Journal of Collective Bargaining in the Academy

April 2019

Panel: Peer-Based Faculty Evaluation v. Student Evaluation of Teaching

Leah Akins
Dutchess Community College

Laura Murphy
Dutchess Community College

Follow this and additional works at: https://thekeep.eiu.edu/jcba

🟦 Part of the Collective Bargaining Commons, and the Higher Education Commons

Recommended Citation

Maintaining peer-based faculty evaluation: a case study involving student surveys of teaching

Presenters:
Dr. Leah Akins, Professor of Engineering and Technology, and
Dr. Laura Murphy, Professor of History
Dutchess Community College (SUNY), Poughkeepsie, NY

Abstract:
Bargaining regarding faculty evaluation is challenging in an environment in which administrators throughout higher education have successfully imposed corporate-style forms of evaluation and supervision that many have come to accept as normal, despite their incompatibility with principles of academic freedom and peer-review. Student surveys of teaching are increasingly central to this management strategy, despite the growing body of evidence indicating bias against historically marginalized groups in student survey results. In our presentation we will discuss our 2016 contract negotiations at Dutchess Community College (SUNY) in Poughkeepsie, New York as a case study. During these negotiations the college administration sought to expand the use of "student evaluations" of teaching (SET) despite significant evidence that student feedback provides limited meaningful evaluative content concerning teaching and is shaped by gender, racial, and ethnic bias, as well as bias against academic rigor. Our presentation will briefly describe our effort to maintain a peer-based evaluation of student survey data, including the published research we used during negotiations, and we will analyze the strengths and weaknesses of our approach and results. These results include:

- a successful effort to maintain the practice of limiting review of qualitative student feedback to peer-based review between faculty and department chairs within academic departments
- limited but significant expansion of administrative oversight of some quantitative student survey data
- contract language that limits the use of student survey results in faculty evaluation
- contract language that requires that all consideration of these data shall be undertaken with the understanding that student feedback is an important but limited vehicle for understanding the effectiveness of an individual's teaching
- contract language that established an all-faculty committee of full-time and part-time faculty that is charged with evaluating the new survey form and process

Background:
During the last contract negotiations undertaken between Dutchess United Educators (DUE), the union representing faculty and most professional staff at Dutchess Community College, and the

college administration, DUE negotiators were confronted with a demand to change the long established process of faculty evaluation and use of student surveys of teaching that would affect both full-time and part-time faculty. The decades-old full-time faculty evaluation process involved two reports produced by the faculty member's department chair and submitted to the Dean of Academic Affairs: (1) a classroom observation report and (2) a professional development report (PDR) that covered teaching effectiveness, student advisement, professional activities, and contributions to the department and college. Additionally, student surveys of teaching were administered on paper in all course sections every spring with results going to the faculty member's department chair and eventually returned to the faculty member. These surveys included statements to be rated on a Likert scale as well as opportunities for students to respond to reflective questions. Survey data was not compiled or quantified but were used by department chairs to inform their commentary on teaching effectiveness in the PDR and/or to generate conversations about teaching. No results were submitted to Academic Affairs.

The demand from administration during negotiations was that survey data needed to be submitted to Academic Affairs to assure that student voice was clearly a part of the process of evaluating faculty. Even though the negotiations went on for two years (2014 - 2016) with a one-year contract agreed to for 2015 - 2016 before eventually agreeing to a four-year contract for 2016 - 2020, pressure concerning student surveys of teaching was a recurring theme. To address this concern, union and college negotiators agreed to research the matter and make evidence-based decisions to resolve the disagreement. At the time, Dr. Akins served on the negotiating team for the Full-time Faculty and Staff 2016 - 2020 contract and Dr. Murphy served on the negotiating team for the Part-time Faculty and Staff 2016-2020 contract. For the purpose of collecting research and formulating our arguments for negotiations, a joint subcommittee on Faculty Evaluations made up of members of both the full-time and part-time teams was formed. Both Akins and Murphy served on this sub-committee and led the research effort.

