



Research article

Does professors' gender impact how students evaluate their teaching and the recommendations for the best professor?



Arturo Arrona-Palacios^a, Kingsley Okoye^a, Claudia Camacho-Zuñiga^b, Nisrine Hammout^c,
Emilia Luttmann-Nakamura^d, Samira Hosseini^{a,e,*}, Jose Escamilla^f

^a Writing Lab, TecLabs, Tecnológico de Monterrey, Monterrey, Mexico

^b School of Engineering and Sciences, Tecnológico de Monterrey, Toluca Campus, Mexico

^c Mines- School of Industrial Management, Mohammed VI Polytechnic University, Benguerir, Morocco

^d Institutional Effectiveness Department, North Region ECOA, Tecnológico de Monterrey, Mexico

^e School of Engineering and Sciences, Tecnológico de Monterrey, Monterrey Campus, Mexico

^f TecLabs, Tecnológico de Monterrey, Monterrey, Mexico

ARTICLE INFO

Keywords:

Education
Student evaluation of teaching
Gender bias
Higher education
Educational innovation
Performance evaluation

ABSTRACT

This study examined the impact of the professors' gender according to a student evaluation of teaching (SET) in a private university. The study took place in a private university (n = 103,833) on six different campuses in the north region of Mexico. The distribution of the professors' gender was analyzed according to semesters, campuses, and schools. Our findings suggested that when undergraduates evaluated their professors on specific criteria concerning teaching performance, they expressed their opinion regardless of the professors' gender. However, when being asked for a single overall evaluation, as whether they would recommend the professor as one of their best professors, the students tended to favor male professors over their female peers by a slight margin. While such perceptions might not be representative of the actual teaching quality, it would be interesting in the future to delve deeper into the causes of possible biases.

1. Introduction

Student evaluations of teaching (SET) is an important assessment tool to evaluate the quality of teaching and provide feedback on the perceived teaching effectiveness of faculty members, as well as to report useful information to administrators (Boring, 2017). These evaluations are typically administered by institutions and completed by students anonymously. Universities use SET when considering promotions, long-term contracts, merit and award-related decisions, salary increases, and contract renewals for their faculty members and staff (Davis, 2009). However, one of the negative aspects of the application of SET is that students become the sole subjective evaluators of the professors' productivity. Previous research has reported that students' criteria for judging their professors are, in part, exogenous or unrelated to actual teaching qualities (Boring, 2017; Kristof De Witte & Rogge, 2011; McPherson, 2006). Despite the widespread use of SET, when students evaluate subjectively, they possibly could be stereotyping or expressing biases. This factor should be of concern to faculty members and universities because SET results may influence the overall assessment of their teaching and

instruction effectiveness (Barth, 2008; Marsh and Bailey, 1993; Wiley, 2019).

However, other research results have indicated several positive associations. For example, they reported higher average SET ratings when the professor was viewed by students as knowledgeable, friendly, clear, enthusiastic, and fair (Barth, 2008; Hills et al., 2009; Tang et al., 2005; Ogier, 2005). Also, professors who use humor and are physically attractive receive higher SETs scores (Felton et al., 2004; Fortson and Brown, 1998; Freng and Webber, 2009). On the other hand, disorganization, lack of clarity in teaching and inaccessibility were linked to lower SET scores of the professors (Barth, 2008; Sitzman, 2010).

Regarding the perception of SET by faculty, literature has shown divided opinions. Some perceive SET to be unreliable or invalid and a measure of popularity rather than of effective teaching. Others consider it meaningful and reported they have made improvements based upon such results (Annan et al., 2013; Balam and Shannon, 2010; Beran and Rokosh, 2009). Recently, Spooren and Christiaens (2017) reported that students have a positive view of SET practice because they agree it can provide accountability for teaching quality. However, conclusions from

* Corresponding author.

E-mail address: samira.hosseini@tec.mx (S. Hosseini).

student responses about the use of SET reports for administrative decision-making remain unclear.

In recent decades, higher education sector has shifted towards a more business-oriented model of operation (Mazzarol et al., 2003). Such shift of paradigm has made the demonstration of institutional quality to become the main priority and a constant routine in academic life. For that reason, it is unlikely that the current and common usage of SETs as a tool for measuring the quality of teaching will diminish (Blackmore, 2009) even though there is evidence suggesting the opposite (Uttl et al., 2017). Therefore, due to the major reliance on SETs and their impact on career advancement, any potential gender bias in these evaluations is still a matter of great concern (MacNell et al., 2015).

As examples of such gender bias, students often expect their male and female professors to behave in different ways or to act according to stereotypes (Anderson and Smith, 2005). Men are commonly characterized by ambition, domination, and independence, while women are often characterized for their compassion, emotional expressiveness, and other care-related characteristics (Koch et al., 2015). Students may hold their professors accountable to these gender-based behaviors and can even be critical of those professors who violate these expectations (Dalmia et al., 2005; Sprague & Massoni, 2005). On the contrary, students perceive professors fulfilling their expectations of gender more favorably (Andersen & Miller, 1997). Moreover, female professors not meeting such gender-based expectations are viewed less favorably; while male professors not exhibiting strong interpersonal traits are not categorized the same (Basow and Montgomery, 2005).

