Engineering Faculty's Perception of Student Bias in Higher Education

Kyra Twohy
Department of Mechanical and Aerospace Engineering
University of Dayton
Dayton, OH 45469
Email: twohyk1@udayton.edu

Abstract

This study investigated engineering faculty member's perceptions of student bias in a variety of classroom settings, specifically from the point of view of both male and female students. The goal was to show perceived differences between how male and female students rate professors. This is especially relevant as teaching evaluations are a crucial part of tenure and promotion cases. The study used a survey comprised of 25 questions that assessed the level of positive and negative bias for male and female students when evaluating teaching. Forty-six faculty members from the engineering department of a midwestern university completed the survey. There was also an interview that asked 8 questions about each faculty member's experience with bias and gender bias. Of the forty-six faculty members who completed the survey, 10 were interviewed. Data from the survey was analyzed using one-way ANOVA tests, as well as descriptive statistics. Qualitative data was obtained from the coded interviews. Results showed that the faculty's perceptions positioned male students as more positively biased when evaluating male faculty and were more negatively biased towards female faculty, than their female counterparts. Class size also was found to contribute to bias, as class size increases led to increasing perceived negative bias. This study finds that more research should be done to establish a method of removing bias from teaching evaluations, so that course evaluations only reflect teaching ability.

Introduction

Debates over fair methods of measuring a teacher's success have been a source of division in both education research and politics, especially with the redesign of evaluation systems as part of Race to the Top[1]. In secondary education, methods of teacher evaluation include student outcomes on standardized tests, in-class evaluations, and teacher portfolios[2]. However, in higher education a teacher's evaluation can be solely based on student evaluations[3]. Most universities have implemented an online survey that students are asked to complete at the end of the course[4]. This survey asks a variety of questions that are aimed at assessing a given teacher's effectiveness in teaching that class's source material[3].

In tenure evaluations and promotion cases, teaching evaluations are part of the package presented to the committee that determines the outcome of the promotional case[5]. The weighting of teacher evaluations may vary from university to university, but they are always considered for tenure positions that involve classroom teaching. Teaching evaluations that do not meet the criteria for a given department at a given university can lead to tenure or promotion being denied.

However, research has also shown that student bias can factor into these evaluations, which could lead to a promotional structure that favors specific candidates based on evaluation scores. A study showed that the evaluations from a person who only watched 30 seconds of a teacher in practice predicted the students' and colleagues' evaluations of that teacher. This implies that visual perceptions of a teacher even over a small period of time can influence evaluations that take place at a much later date[6]. Other studies have purposely activated stereotypes in evaluations of teaching based on exclusively auditory input. It was found that the professor was rated higher if the student believed the person they had listened to was a man under the age of 35, regardless of what the actual age or gender was on the recording[7].

Similarly, blinded studies have specifically targeted evaluation differences between genders. Studies have shown that male students will score teaching from a male professor significantly

2

higher than they score teaching from a female professor[9]. This is compared to female students that score teaching from male and female professors approximately the same[10]. These findings were for teaching through an online forum[9], where gender was only assumed based on the name of the professor. In addition, when rated by students on qualities such as scholarship and overall teaching ability, female professors receive more negative scores from both male and female students[8].

Race has also been investigated as a prevalent form of bias in higher education. In a study of the evaluations of over 3,000 professors on RateMyProfessor.com, it was found that racial minorities, specifically Black or Asian professor, had more negative evaluations than their peers[11]. The presence of multiple forms of bias could then influence the possibility of promotion, unless review committees or the professors themselves are actively compensating for this.

Although faculty members are in general aware of the prevalence of this bias, the impact is varied. A study showed that female faculty members are more negatively impacted by student evaluations and are more likely to self-select negative emotions relating to "unhappy" or "mad" when reading student evaluations[12]. However, student feedback is generally found to be helpful across higher education[13]. This means that faculty members thought students had enough basis to assess the quality of instruction and incorporated the students' feedback into restructuring courses. They remained neutral on whether these evaluations should play a role in tenure decisions. Current studies of faculty perception have not investigated how perceptions change in different subject areas and they do not differentiate between the evaluations of male and female students.

Given this problem space, this research study explores the faculty perceptions of student bias in course evaluations. Specifically, the influence of faculty gender and position level within engineering on perception was investigated through this study. Faculty perception was broken down into their views on the bias of their male and female students.