Research:
Our study of the research on student surveys quickly led to an ever-expanding body of scholarship indicating that student surveys do not reliably measure the quality of teaching. While student opinions about their educational experiences are important and valuable, numerous studies demonstrate that students are not qualified to judge teaching effectiveness. In addition, research indicates that survey results are influenced by the gender, race, ethnicity, and perceived attractiveness of the instructor. For example, Anne Boring, Kellie Ottoboni, and Philip B. Stark write that "SET are biased against female instructors by an amount that is large and statistically significant," and "gender biases can be large enough to cause more effective instructors to get lower SET than less effective instructors." Boring, Ottoboni, and Stark argue that "it is not

possible to adjust for the bias, because it depends on so many factors."[1] There is also evidence to suggest that the academic rigor of a course, as well as a student's desire to take a course and how much prior knowledge student has about a subject impact survey results.[2] Any implementation of student surveys with low response rates, such as commonly occurs when surveys are delivered online, is statistically problematic, further distorting the data.[3]

Based on our review of the available scholarship we concluded that our then-present system of faculty evaluation utilized student surveys in a way that was best-suited to provide faculty with the opportunity to gain what is valuable about student opinions expressed in surveys while minimizing damage to academic freedom and academic integrity, and minimizing discrimination against faculty who are rated lower on surveys for reasons that have nothing to do with their teaching effectiveness, or are even rated lower *because* they are effective teachers.

A list of articles consulted for contract negotiations and subsequent Evaluation Committee work is provided at the end of this paper.

Negotiations:

As described in the previous section, for the purpose of negotiations we researched important questions about evaluations of teachers and the use of Student Evaluation of Teaching (SET), or, as we prefer to call it Student Surveys of Teaching (SST). This preferred name is what the form has been called historically at our campus and we believe "survey" better reflects its output. The negotiations process was based on an interest-based bargaining (IBB) model without formal training for either side to assure we understood how to apply IBB principles fairly and equitably, which was a challenge to effective negotiations.

Another challenge to effective negotiations was a silence agreement that limited our ability to mobilize faculty. Within this framework for discussion, however, there was agreement around the table that student voice had a role in evaluating the effectiveness of a faculty member. Our conflict was in the appropriate extent of that role and the process to include that voice as well as other relevant voices in the evaluation process. That conflict was not only with management but also with other faculty negotiators, which was further complicated by the silence agreement because it limited these discussions to a very small group of faculty.

---

[1] Anne Boring, Kellie Ottoboni, and Philip B. Stark, "Student evaluations of teaching (mostly) do not measure teaching effectiveness," *ScienceOpen Research* (2016) https://www.scienceopen.com/document?vid=818d8ec0-5908-47d8-86b4-5dc38f04b23e

[2] Stephen L. Benton and Kenneth R. Ryalls, "Challenging Misconceptions About Student Ratings of Instruction", The IDEA Center*, *IDEA*, IDEA paper #58, (April 2016), https://www.ideaedu.org/Portals/0/Uploads/Documents/Challenging_Misconceptions_About_Student_Ratings_of_Instruction.pdf

[3] Philip B. Stark and Richard Freishtat, "An evaluation of course evaluations." *ScienceOpen Research,* (2014) https://www.stat.berkeley.edu/~stark/Preprints/evaluations14.pdf

L. Akins and L. Murphy, NCSCBHEP 2019   3

Our limited understanding of the IBB process meant that we would agree on a shared interest, separately perform research on the topic including data collection and analysis, share with the other side our research and analysis, then discuss our differences to find what we all could agree to. While we did a significant review of literature on the topic, as described above, the administration shared a limited number of articles which mostly focussed on outdated data and analysis or were crafted by organizations that benefit financially from "quantifying" student feedback.

