goals effectively and efficiently. The feedback afforded by these systems is used to facilitate employee development and inform personnel decisions. Consistent with best practices in performance management, the current and revised SET instrument's method was guided by the university's mission and goals. The Committee succeeded in transitioning the process from paper-and-pencil to online, developing a revised Student Evaluation of Teaching (SET) instrument that will facilitate evidence-based personnel decisions (e.g., appointment, promotion, & tenure), training, and development opportunities, and streamlining online course evaluations and response rates.

Acknowledgements

Special thanks to the Blue Ribbon Committee for their dedication to redesigning the Course & Teacher Ratings. The Committee members include

S. Stavros Valenti (Co-Chair) – Vice Provost for Accreditation and Assessment

Kevin P. Nolan, (Co-Chair) – Associate Professor of Psychology, HCLAS

Lynn A. Albers – Assistant Professor of Engineering, SEAS

Corinne Kyriacou – Associate Professor of Health Professions, SHPHS

Sabine Loucif – Professor of Romance Languages and Literatures, HCLAS

Renee Mcleod-Sordjan – Professor of Nursing, SONPA

Mario Murillo – Professor of Radio, Television, Film, LHSC

Uzo N. Osuno – Adjunct Assistant Professor of Computer Science, SEAS

Daniel Seabold – Acting Dean, HCLAS and School of Education

Chavon C. Stuparich – Executive Director of Institutional Research & Assessment

Shawn T. Thelen – Professor of Marketing and International Business, ZARB

References

- [1] Chavez, K., & Mitchell, K. M. W. (2020). Exploring bias in student evaluations: Gender, race, and ethnicity. *PS: Political Science and Politics*, 53(2), 270-274. DOI: <https://doi.org/10.1017/S1049096519001744>
- [2] Goodman, J., Anson, R. and Belcheir, M., (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.
- [3] Kreitzer, R.J., & Sweet-Cushman, J. (2022). Evaluating Student Evaluations of Teaching: a Review of Measurement and Equity Bias in SETs and Recommendations for Ethical Reform. *Journal of Academic Ethics*, 20, 73–84. Retrieved from <https://doi.org/10.1007/s10805-021-09400-w>
- [4] Marcham, C.L., Ade, A.M., Clark, P. & Marion J. (2020). Bias and Trends in Student Evaluations in Online Higher Education Settings. *Collegiate Aviation Review International*, 38(2), 34-50. Retrieved from <http://ojs.library.okstate.edu/osu/index.php/CARI/article/view/8036/7417>
- [5] Mengel, F., Sauermann, J., & Zolitz, U. (2019). Gender bias in teaching evaluations. *Journal of the European Economic Association*, 17(2), 535–566. DOI: 10.1093/jeea/jvx057
- [6] Statham, A., Richardson, L., & and Cook, J. (1991). *Gender and University Teaching: A Negotiated Difference*. Albany: SUNY Press.
- [7] Wagner, N., Rieger, M., & Voorvelt, K. (2016). Gender, ethnicity, and teaching evaluations: evidence from mixed teaching teams. *International Institute of Social Studies, Working Paper 617*. ISSN 0921-0210. Retrieved from www.iss.nl