Methods

This study was conducted at a private Midwestern university in the college's engineering department, which has a faculty population of approximately 200 people. Roughly, 10% of the faculty population is female. The size of the university allows for diversity of staff, both in gender and position level. This provided a large enough sample size, specifically of women, for any justifications to be statistically significant. After receiving consent from the dean's office of the engineering department, every faculty member in the engineering department was invited to participate in the survey through an email sent from the Dean's Office. The email included a brief description of the study and a link to the online survey, completed through Qualtrics. The initial screen introduced the study and its purpose before asking for consent. Instructions were given before the questions and participants were informed that they could exit the survey at any time. All partially completed surveys were included in the final data set. The survey was administered over a two-week period in April to guarantee that all respondents had a minimum of 3 months of teaching experience at the private university, and more than likely at least one cycle of course instructor feedback.

Forty-six faculty members completed all or part of the online survey. They were composed of 8 female faculty members and 38 male faculty members. They also had a range of positions,

3

including 7 junior faculty members, 29 senior faculty members, and 10 non-tenure track research or teaching professors. All participants were asked to verify that they were over the age of 18 and to give consent before taking the online survey. Break down of demographics is shown in Figure 1. The sample was limited to engineering faculty in order to focus on an area where bias is widely sighted an extremely prevalent.
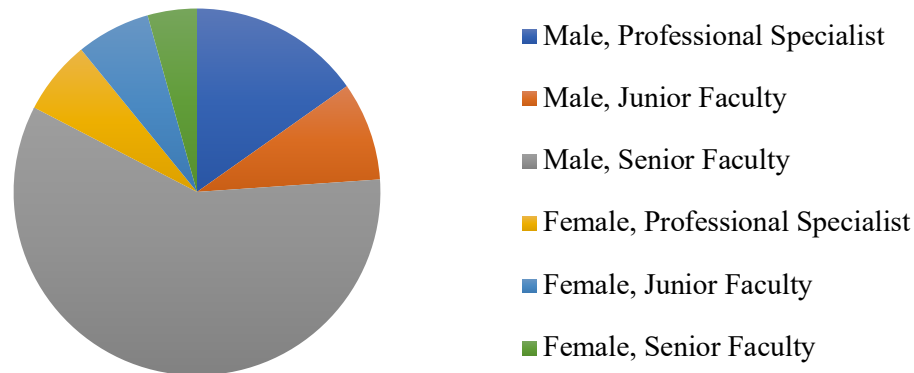

Figure 1: Demographic data for participants in the survey.

The first 2 questions were asked to gather descriptive information on the participants. The next 22 questions assessed the perceptions of student bias by asking the participants to rate the level of bias in given situations from the perspective of male and female students. These situations included professor gender (male or female), class size (1-30 students, 31-60 students, or 60+students), professor tenure status (senior faculty, junior faculty, or professional specialist), and class discipline (science, engineering, or philosophy). They were scored on a 5 point Likert scale, from high negative bias to high positive bias. Questions were adapted from similar studies assessing faculty perceptions of student evaluations[12]. The final question required participants to rank the factors that generate bias from most prevalent to least from the list of faculty gender, tenure status, faculty teaching experience, class size, and class discipline. They were given a space to justify this selection.

At the end of the survey participants were asked to mark if they were interested in being interviewed as part of the study as well. Eleven participants marked that they were willing to take part in a 15 minute interview and of those 11, 10 were interviewed. Before the interview was completed the participants were asked to complete a separate consent form. They were informed of their right to speak off the record at any time and ability to choose to answer or not answer any questions asked during the interview. The recorded interview was approximately 15 minutes long and was composed of a script of 8 questions. The introductory question established the teaching experience of the faculty member at the current institution, as well as any previous institutions. The next three questions explored the faculty member's opinions and experience with student bias. Specifically, participants were asked which form of bias is the most prevalent in higher education, if they had ever experienced bias from students, and if there are any implications of bias on their career. The last set of questions focused specifically on gender bias. Participants were asked if they had personal experience with gender bias, if they take action to avoid gender bias, and if they had ever felt biased to a professor during their own higher education career. All interview questions were worded to be open-ended, allowing for a variety

of answers. The interviews were digitally transcribed using Temi online software and the transcripts were stored in a Box folder.

The survey data was analyzed in SPSS. For the Likert scaled questions, no bias was equivalent to a zero, while high positive and negative bias were associated with a positive 2 and negative 2 respectively. The difference between the level of bias from a male student versus a female student was taken for each situation in order to run statistical tests. Descriptive statistics were generated for the last question, in which participants ranked the perceived influence of different biases on teacher evaluations. They were also used to obtain information on the general population studied, from the first two questions. One-way ANOVA tests were conducted on the difference variables created after the study was completed using both faculty gender and faculty status as the comparison factor.