When confronted with evidence, including from the very same documents supplied by the administration, that the decades-long trend outside of our institution to collect, quantify, and elevate the numerical significance of student feedback is problematic, the administration was not swayed. In negotiations, we focused on two conclusions from the research: (1) students are not effective evaluators of teaching practice, and (2) student bias stemming from a multiple of possible sources leads to data biased against women, faculty of color, faculty teaching in STEM fields, faculty teaching core requirement courses, and more. We pointed out that while DCC's use of SSTs within the faculty evaluation process was outside the current norm in higher education where SET data is quantified and used to make important decisions of tenure and promotion, the research is now showing that our model may have been more effective, in that it likely minimized the impact of bias and was more consistent with professional standards of academic freedom and peer-review.

With some unintended help from the administration, we were eventually able to convince faculty negotiators to support our position that the administration's proposal could inevitably be damaging to faculty. Even though there was a "gag order" on negotiation topics, the college president spoke with a faculty member about the student survey discussion, which got back to the union president and lead negotiator. This breach of the silence agreement created an opportunity for union negotiators to gain some faculty feedback and leverage to push back on the administration's plan. The administration's lack of receptiveness to evidence that challenged their assumptions, however, along with the silence agreement limiting the extent of our ability to mobilize faculty, constrained our negotiating position. We were told that the student voice, through the student surveys of teaching, must be heard and that there was no bargaining on the point that data from those surveys had to be given to the Office of Academic Affairs.

With no possibility of getting a signed contract without dealing with the student surveys of teaching and thus the faculty evaluation process, we were cornered into developing a proposal that included providing some direct student feedback to the administration. Although the results were a compromise, and therefore included content that we had not championed, the research we

undertook was central in crafting contract language supported by evidence. The primary features of the contract language are:

- a statement of purpose for the faculty evaluation process that focuses on assisting faculty with their growth and development as educators.
    - "The purpose of professional evaluations shall be to recognize and encourage outstanding professional performance by providing a process that includes supervisory, peer, and self-review." (2016-2020 Full-time Contract, p. 24)
- a statement of purpose for collecting student feedback through the student survey of teaching process.
    - "The Student Survey of Teaching process provides a mechanism to bring the student voice in to the faculty evaluation process (see section 7.02). As is the case throughout the evaluation process, the intent and purpose is to use this information to assist the faculty member in his/her growth and development as an educator." (2016-2020 Full-time Contract, p. 53)
- a statement of concern about the inherent biases that research shows impact student survey responses.
    - "All consideration of these data shall be undertaken with the understanding that student feedback is an important but limited vehicle for understanding the effectiveness of an individual's teaching. All faculty and administrative supervisors' evaluation of student survey results will be informed by a clear understanding of the research that demonstrates that student survey responses may reflect biases based on gender, race, sexual orientation, appearance, academic rigor, subject matter of the course, and students' desire to take the course, work habits, and confidence about and prior knowledge of the subject matter. Therefore, data can be used to guide future professional development and shall not be used to initiate disciplinary procedures." (2016-2020 Full-time Contract, p. 53)
- added contract language to provide more detail about the faculty evaluation process as a peer-based system that considers student surveys as a factor for discussion concerning professional development but not a main factor in tenure and promotion decisions.
- survey questions that research suggests could contextualize bias in results. (See current form of survey in appendix of this paper).
- limitations on the portions of the survey responses going to the administration (Part A).
    - "Part A of the Student Survey of Teaching must consist of statements that the union and management have agreed are more likely to lead to reliable student response." (2016-2020 Full-time Contract, p. 54)
- limitations on the use of surveys in the faculty evaluation process.