Several studies have shown that women receive lower SET assessments than their male colleagues (Heckert et al., 2006; Tatro 1995). MacNell, Driscoll, and Hunt (2015) showed that students rated female professors more harshly than male professors which suggested that the former would have to work harder than the latter to receive comparable ratings. In another study, female professors were less favorably rated in SET scores, but the results heavily depended on the degree program. For example, in engineering courses, they received the lowest rating; while, in business courses, they received the highest scores (Bianchini et al., 2013). Wagner, Rieger, and Voorvelt (2016) suggested a gender bias against female professors in SET and therefore, using them in hiring and promotion decisions might put them under disadvantage. Also, Mengel et al., (2019) found that female professors systematically received lower evaluations from both female and male students. This finding was stronger for male students and junior female professors in general; but, particularly, those in math-related courses consistently received lower evaluation scores. They found no evidence that these differences were driven by gender disparities in teaching skills.

In a different address, Heck et al. (2002) showed that female professors received higher SET ratings, particularly with respect to the quality of instruction and the ability to communicate. One explanation for these mixed results with respect to gender may be physical attractiveness. In two studies with large samples taken from RateMyProfessor.com and performed across different university disciplines, a significant positive correlation was reported between the perceived attractiveness of the professors and their corresponding ratings (Freng and Webber, 2009; Hamermesh and Parker, 2005). Another reason for the mixed results may be the interactions between genders and the types of classes that are being taught (Bianchini et al., 2013).

Studies in Mexico have been scarce. To our knowledge, there are only a few papers analyzing gender bias in SET scores. The first study carried out in Mexico City by Martínez-González et al. (2012) studied a large sample ($n = 14,646$) of medical students from a SET questionnaire of their own creation during the academic year of 2007–2008. They showed female professors obtained better scores than their male counterparts; however, the ratings depended on the subject (Cellular Biology and Microbiology for females and Anatomy and Pharmacology for males). The second study was carried out in Ciudad Juarez, Chihuahua by Ayalá-Hernández (2013) as a qualitative study. The sample were 30 undergraduates and the findings indicated that the professor's personality,

the teaching skills, the type of course, and the grades obtained by the students had a significant impact on SET. The third study was performed in Baja California by Arámburo-Vizcarra and Luna-Serrano (2013), with a sample of 44,599 participants during four semesters in 2006–2007. According to their results, male professors were better rated than female professors; furthermore, those with eight or more years of experience and, especially, those belonging to the Mexican National System of Researchers (Sistema Nacional de Investigadores in Spanish) received higher scores. The fourth study was carried out in Monterrey, Nuevo Leon by Galván-Salinas and Fariás-Martínez (2018) with a sample size of 9,300 undergraduates in the academic year 2015–2016. They reported that professors were better rated by students if their failing index was low, and female professors were more likely to score higher than male professors. Lastly, Arceo-Gomez and Campos-Vazquez (2019) studied a large sample ($n = 600,000$) of university students from a platform called MisProfesores.com (the Mexican equivalent of RateMyProfessors.com) from 2008 to 2018 and found that female professors received lower scores than their male counterparts. They found that students commented on the appearance and personalities of female professors and referred to them often as “bad” or “strict.” They also reported students referred to women in less respectful terms, calling them “teacher”; but calling men “professor” or the title corresponding to their academic degrees; moreover, they used less positive language for female professors (“good”) when compared to their male peers (“excellent”).

As we can see in the aforementioned studies in Mexico, the analyses of SET scores according to gender indicate bias against and in favor of female professors. This fact seems contradictory because in Mexico the gender stereotype and machismo are applied to a large extent (Arciniega and Anderson, 2008; Mena and Rojas, 2010). Therefore, this topic demands further and deeper research.

The literature about gender bias reflected in SET scores has only focused on specific issues and teaching abilities; however, to our knowledge, there have not been studies looking at how gender affects the recommendations of students for “best professor”. Furthermore, studies on SET evaluations and gender bias are still scarce in Mexico, and they have generally concentrated on a specific city or state. For this reason and based on previous research, we considered three hypotheses for this study. Hypothesis 1 predicts a main effect of professors' gender, considering male professors receiving higher ratings on the evaluation criteria. Hypothesis 2 predicts a main effect of university schools, with the school of engineering and science, architecture and design and medicine and health sciences receiving higher ratings in the evaluation criteria and school of social sciences and government, humanities and education and business the lowest. With respect to determining the impact of the professors' gender on the students' recommendations for the best professor, Hypothesis 3 predicts male professors receive higher ratings than female professors.

2. Method

2.1. Participants

The sample was comprised of faculty professors and students (from their first semester to their last semester of their careers) from a private university in Mexico. Even though this private university has several campuses across Mexico, in this study, only data generated by campuses in the north region of the country were analyzed (Tampico, Monterrey, Saltillo, Laguna, Chihuahua, and Ciudad Juarez). A total of 103,833 surveys were answered by students during three semesters: January–May 2017 (28,091 surveys), August–December 2017 (47,120 surveys), and January–May 2018 (28,622 surveys). The surveys were administrated at the end of the semesters before the classes had ended. Furthermore, a total of 5,083 faculty members were evaluated: 1,522 from January–May 2017; 2,110 from August–December 2017, and 1,541 from January–May 2018. The survey covered six schools, namely: Engineering and Science, Architecture and Design, Medicine and Health Sciences, Humanities and

Education, Social Sciences and Government, and Business; with a total of 78 departments and 1,082 courses. The inclusion criteria considered for this sample was the following: those students who completed the survey at a 100 percent. Informed consent was obtained from all individual participants included in the study, as well as the inform consent of the school administrators for obtaining access to this data. In addition, the research project was approved by the Institutional Research and Ethics Review Committee from the Office of the Vice President for Research and Technology Transfer from Tecnológico de Monterrey, and it complied with the principles of the Declaration of Helsinki of research on human participants.