The transcribed interview data was coded for presence of specific forms of bias, visual methods to combat bias, and personal satisfaction with course evaluations based on findings in similar studies[12]. All interviews were coded by the principal investigator and another coder to ensure reliability to the coding scheme and between coders.

Results

The calculated mean values for the perceived amount of bias from both male and female students for specific situations are given in Table 1. The faculty rated the perceived bias on a Likert scale from negative 2 to positive 2, where high negative bias corresponds to negative 2 and high positive bias corresponds to positive 2. In general, faculty perceived that male students had a negative bias, if any bias at all. The two exceptions to this were larger positive bias for male students evaluating male faculty members and male students evaluating teachers of small classes.

Table 1: Mean Values of Perceived Bias for Male and Female Students

Bias Category		Male Students	Female Students
Tenure Status	Professional Specialist °	-0.32	-0.13
	Junior Faculty	-0.08	0.08
	Senior Faculty #	-0.03	0.08
Gender	Male Faculty °	0.35	0.19
	Female Faculty	-0.32	0.00
Class Discipline	Engineering Class	-0.03	-0.11
	Science Class	-0.03	-0.06
	Philosophy Class *	-0.17	0.11
Class Size	Small Class (1-30 students)	0.51	0.54
	Medium Class (31-60 students)	-0.08	-0.11
	Large Class (61+ students)	-0.58	-0.63

° $p < 0.1$, * $p < 0.05$, # $p < 0.01$

Perceptions of female students' course evaluations remained on the same magnitude as their male counterparts' evaluations but are more polarized. In cases where male students were perceived to have negative bias, such as large class sizes, the female students were perceived to have more negative bias. This also applied to medium classes, science classes, and engineering classes in the negative direction, as well as small classes in the positive direction. Occasionally the sign actually switches when comparing perceived bias in male and female student evaluations, including perceptions of evaluations of female faculty and professors of philosophy classes. In these cases, male students were assumed to have negative bias, while female students were assumed to have positive bias.

The average rank of the five forms of bias given as Question 24a was also calculated and shown in Table 2, where 1 indicates the most prevalent form of bias and 5 indicates the least prevalent form of bias. Tenure status had a much lower rank than the other forms of bias. In general, this was due to faculty perception that students did not know the tenure status of their professors, so there could not be bias in this area. Gender and class size had the highest average of rank and then there is a drop to experience level and discipline/subject.

Table 2: Average Rank of Specific Bias

Gender	Class Size	Disciple/Subject	Experience Level	Tenure Status
2.59	2.62	2.71	2.79	4.29

The ANOVA tests showed statistical significance only for the difference between how male and female faculty members perceived the bias of male and female students in large classrooms. Male faculty (M=0) perceived no difference in bias from male students to female students, as compared to female faculty (M=0.2) who perceived more bias from female students in large classrooms, $F(df)=7.071$, $p<0.05$. All other differences between male and female faculty member's perceptions of student bias were not significant, with $p>0.4$.

Although, in the survey faculty responded that there was bias present, specifically for gender, the interview data did not reinforce this finding. Most faculty members said that they did not experience gender bias. Even if they had experience with gender bias in the past, they were quick to emphasize that it was only small offenses and this type of bias no longer occurred in their course evaluations. One faculty member said, "I've seen it in little ways. I would say it's not overwhelming and I feel like as I've gone further in my career, I see less of it in my evaluations."

This could be related to the general faculty who agreed to participate in the interview portion of the study. Of the ten faculty members interviewed, five were male, senior faculty members. The rest of the interviewees were two female non-tenure track faculty, two male junior faculty members, and one female junior faculty member. This is a biased sample to those who have received numerous course evaluations and presumably achieved tenure based in part on their teaching practice. Even the professors who have not achieved tenure mostly pointed out that they also received sufficient evaluation scores. Overall, eight expressed satisfaction with their course evaluations, while one expressed dissatisfaction and one did not indicate their feelings towards their own evaluations.

Seven faculty members expressed different ways they counteract bias in the classroom. Two of the three female faculty members mentioned clothing choice as a method of counteracting bias.

One said, "when I started…I would dress more masculine, business casual. I wouldn't wear frilly dresses, because I didn't want to kind of bring out the femininity." Others specified grading techniques that were blinded from student name or gender in order avoid bias or general procedures that were instituted to make sure all students were part of a safe learning environment that did not cater to a specific type of student.