- ○ "These data can only be used by OAA to initiate a conversation with the department chair to discuss institutional and departmental trends." (2016-2020 Full-time Contract, p. 54)
  - ○ "The quantitative data collected from Part A of the Student Survey of Teaching shall not be included in any PDR [Professional Development Report], tenure application, or promotion application." (2016-2020 Full-time Contract, p. 54)
- a faculty-only committee to review and offer recommendations for labor-management negotiations about the student survey of teaching form and the evaluation process.
  - ○ "The Student Survey of Teaching form and process shall be annually reviewed by an all-faculty committee. This DUE committee will consist of faculty members from a range of academic disciplines, including at least two part-time faculty members, and including two faculty who specifically represent DUE. Hereafter, the committee is referred to as the Evaluation Committee." (2016-2020 Full-time Contract, p. 53)
  - ○ "Any revision to any portion of the form or process that is recommended by the Evaluation Committee and approved by the Office of Academic Affairs shall be formally negotiated between DUE and the College prior to implementation." (2016-2020 Full-time Contract, p. 53)

Since the student survey form is the same for all faculty, much of the same contract language is included in the part-time faculty contract. Additionally, the part-time faculty contract states: "[survey] data can be used to guide future professional development and shall not be used to initiate disciplinary procedures".

The current student survey of teaching form is provided at the end of this paper. The full-time and part-time contracts are publicly available at http://www.dutchessunitededucators.org/due-files.

Strengths and Weaknesses of Approach and Outcomes:
Strengths:
1. Our contract maintained a faculty evaluation system based primarily on peer-review by department chairs, in which peer observation of faculty in the classroom is central.
2. While the Office of Academic Affairs has access to a part of the student survey, those questions are framed carefully based on evidence found in scholarship.
   - The data that goes to Academic Affairs is mainly about procedures.
   - This section includes questions to help contextualize bias in results.
   - The qualitative question section remains in peer-review relationship, at the department level.

3. Surveys still cannot be used for "summative" evaluation. The administration cannot use surveys for promotion and tenure, though faculty can introduce qualitative results if they wish.

4. We maintained paper-based, in-class survey delivery.

5. There is an all-faculty standing union committee, consisting of "faculty members from a range of academic disciplines, including at least two part-time faculty members, and including at least two faculty who specifically represent DUE," that is evaluating new uses of surveys:

   ● This committee has made some adjustments to questions.
   ● So far faculty have not asked to substantially change new questions or procedures, which may indicate that the changes have not had negative effects.

Weaknesses:

1. Without continued education of faculty, chilling effects may occur without faculty even realizing it. Because surveys are normalized in higher education, and, more broadly, market-based norms of a consumer society are so pervasive, we may not be fully aware of how the surveys shape our teaching. For example, in disciplines in which faculty teach controversial topics, students exposed to new ideas at times experience discomfort or defensiveness, which may translate into assumptions about teaching effectiveness, particularly toward those professors who are not male and/or not white.

2. Department chairs, despite their role as peer-evaluators, are not immune to being influenced by problematic student survey data. For example, some DCC department chairs recently agreed to new usage of poorly designed student surveys in faculty job searches for some candidates for tenure-track positions.

3. Campus decision makers are often not from groups that regularly face the discrimination that appears in student surveys and thus tend to view the feedback as a fault of the faculty member rather than the inherent bias. As decision makers, this viewpoint is more problematic and can have dramatic negative impact on marginalized faculty.

4. Even though we have educated the administration about the problems with survey data, and they acknowledge that the data are questionable, they still insist on the importance of their access to the data without any rationale for why or how they would use the data.

Lessons Learned

It is still too early to fully understand the impact of these changes in the use of student surveys in our faculty evaluation system at Dutchess Community College, however our experience researching and negotiating this matter suggest to us these preliminary lessons:

● Though difficult, it is possible to challenge the dominant corporate narrative that shapes so much of higher education, including faculty evaluation.

L. Akins and L. Murphy, NCSCBHEP 2019　　　　　　　　　　　　　　　　　　　　　　7

- The excellent, important scholarship that colleagues are pursuing made it possible for us to challenge the imposition of these practices at DCC, and this scholarship is laying a foundation for challenging these practices where they are already in place.
- Educating faculty on the research concerning student surveys is at least as important as educating administrators so that all involved are making evidence-based decisions that impact the educational workforce.
- Silence agreements in collective bargaining make it difficult to mobilize faculty to support or oppose contract proposals that have significant bearing on the the faculty working conditions that are essential to the academic integrity of institutions.
- Negotiating for full-time and adjunct/part-time faculty together on this matter helped us to foreground the particularly vulnerable position that adjunct/part-time faculty are in, which allowed us to create policies that protected all faculty. This case study illustrates how addressing the exploitation of adjunct/part-time faculty strengthens the profession as a whole.
- It was our experience that without proper training in Interest Based Bargaining (IBB), if management and/or labor are not fully committed to the interest-based principle that uses reliable evidence for decision-making, the "interest" can be used to silence dissent.