2.2. ECOA evaluation

For this study, the Student Opinion Survey (ECOA, for its acronym in Spanish: Encuesta de Opinión de Alumnos) was used. Otherwise known as the Students Evaluation of Teaching (SET) (Boring, 2017), it is a survey designed and applied in Spanish which collects opinions from university students about their professors' performance with regards to the quality of the courses delivered, the competencies of their professors, and the academic services offered across the different campuses (Dirección de Servicios Académicos, 2018). The ECOA is an institutional survey elaborated and owned by the private university and has been used for several years within the university for the evaluation of the professors' performance. It has been validated in previous studies (Ayala-Hernández, 2013; Galván-Salinas and Farías-Martínez, 2018; Montemayor-Gallegos, 2002) and has shown good scale reliability with a Cronbach's alpha value of 0.89 (Montemayor-Gallegos, 2002). In the current study, the overall scale reliability statistics indicated that Cronbach's alpha was 0.97.

The evaluation criteria of the ECOA include seven questions that assess different aspects of the professors' performance: (1) *methodology and learning activities used by the professor during the course*; (2) *applied concepts taught in terms of their application in the real world*; (3) *their advising role*; (4) *the evaluation and grading throughout the semester*; (5) *the applied intellectual challenge in teaching*; and (6) *the learning guidance offered by the professor*, and (7) *the question: "Would you recommend this professor to other students as the best professor?"* The answers were recorded through an

eleven-point Likert-type scale and scored from 0 to 10, where 0 meant dreadful and 10 meant exceptional.

2.3. Data analysis

A Chi-squared test was used to analyze the distribution of the professors' gender for the different semesters, campuses and schools. Furthermore, a multivariate analysis of variance (MANOVA) was executed to know the differences between the professors' gender (female and male) and the university schools (six schools described above) as independent variables and the ECOA evaluation criteria scores (the first seven items described above) as dependent variables. A hierarchical regression analysis was applied to determine the prediction effect of the student's recommendation for the best professor considering the ECOA evaluation scores and the professors' gender. The Statistical Program for the Social Sciences (version 25) (IBM Corp., Armonk, NY) was used to perform the analysis.

3. Results

Table 1 shows the distribution of the professors' gender for the different semesters, campuses, and schools. The chi-squared test showed no significant difference in the gender distribution with respect to the semesters, ($\chi^2_{(2)} = 2.298$, $p = 0.317$), but a significant difference in gender distribution on the six different campuses ($\chi^2_{(5)} = 20.980$, $p < 0.001$). This means that there are more male professors than female ones across all the campuses. Noteworthy are the small campuses, like Ciudad Juárez and Tampico, where the number of male professors was about double that of the female professors. The number of male professors is also larger across most of the schools, except for the School of Architecture and Design and the School of Humanities and Education ($\chi^2_{(5)} = 254.574$, $p < 0.001$). In the latter, the number of female professors doubled the number of male professors.

Concerning the multivariate analysis of variance (MANOVA), it shows a significant but slight difference in the ECOA evaluation scores considering gender. The results reported that, in general, female professors obtained slightly higher scores from their students than male professors in all the measured criteria of the ECOA (Tables 2 and 3). Moreover, no

Table 1. Distribution (n and %) of professors' gender across the three semesters, campuses, and schools.

	Gender		χ^2	Total
	Female	Male		
Semesters				
Semester Jan–May 2017	669 (44.0)	853 (56.0)	2.29	1522
Semester Aug–Dec 2017	981 (46.5)	1129 (53.5)		2110
Semester Jan–May 2018	660 (45.5)	791 (54.5)		1451
Campuses				
Tampico	32 (31.7)	69 (68.3)	20.98***	101
Monterrey	1994 (46.6)	2547 (53.4)		4280
Saltillo	88 (39.1)	137 (60.9)		225
Laguna	92 (39.1)	143 (60.9)		235
Chihuahua	57 (48.7)	60 (51.3)		117
Ciudad Juárez	47 (37.6)	78 (62.4)		125
Schools				
Engineering and Science	652 (35.4)	1192 (64.6)	254.57***	1844
Architecture and Design	208 (51.9)	193 (48.1)		401
Medicine and Health Sciences	351 (48.5)	373 (51.5)		724
Humanities and Education	530 (67.2)	259 (32.8)		789
Social Sciences and Government	141 (34.7)	265 (65.3)		406
Business	428 (46.6)	491 (53.4)		919
Total	2310	2773		5083

*** $p < 0.001$.

Table 2. Professors scores in seven criteria included in the SET according to their gender and their university schools.