Widely mentioned throughout the interviews, were class size, race and age of the professor. Each were mentioned in over half of the interviews, with class size addressed the most often in eight interviews, followed by race and age at 6 and 5 times, respectively. Similar to the survey information, professors expressed that larger classes lead to more negative feedback. They described this phenomenon by saying, "the student doesn't have the feeling like they're not being cared for as much and there's less personal connection. They start to feel a disconnect and so they rate people a little bit lower." The other form of bias that was frequently mentioned was race, specifically the difference between international and domestic faculty members. One faculty member thought this was more prevalent in younger students, as they would not have had experience with international teachers in high school. They postulated that these students, "expect the onus is on the faculty member to change and not the other way around."

Discussion

This study investigated the perceptions of student bias in course evaluations from the perspective of engineering faculty. The bias of male and female students was evaluated individually to allow for distinctions to be made for student gender. Faculty members who completed the survey were asked to indicate their gender and tenure status so similar distinctions could be made for faculty members. The perception of bias was assessed in a variety of situations, including faculty gender, faculty tenure status, class size, and class subject for both positive and negative bias. Faculty members were also asked to rank specific forms of bias by which is the most prevalent.

In the data shown in Table 1, the mean perceived evaluation shows differences in how faculty think male and female students rate different types of teachers. For the evaluations of different faculty positions, the general perception is that female students have less negative bias. This is not to say that female students are perceived to have less bias because faculty perceived that female students had more positive bias than male students have negative bias. In terms of application, this indicates that professors should focus on ways to diminish negative responses from male students based on promotional level. These results insinuate that male students respond more positively to more authority, so course evaluation scores could be increased by presenting authority in a different way.

This assessment does not apply to differences in how male and female students score professors of different genders. Faculty found that female students showed less overall bias, regardless of gender, while male students were seen as polarizing by the faculty who completed the survey. This is especially important because the undergraduate population in engineering programs is comprised of more male students than female students. If male professors experience positive bias from a larger percentage of the population, while female professors simultaneously experience greater amounts of negative bias, the gap between average scores for male and female professors would grow. Because the professors perceive differences between how male and female students score them in course evaluations, stress due to the implications of lower scores from the predominant gender of their students could increase. However, the results of the

7

interviews with female faculty members did not support this and instead only supported small slights that, in general, did not impact overall teaching evaluation scores.

When the survey asked about bias in the form of class dynamics, including size and subject matter, the trends are different. In both cases, the gender of the student that is perceived to have the most bias, switches depending on the class or the subject. While faculty perceived that female students would be more negatively biased in evaluation of engineering faculty members, the reverse was true of a professor teaching a philosophy class. This indicates a different perception of the classroom environment for different subjects. For the purpose of increasing teaching evaluation scores, engineering teachers could incorporate common teaching practices from other disciplines to which a greater number of students would respond positively. This is important not only to increase student evaluation scores, but also to increase faculty members' satisfaction with their scores. Although, satisfaction was present in most of the interviews, large assessments of bias, especially negative bias, indicates that faculty members are aware of inequalities. Based on the population who completed the survey, many acknowledged the presence of the bias in situations where they would not have personal experience. This highlights the split between using evaluations to influence future iterations of the course and using them for performance evaluations that influence tenure decisions[12].

Although only small differences occurred between male and female students, female students were seen to be more positively biased in small classes and more negatively biased in large classes. This indicates that small classes are perceived to create a better environment for the student and teacher. Investing in more faculty members to decrease class size could result in increased course evaluation scores or perceptions of evaluation scores. An important note about these descriptive statistics is the population of the survey, half the participants are male, senior faculty members, statistics that do not desegregate these descriptors, such as ANOVA tests, would be weighted to the perceptions of this population.

The average rank of the different forms of bias as assigned by the faculty members, showed other signs of perceptions of bias. Tenure status received the lowest average rank insinuating that the faculty members do not think that students are aware of the tenure status of their professors and if they are aware it does not heavily influence their evaluations. This is intriguing because of the weighting of evaluation scores that go into a faculty member receiving tenure[3].

As seen with the ANOVA data, the small sample size, especially of female faculty members, influenced the data of this study. Future research would expand the sample pool to verify that all findings are statistically significant. An expanded sample pool should include more universities of a variety of sizes and institution types. Another further form of research would be to give the same test to other STEM faculty, including science and math professors at the same university. This could help to strengthen results across the field or show differences between STEM disciplines that are already perceived by engineering faculty.