Appendix: Articles Consulted for Contract Negotiations and Subsequent Evaluation Committee Work

- K. Mitchell, & J. Martin, "Gender bias in student evaluations," *PS: Political Science & Politics,* 1-5 (2018), https://doi.org/10.1017/S104909651800001X
- Michael Hessler, Daniel M Pöpping, Hanna Hollstein, et al, "Availability of cookies during an academic course session affects evaluation of teaching," *Medical Education* Volume 52, Issue10 (2018), https://onlinelibrary.wiley.com/doi/abs/10.1111/medu.13627
- Colleen Flaherty, "Teaching Eval Shake-Up: Most institutions say they value teaching. But how they assess it tells a different story," *Inside Higher Ed,* May 22, 2018, https://www.insidehighered.com/news/2018/05/22/most-institutions-say-they-value-teaching-how-they-assess-it-tells-different-story
- Michelle Falkoff, "Why We Must Stop Relying on Student Ratings of Teaching," *Chronicle of Higher Education,* April 25, 2018 https://www.chronicle.com/article/Why-We-Must-Stop-Relying-on/243213
- B. Uttl, C.A. White, and D.W. Gonzalez, "Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related." *Studies in Educational Evaluation*, 54 (2017), https://www.sciencedirect.com/science/article/pii/S0191491X16300323
- Henry A. Hornstein, "Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance," *Cogent Education*, 4:1, (2017), https://www.tandfonline.com/doi/full/10.1080/2331186X.2017.1304016?scroll=top&needAccess=true
- Anne Boring, Kellie Ottoboni, and Philip B. Stark, "Student evaluations of teaching (mostly) do not measure teaching effectiveness," *ScienceOpen Research* (2016) https://www.scienceopen.com/document?vid=818d8ec0-5908-47d8-86b4-5dc38f04b23e
- Colleen Flaherty, "Bias against female instructors," *Inside Higher Ed,* January 11, 2016, https://www.insidehighered.com/news/2016/01/11/new-analysis-offers-more-evidence-against-student-evaluations-teaching#.Ws_LBxwytC8.link
- Philip B. Stark and Richard Freishtat, "An evaluation of course evaluations." *ScienceOpen Research,* (2014) https://www.stat.berkeley.edu/~stark/Preprints/evaluations14.pdf
- L. MacNell, A. Driscoll, & A.N. Hunt, "What's in a Name: Exposing Gender Bias in Student Ratings of Teaching," *Innov High Educ* 40: 291 (2015) https://doi.org/10.1007/s10755-014-9313-4
- M. Braga, M. Paccagnella, & M. Pellizzarim, "Evaluating students' evaluations of professors," *ScienceDirect,* (2014) https://doi.org/10.1016/j.econedurev.2014.04.002
- Guy A. Boysen, Timothy J. Kelly, Holly N. Raesly and Robert W. Casner, "The (mis)interpretation of teaching evaluations by college faculty and administrators", *Assessment & Evaluation in Higher Education*, Vol. 39, No. 6, 641-656 (2014) http://dx.doi.org/10.1080/02602938.2013.860950
- Stephen L. Benton and Kenneth R. Ryalls, "Challenging Misconceptions About Student Ratings of Instruction", The IDEA Center*, *IDEA*, IDEA paper #58, (April 2016), https://www.ideaedu.org/Portals/0/Uploads/Documents/Challenging_Misconceptions_Abo