		Gender		Total
		Female	Male	
Methodology and learning activities	Engineering and Science	8.88 ± 0.88	8.82 ± 0.93	8.85 ± 0.91
	Architecture and Design	8.90 ± 0.80	8.82 ± 0.88	8.85 ± 0.85
	Medicine and Health Sciences	8.90 ± 0.89	8.81 ± 0.92	8.84 ± 0.91
	Humanities and Education	8.87 ± 0.87	8.77 ± 1.02	8.81 ± 0.96
	Social Sciences and Government	8.95 ± 0.88	8.72 ± 1.05	8.81 ± 0.99
	Business	8.88 ± 0.91	8.74 ± 1.05	8.81 ± 0.99
	Total	8.89 ± 0.88	8.79 ± 0.98	8.83 ± 0.94
Concepts in terms of their application in real world scenario	Engineering and Science	9.02 ± 0.79	8.95 ± 0.84	8.98 ± 0.81
	Architecture and Design	9.05 ± 0.70	8.96 ± 0.79	8.99 ± 0.76
	Medicine and Health Sciences	9.05 ± 0.77	8.93 ± 0.83	8.97 ± 0.81
	Humanities and Education	9.00 ± 0.76	8.90 ± 0.91	8.94 ± 0.85
	Social Sciences and Government	9.06 ± 0.74	8.86 ± 0.96	8.94 ± 0.89
	Business	9.03 ± 0.80	8.89 ± 0.97	8.96 ± 0.90
	Total	9.03 ± 0.78	8.92 ± 0.88	8.97 ± 0.84
Adviser	Engineering and Science	9.22 ± 0.73	9.15 ± 0.79	9.18 ± 0.76
	Architecture and Design	9.29 ± 0.61	9.15 ± 0.76	9.20 ± 0.71
	Medicine and Health Sciences	9.21 ± 0.76	9.13 ± 0.84	9.16 ± 0.81
	Humanities and Education	9.24 ± 0.68	9.10 ± 0.90	9.16 ± 0.81
	Social Sciences and Government	9.26 ± 0.66	9.08 ± 0.92	9.15 ± 0.83
	Business	9.23 ± 0.76	9.10 ± 0.89	9.16 ± 0.83
	Total	9.23 ± 0.72	9.12 ± 0.84	9.17 ± 0.79
Evaluation system	Engineering and Science	9.07 ± 0.79	9.00 ± 0.81	9.03 ± 0.80
	Architecture and Design	9.10 ± 0.72	9.01 ± 0.79	9.04 ± 0.77
	Medicine and Health Sciences	9.08 ± 0.84	9.01 ± 0.82	9.04 ± 0.83
	Humanities and Education	9.07 ± 0.76	8.96 ± 0.93	9.01 ± 0.86
	Social Sciences and Government	9.12 ± 0.71	8.93 ± 0.93	9.00 ± 0.86
	Business	9.11 ± 0.77	8.94 ± 0.95	9.02 ± 0.88
	Total	9.09 ± 0.77	8.97 ± 0.87	9.02 ± 0.83
Intellectual challenge	Engineering and Science	9.02 ± 0.78	9.06 ± 0.74	9.04 ± 0.76
	Architecture and Design	9.07 ± 0.66	9.03 ± 0.73	9.04 ± 0.71
	Medicine and Health Sciences	9.05 ± 0.76	9.03 ± 0.80	9.03 ± 0.78
	Humanities and Education	9.02 ± 0.77	8.99 ± 0.88	9.00 ± 0.83
	Social Sciences and Government	9.09 ± 0.66	8.98 ± 0.88	9.02 ± 0.81
	Business	9.01 ± 0.81	9.00 ± 0.85	9.01 ± 0.84
	Total	9.03 ± 0.77	9.02 ± 0.81	9.03 ± 0.79
Learning guide	Engineering and Science	9.05 ± 0.87	9.00 ± 0.91	9.02 ± 0.90
	Architecture and Design	9.08 ± 0.76	9.00 ± 0.81	9.03 ± 0.79
	Medicine and Health Sciences	9.04 ± 0.88	9.00 ± 0.90	9.02 ± 0.89
	Humanities and Education	9.06 ± 0.82	8.95 ± 0.95	9.00 ± 0.90
	Social Sciences and Government	9.10 ± 0.82	8.89 ± 1.09	8.97 ± 1.00
	Business	9.07 ± 0.86	8.94 ± 0.99	9.00 ± 0.93
	Total	9.06 ± 0.85	8.97 ± 0.95	9.01 ± 0.91
Recommend the professor	Engineering and Science	8.76 ± 1.14	8.73 ± 1.19	8.74 ± 1.17
	Architecture and Design	8.77 ± 1.08	8.72 ± 1.13	8.74 ± 1.11
	Medicine and Health Sciences	8.73 ± 1.23	8.72 ± 1.21	8.72 ± 1.22
	Humanities and Education	8.75 ± 1.17	8.69 ± 1.24	8.71 ± 1.21
	Social Sciences and Government	8.82 ± 1.21	8.64 ± 1.19	8.71 ± 1.20
	Business	8.78 ± 1.18	8.65 ± 1.29	8.71 ± 1.24
	Total	8.77 ± 1.16	8.69 ± 1.2	8.73 ± 1.19

significant differences were found in the ECOA scores considering the university schools. Finally, no interaction effect between gender and university schools was found (Table 3).

For the hierarchical regression analysis, we check the multicollinearity through the variance inflation factor values (VIF) where all the values were below 10. This, according to Hair et al. (1995), is considered to be the maximum level of VIF. The criterion p-value for significance was < 0.05. Considering the predictor variable, the recommendation of the professor revealed that the almost all ECOA evaluation

scores contributed significantly to the regression model: $F(6, 5069) = 5530.38, p < 0.001$, accounting for 86% of the variance (Table 4). The model explained that professors with good methodology and learning activities, teaching concepts by using real-world applications, using a valid assessment system, and offering learning guidance were positively recommended by the students as the best professors. On the contrary, if they incorporate difficult intellectual challenges, it correlated negatively with the recommendation of the students as the best professors.