The ANOVA tests also showed statistical difference between how female and male faculty perceive student evaluations from students of opposite genders. The only statistically significant difference was in how male and female faculty perceived bias from students in large classrooms. This is due to one female faculty member scoring moderate negative bias from female students in large classes, as compared to all other participants who perceived no difference between male and female student's bias in large classrooms. Although, there is a statistically significant

difference in the means of the male and female students the sample size of female faculty was low. This allowed the data from one participant to influence the analysis. However, this phenomenon is also present in course evaluations. In this case a minority faculty member influenced the data because perceptions were compared across gender. If course evaluations were separated based on attributes of the student population, including gender, the bias of the majority could be counteracted. This could diminish the influence of negative bias on specific groups within academia and could also highlight areas of growth for specific teachers. If specific groups feel disenfranchised in the classroom, but the population is low, their influence on mean course evaluation scores could be small enough to not influence a change. This research also helped to highlight the bias of faculty members that should be considered with the intention of avoiding these pit falls in the future.

Conclusion

As expected, faculty perceived statistical differences in how male and female students evaluate their professors in a variety of situations. The differences in the perception of how male students score male and female faculty was the largest difference. These perceptions could influence classroom dynamics, as professors, specifically female professors, are actively adjusting their classrooms to avoid bias. Methods of counteracting bias mentioned in the interviews were aimed at increasing the perceived authority level of the female faculty members in the eyes of their students. Similar to gender differentiation in the STEM industry, authority, whether perceived or in practice, and its implications create differences both in actual evaluation scores and the perceptions of them. This leaves room for improved methods of considering teaching evaluation scores within the context of faculty gender, student population attributes, as well as other forms of bias including class size.

References

1. Hunt, K., Gurvitch, R., & Lund, J. L. (2016). Teacher evaluation: Done to you or with you? *Journal of Physical Education, Recreation & Dance, 87*(9), 21-27.
2. Millman, J., & Darling-Hammond, L. (1991). *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers* Corwin Press.
3. Goos, M., & Salomons, A. (2017). Measuring teaching quality in higher education: Assessing selection bias in course evaluations. *Research in Higher Education, 58*(4), 341-364. 10.1007/s11162-016-9429-8 Retrieved from https://doi.org/10.1007/s11162-016-9429-8
4. Greenlee, L. R. (2014). *A tale of two gender roles: The effects of implicit bias on the perception of others* (Ph.D.). Available from GenderWatch, ProQuest Dissertations & Theses Global. (1620539583). Retrieved from https://search-proquest-com.proxy.library.nd.edu/docview/1620539583?accountid=12874
5. Chen, Y., & Hoshower, L. B. (2003). Student evaluation of teaching effectiveness: An assessment of student perception and motivation. *Assessment & Evaluation in Higher Education, 28*(1), 71-88. 10.1080/02602930301683 Retrieved from https://doi.org/10.1080/02602930301683
6. Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology, 64*(3), 431-441. 10.1037/0022-3514.64.3.431
7. Arbuckle, J., & Williams, B. D. (2003). Students' perceptions of expressiveness: Age and gender effects on teacher evaluations. *Sex Roles, 49*(9), 507-516. 1025832707002 Retrieved from https://doi.org/10.1023/A:1025832707002

8. Basow, S., & Silberg, N. (1987). Student evaluations of college professors: Are female and male professors rated differently? *Journal of Educational Psychology, 79*(3), 308-314. 10.1037/0022-0663.79.3.308
9. MacNell, L., Driscoll, A., & Hunt, A. N. (2015). What's in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education, 40*(4), 291-303. 10.1007/s10755-014-9313-4 Retrieved from https://doi.org/10.1007/s10755-014-9313-4
10. Maura Ryan, Timothy Judge, James Brockmole, Mark Gunty, Kevin Barry, James   Frabutt, . . . Christina Wolbrecht. (2017). *Final report of the ad hoc committee on the evaluation of teaching.* (). University of Notre Dame:
11. Reid, L. D. (2010). The role of perceived race and gender in the evaluation of college teaching on RateMyProfessors.com. *Journal of Diversity in Higher Education, 3*(3), 137-152. 10.1037/a0019865
12. Kogan, L. R., Schoenfeld-Tacher, R., & Hellyer, P. W. (2010). Student evaluations of teaching: Perceptions of faculty based on gender, position, and rank. *Teaching in Higher Education, 15*(6), 623-636. 10.1080/13562517.2010.491911 Retrieved from https://doi.org/10.1080/13562517.2010.491911
13. Nasser, F., & Fresko, B. (2002). Faculty views of student evaluation of college teaching. *Assessment & Evaluation in Higher Education, 27*(2), 187-198. 10.1080/02602930220128751 Retrieved from https://doi.org/10.1080/02602930220128751