L. Akins and L. Murphy, NCSCBHEP 2019   9

[ut_Student_Ratings_of_Instruction.pdf](https://...)[4]

- Susanna Calkins and Marina Micari, "Less-Than-Perfect Judges: Evaluating Student Evaluations," *Thought and Action*, (Fall 2010),

  http://www.nea.org/assets/img/PubThoughtAndAction/TA10CalkinsMicariR.pdf

- Heather Laube, Kelley Massoni, Joey Sprague, Abby L. Ferber, "The Impact of Gender on the Evaluation of Teaching: What We Know and What We Can Do" *NWSA Journal*, (Fall 2007)

- James England, Pat Hutchings, and Wilbert J. McKeachie, "The Professional Evaluation of Teaching" *American Council of Learned Societies*, Occasional Paper No. 33, (1996), http://archives.acls.org/op/33_Professonal_Evaluation_of_Teaching.htm

---

[4] From their website (http://www.ideaedu.org/), "IDEA is a nonprofit organization dedicated to improving student learning in higher education through analytics, resources, and advice." They are a vendor of SETs.

Appendix: Current Form of Student Survey of Teaching

Class Climate	Student Survey of Teaching	SCANTRON.
		DUTCHESS COMMUNITY COLLEGE

Mark as shown: ☐ ☒ ☐ ☐ ☐ Please use a ball-point pen or a thin felt tip. This form will be processed automatically.

Correction: ☐ ■ ☐ ☒ ☐ Please follow the examples shown on the left hand side to help optimize the reading results.

1. Instructions

This form provides an opportunity to give feedback about this class. Your instructor will tell you how to fill out the form and should then leave the room. After final grades are turned in, these forms will be reviewed by the instructor with his/her supervisor. Your participation is important to and valued by the College.

2. Reason

2.1 Reason for taking this course: ☐ Elective ☐ Requirement

3. Part A of the Student Survey of Teaching Form

3.1	The instructor provided a syllabus, either hard copy or electronically, that included a course outline. (Check Not Applicable if LAB section only.)	☐ Yes	☐ No	☐ Not Applicable
3.2	The instructor provided his/her DCC email address and how best to contact them outside of class.	☐ Yes	☐ No	

Use the scale below to express your opinion on each of the statements listed. Enter your response below.

		Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree	Not Applicable
3.3	The instructor clearly communicated how the course grade was to be determined.	☐	☐	☐	☐	☐	☐
3.4	The instructor usually begins class on-time as scheduled.	☐	☐	☐	☐	☐	☐
3.5	The instructor usually uses the instructional time available.	☐	☐	☐	☐	☐	☐
3.6	The instructor expects students to use the required course materials (such as textbooks, online resources, films, software).	☐	☐	☐	☐	☐	☐
3.7	The instructor is responsive to students' questions and concerns.	☐	☐	☐	☐	☐	☐
3.8	The instructor gives clear explanations.	☐	☐	☐	☐	☐	☐
3.9	I had a strong desire to take this course.	☐	☐	☐	☐	☐	☐
3.10	I was able to access the resources required for the course.	☐	☐	☐	☐	☐	☐
3.11	I was able to commit the time and effort necessary to be successful in the course.	☐	☐	☐	☐	☐	☐
3.12	I feel my prior knowledge in the subject matter was a significant factor in my ability to be successful in the course.	☐	☐	☐	☐	☐	☐

F217U0P1PL0V0

02/12/2018, Page 1/2

L. Akins and L. Murphy, NCSCBHEP 2019

11

Class Climate	Student Survey of Teaching	SCANTRON.

4. Part B of the Student Survey of Teaching Form

Enter your comments for each of the prompts provided below.

4.1 Please comment on your impression of the instructor's teaching effectiveness.

4.2 Please comment on the organization and structure of the course.

4.3 Please comment on your interactions with the instructor.

4.4 Please comment on the text book and materials used in the class.

4.5 Please add any other comments you would like to make, including your overall summary of the course and suggestions for improvement. Give examples where you can.