Table 3. Statistical data (F_s , signification level and partial eta-square) on each variable according to the gender of the professors and the university schools.

	Gender (1,13345)			Schools (6,13345)			Gender * Schools (6,13345)		
	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
Methodology and learning activities	35.303	<.001	0.003	1.041	.39	0.000	1.777	.09	0.001
Concepts in terms of their application in real world scenario	44.394	<.001	0.003	0.977	.49	0.000	1.452	.19	0.001
Adviser	49.098	<.001	0.004	0.384	.60	0.000	1.403	.20	0.001
Evaluation system	44.353	<.001	0.003	0.304	.43	0.000	1.92	.07	0.001
Intellectual challenge	6.152	<.01	0.000	0.979	.43	0.000	2.06	.06	0.001
Learning guide	28.889	<.001	0.002	0.397	.88	0.000	2.12	.06	0.001
Recommend the professor	13.175	<.001	0.001	0.253	.95	0.000	1.473	.18	0.001

Moreover, the professors' gender contributed to the regression model: $F(7, 5068) = 4783.43, p < 0.001$, improving it slightly to account for 87% of the variance. The beta values (or coefficients for the equation of regression) showed that being a female was practically invariant, but being a male professor was a factor favoring the student recommendation for the best professor. This finding provides an insight that while students believed that their female professors are equally competent in all measured values, they seem to favor male professors over their female peers slightly when it comes to overall judgment of the professor and addressing the question of "Would you consider the professor as one of the best professors you have had?"

4. Discussion

The descriptive results showed that male professors tended to be generally dominant in every semester and campus. Reporting a similar trend as previous works, this suggests there is a subtle gender gap that pervades the workplace structure favoring men (Cooper et al., 2007; Fan and Sturman, 2019; Muhs et al., 2012; Valian 1999). Additionally, male professors occupy more positions in the schools of engineering and science according to the well-known gender gap in STEM (science, technology, engineering, and mathematics) fields (Wang & Degal, 2017).

Hypothesis 1 predicted that male professors would receive higher ratings; nevertheless, the analysis dismiss it. Considering the standard deviation of the results and the negligible difference between the two sets, it can be concluded that the participants did not favor one gender over another, and they judged their professors fairly. Also, it agrees with Mengel et al. (2019) who found no difference in teaching skills between males and females, at least from the perspective of the students under the

criteria included in the ECOA. Such results suggest the undergraduates participating in this research are seemingly unbiased, and the outcomes presented here do not follow the gender stereotyping as previous studies reported (Valencia, 2019).

In the past, Mexican society has been represented with a high sexism in and out of the media (Bonavitta and de Garay-Hernández, 2011; Vidal-Correa, 2020) and a higher level of hostile sexism by men towards women (Batista Pereira, 2020). However, Nava-Reyes et al. (2018) reported male students did not have high levels of sexism towards women, and showed that educated men (in particular, university students) had not inherited the same level of sexism as past generations. Also, Díaz-Loving et al. (2015), in a study about the norms and beliefs in Mexico, suggested that Mexican female students support equity and self-affirmation as well as sexual openness and emancipation, which has a major impact on these fundamental cultural changes taking place across the nation. While the new generations incorporate innovative ways to combat sexism in Mexico, the educational system, especially higher education institutions will undoubtedly play a great role in such evolutionary pathways. Besides being known for promoting a culture of equity, solidarity, free expression, and student security, Universities have prioritized the values and empowerment of individuals regardless of gender, sexual preferences or religious beliefs (Villanueva, 2019). In this case, We suggest that both, the philosophy and values of the university as well as the high personal and social development of its students are clearly reflected on the gender equity under which ECOA evaluates its professors which clearly differs from the trends reported in past studies (MacNell et al., 2015; Mengel et al., 2019; Wagner et al., 2016).

For the second hypothesis, the analysis showed no differences in the ECOA scores of the university schools, unlike to what previous studies

Table 4. Hierarchical regression analysis on students' recommendations for the best professor considering ECOA evaluation criteria and the professors' gender.

	Beta	t	R ²	Δ R ²
Step 1				
Methodology and learning activities	0.38	25.26***	0.86	0.86
Concepts in terms of their application in the real world	0.07	5.57***		
Adviser	-0.01	-0.07		
Evaluation system	0.10	9.69***		
Intellectual challenge	-0.08	-8.40***		
Learning guide	0.48	30.53***		
Step 2				
Methodology and learning activities	0.38	25.37***	0.87	0.87
Concepts in terms of their application in the real world	0.07	5.66***		
Adviser	0.03	0.21		
Evaluation system	0.10	9.87***		
Intellectual challenge	-0.09	-8.68***		
Learning guide	0.48	30.35***		
Professors' gender	0.02	4.44***		

Notes: Professors' gender: 0 = female and 1 = male.

* $p < 0.05$, *** $p < 0.001$.

have reported (Basow and Montgomery, 2005; Peterson et al., 2019). Further research will be needed for a better understanding of this result and the reasons behind the grading. Criterion number seven in ECOA has the purpose of giving a global evaluation of the teaching quality, unlike the rest of the questions and most SETs which try to rate specific competencies and performance characteristics. The scores in this criterion were the lowest regardless of the gender and the schools. They also shown the highest dispersion from the mean. This result might be a consequence of the students taking into account subjective factors far from the quality of teaching.

While the reported results in this study showed that female faculty obtained better or similar evaluation grades than their male colleagues, when considering the predictive value of professors gender the students slightly leaned towards the male professors as “the best there is.” Therefore, Hypothesis 3 in did predicted the aim result. Many bodies of work have reported gender bias when students are evaluating their professors via a SET (Boring, 2017; Centra and Gaubatz, 2000; Sax et al., 2016; Potvin and Hazari, 2016; Protivinský and Münich, 2018), including some reports in Mexico (Arámburo-Vizcarra and Luna-Serrano, 2013; Arceo-Gomez & Campo-Vazquez, 2019). However, they rely on specific criteria that are based upon teaching performance, and none of them has analyzed the students’ recommendations for the best professor; such an analysis involves a unique, integral way of evaluating a professor by including concrete, objective criteria in addition to some other subjective factors beyond mere performance in the classroom. Therefore, it is difficult to compare the results of this study with others, as the nature of the assessments is different. We postulate the following two hypotheses to this result: One has to do with the behavior of the professors in classes, as generally, students expect their professors to behave according to their stereotypes (Anderson and Smith, 2005; Kombe et al., 2019) and if these expectations are not met, the students may hold the professors accountable in the final evaluations (Dalmia et al., 2005; Kayas et al., 2020; Sprague & Massoni, 2005). Also, the likability of the professors can have a strong effect on the SET outcomes (Feistauer and Richter, 2018) The second possible explanation is that students may have an ingrained cultural mindset that, despite the equal competencies of men and women, they tend to see a male professor as slightly more recommendable than the female counterpart. Several studies show that gender bias is learned in childhood, and it can be passed on as a belief system through family and school or can be transferred generation after generation through media (Atwood, 2001; Newall et al., 2018; Rajan and Morgan, 2018; Singh, 2017; Shor, van de Rijt and Fotouhi, 2019). Furthermore, religious beliefs often play a major role in assuming certain tasks for a female that are less relevant for males and vice versa. Generally, strong religious beliefs are part of the traditional values of Mexican society (Choi et al., 2019; León-Ramírez and Ferrando, 2013). Therefore, the cultural aspects could make it complicated for both female and male adolescents to distance themselves from the common belief system in Mexican society.

There are several implications to this study from which we point out a few: 1) for the first time, we have a complete assessment of a SET that brings together all the north regions of Mexico in a united study; 2) contrary to some of the reports of the literature which indicates a degree of gender bias, our results demonstrate a fair judgment of professors by their students regardless of the gender of the professors. This opens a window of opportunity to further exploration of the reasons behind biased or non-biased students’ evaluation of their professors’ performance; 3) this study is the first to also consider the assessment of the variable “recommendation for the best professor” which has not been previously addressed in the literature. By considering this variable we reported that although students fairly evaluated their professors with no traces of gender bias, when asked this single question, they have shown a slight preference towards male professors, which is a call for future in-depth analysis.

5. Conclusion

The findings of this study suggest insights into the SET applied in a private university which has campuses all over the north region of Mexico, reflecting that students provide fair evaluations to their professors and with almost no gender bias, as long as they assess specific teaching characteristics and skills. While the findings are promising and the level of gender bias is minor in our assessment, there is a crucial need for further investigation of the results at a national level and with consideration of many other factors including demographics, socioeconomic backgrounds of the students, and the type of academic support they receive with gender bias. Nevertheless, there are some limitations. First, the ECOA is an institutional survey designed for purposes aside from this research, and therefore, there are a few shortcomings, specifically, with its design, as mentioned previously by Centra and Gaubatz (2000). Second, factors like age, gender, academic background of the participants as well as any sociodemographic characteristics from the students and professors were not taken into consideration due to privacy issues from part of the institution and the application conditions of the ECOA itself. Third, campus Monterrey represented the majority of participants in this study, which may be a confounding factor to the results of this analysis, therefore future studies should consider a more balanced sample when considering campuses. Finally, this research did not include any analysis of the comments explaining why the students would recommend or not the teacher as the best professor they ever had, which may give us a deeper insight into the impact of gender on the students’ evaluations of professor performance.

Our findings suggest that when undergraduates evaluate their professors on specific criteria concerning teaching performance, they do it without considering gender as a relevant factor, which might derive from a favorable impact of the university’s principles and values. However, when being asked for a single global evaluation as to whether or not the students would recommend the professor as the best they ever had, the students tended to favor male professors slightly. Therefore, this sort of general perception might not be representative of the real quality of teaching.

Declarations

Author contribution statement

A. Arrona-Palacios: Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

K. Okoye: Performed the experiments; Analyzed and interpreted the data.

C. Camacho-Zuñiga: Conceived and designed the experiments; Wrote the paper.

N. Hammout: Analyzed and interpreted the data.

E. Luttmann-Nakamura: Contributed reagents, materials, analysis tools or data.

S. Hosseini: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper.

J. Escamilla: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data.

Funding statement

This work was supported by Writing Lab, TecLabs, Tecnológico de Monterrey, Mexico.

Competing interest statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

- Andersen, K., Miller, E., 1997. Gender and student evaluations of teaching. *PS Polit. Sci. Polit.* 30, 216–219.
- Anderson, K.J., Smith, G., 2005. Students' preconceptions of professors: benefits and barriers according to ethnicity and gender. *Hisp. J. Behav. Sci.* 27, 184–201.
- Annan, S.L., Tratnack, S., Rubenstein, C., Metzler-Sawin, E., Hulton, L., 2013. An integrative review of Student Evaluations of Teaching: implications for the evaluation of nursing faculty. *J. Prof. Nurs.* 29, e10–e24.
- Arámburo-Vizcarra, V., Luna-Serrano, E., 2013. La influencia de las características del profesor y del curso en los puntajes de evaluación docente [The influence of teacher and course characteristics on the evaluation of teaching]. *Rev. Mex. Invest. Educ.* 18, 949–968.
- Arceo-Gomez, E.O., Campos-Vazquez, R.M., 2019. Gender stereotypes: the case of MisProfesores.com in Mexico. *Econ. Educ. Rev.* 72, 55–65.
- Arciniega, M., Anderson, T.C., 2008. Toward a fuller conception of machismo: development of a traditional machismo and caballerismo scale. *J. Counsel. Psychol.* 55, 19–33.
- Atwood, N.C., 2001. Gender bias in families and its clinical implications for women. *Soc. Work* 46, 23–36.
- Ayala-Hernández, P., 2013. Factores que inciden en la evaluación del desempeño docente por los alumnos de nivel superior en la Universidad TecMilenio, campus Ciudad Juárez [Factors for teachers performance assessment for upper level students at University TecMilenio, campus Ciudad Juarez]. *Noesis* 43, 190–224.
- Balam, E.M., Shannon, D.M., 2010. Student ratings of college teaching: a comparison of faculty and their students. *Assess Eval. High Educ.* 35, 209–221.
- Barth, M., 2008. Deciphering student evaluations of teaching: a factor analysis approach. *J. Educ. Bus.* 84, 40–46.
- Batista Pereira, F., 2020. Do female politicians face stronger backlash for corruption allegations? Evidence from survey-experiments in Brazil and Mexico. *Polit. Behav.*
- Basow, S.A., Montgomery, S., 2005. Student ratings and professor self-ratings of college teaching: effects of gender and divisional affiliation. *J. Person. Eval. Educ.* 18, 91–106.
- Beran, T.N., Rokosh, J.L., 2009. The consequential validity of student ratings: what do instructors really think? *Alberta J. Educ. Res.* 55, 497–511.
- Bianchini, S., Lissoni, F., Pezzoni, M., 2013. Instructor characteristics and students' evaluation of teaching effectiveness: evidence from an Italian engineering school. *Eur. J. Eng. Educ.* 38, 38–57.
- Blackmore, J., 2009. Academic pedagogies, quality logics, and performative universities: evaluating teaching and what students want. *Stud. High Educ.* 34, 857–872.
- Bonavitta, P., de Garay-Hernández, J., 2011. De estereotipos, violencia y sexismo: la construcción de las mujeres en los medios mexicanos y argentinos [Of stereotypes, violence and sexism: the construction of women in the Mexican and Argentine media]. *Anagramas Rumbos y Sentidos de la Comunicación* 9, 15–29.
- Boring, A., 2017. Gender biases in student evaluations of teaching. *J. Publ. Econ.* 145, 27–41.
- Centra, J.A., Gaubatz, N.B., 2000. Is there gender bias in students' evaluations of teaching? *J. High Educ.* 71, 17–33.
- Cooper, J., Eddy, P., Hart, J., Lester, J., Lukas, S., Eudey, B., Madden, M., 2007. Improving gender equity in postsecondary education. In: Klein, S.S., Richardson, B., Grayson, D.A., Fox, L.H., Kramarae, C., Pollard, D.S., Dwyer, C.A. (Eds.), *Handbook of Achieving Gender Equity through Education*, second ed. Lawrence Erlbaum, Mahwah, NJ, pp. 631–653.
- Choi, N.Y., Kim, H.Y., Gruber, E., 2019. Mexican American women college student's willingness to seek counseling: the role of religious cultural values, etiology beliefs, and stigma. *J. Counsel. Psychol.* 66, 577–587.
- Dalmia, S., Giedeman, D.C., Klein, H.A., Levenburg, N.M., 2005. Women in academia: an analysis of their expectations, performance and pay. *Forum Public Policy* 1, 160–177.
- Davis, B.G., 2009. *Tools for Teaching*, second ed. Jossey-Bass, San Francisco, CA.
- Díaz-Loving, R., Saldívar, A., Armenta-Hurtarte, C., Reyes, N.E., López, F., Moreno, M., Romero, A., Hernández, J.E., Domínguez, M., Cruz, C., Correa, F.E., 2015. Creencias y normas en México: una actualización del estudio de las premisas psico-socio-culturales [Beliefs and norms in Mexico: an update of the study of psycho-socio-cultural premises]. *Psykhe* 24, 1–25.
- Dirección de Servicios Académicos, 2018. *Manual de Encuesta de Opinión de Alumnos (ECO)*. Tecnológico de Monterrey, Monterrey.
- Fan, X., Sturman, M., 2019. Has higher education entered the problem? Examining the gender wage gap of recent college graduates entering the workplace. *Compensat. Benefit Rev.* 51, 5–12.
- Feistauer, D., Richter, T., 2018. Validity of students' evaluations of teaching: biasing effects of likability and prior subject interest. *Stud. Educ. Eval.* 59, 168–178.
- Felton, J., Mitchell, J., Stinson, M., 2004. Web-based student evaluations of professors: the relations between perceived quality, easiness, and sexiness. *Assess Eval. High Educ.* 29, 91–108.
- Fortson, S.B., Brown, W.E., 1998. Best and worst university instructors: the opinions of graduate students. *Coll. Student J.* 32, 572–576.
- Freng, S., Webber, D., 2009. Turning up the heat on online teaching Evaluations: does 'hotness' matter? *Teach. Psychol.* 36, 189–193.
- Galván-Salinas, J.O., Fariás-Martínez, G.M., 2018. Características personales y práctica docente de profesores universitarios y su relación con la evaluación del desempeño [Personal characteristics and teaching practice of university professors and their relationship with the performance evaluation]. *Revista Iberoamericana de Evaluación Educativa* 11, 9–33.
- Hair Jr., J.F., Anderson, R.E., Tatham, R.L., Black, W.C., 1995. *Multivariate Data Analysis*, third ed. Macmillan, New York.
- Hamermesh, D.S., Parker, A.M., 2005. Beauty in the Classroom: instructors' pulchritude and putative pedagogical productivity. *Econ. Educ. Rev.* 24, 369–376.
- Heck, J.L., Todd, J., Finn, D., 2002. Is student performance enhanced by perceived teaching quality? *J. Fin. Educ.* 28, 54–62.
- Heckert, T.M., Latier, A., Ringwald, A., Silvey, B., 2006. Relation of course, instructor, and student characteristics to dimensions of student ratings of teaching effectiveness. *Coll. Student J.* 40, 195–203.
- Hills, S.B., Naegle, N., Bartkus, K.R., 2009. How important are items on a student evaluation? A study of item salience. *J. Educ. Bus.* 84, 297–303.
- Kayas, O.G., Assimakopoulos, C., Hines, T., 2020. Student evaluations of teaching: emerging surveillance and resistance. *Stud. High Educ.*
- Koch, A.J., D'Mello, S.D., Sackett, P.R., 2015. A meta-analysis of gender stereotypes and bias in experimental simulations of employment decision making. *J. Appl. Psychol.* 100, 128–161.
- Kombe, D., Che, S., Bridges, W., 2019. Students gendered perceptions of mathematics in middle grades single-sex and coeducational classrooms. *Sch. Sci. Math.* 119, 417–427.
- Kristof De Witte, K., Rogge, N., 2011. Accounting for exogenous influences in performance evaluations of teachers. *Econ. Educ. Rev.* 30, 641–653.
- León-Ramírez, B., Ferrando, P.J., 2013. Assessing sexism in a sample of Mexican students: a validity analysis based on the Ambivalent Sexism Inventory. *Anuario de Psicología* 43, 335–347.
- MacNeill, L., Driscoll, A., Hunt, A.N., 2015. What's in a name: exposing gender bias in student ratings of teaching. *Innovat. High. Educ.* 4, 291–303.
- Marsh, H.W., Bailey, M., 1993. Multidimensional students' evaluations of teaching effectiveness: a profile Analysis. *J. High Educ.* 64, 1–18.
- Martínez-González, A., Martínez-Stack, J., Buquet-Corneto, A., Díaz-Bravo, P., Sánchez-Mendiola, M., 2012. Satisfacción de los estudiantes de medicina con el desempeño de sus docentes: género y situaciones de enseñanza [Medical students satisfaction with their teachers educational performance: gender and teaching situations]. *Investigación en Educación Médica* 1, 64–74.
- Mazzarol, T., Soutar, G.N., Seng, M.S.Y., 2003. The third wave: future trends in international education. *Int. J. Educ. Manag.* 17, 90–99.
- McPherson, M.A., 2006. Determinants of how students evaluate teachers. *J. Econ. Educ.* 37, 3–20.
- Mena, P., Rojas, O., 2010. Padres solteros de la ciudad de México: un estudio de género [Single fathers in Mexico City: a gender study]. *Papeles de Población* 16, 41–74.
- Mengel, F., Saueremann, J., Zöllitz, U., 2019. Gender bias in teaching evaluations. *J. Eur. Econ. Assoc.* 17, 535–566.
- Montemayor-Gallegos, J.E., 2002. Confiabilidad y validez de la encuesta de opinión realizada a los alumnos para evaluar y retroalimentar el desempeño de los profesores del ITESM [Reliability and validity of the opinion survey carried out on students to evaluate and provide feedback on teacher performance from ITESM] (Master thesis). Tecnológico de Monterrey, Mexico.
- Muhs, G.G., Niemann, Y.F., González, C.G., Harris, A.P., 2012. *Presumed Incompetent: The Intersections of Race and Class for Women in Academia*. Utah State University Press, Logan.
- Nava-Reyes, M.A., Rojas-Solís, J.L., Greathouse, L.M., Morales, L.A., 2018. Gender roles, sexism, and myths of romantic love in Mexican adolescents. *Interam. J. Psychol.* 52, 102–111.
- Newall, C., Gonsalkorale Walker, E., Forbes, G.A., Highfield, K., Sweller, N., 2018. Science education: adult biases because of the child's gender and gender stereotypicality. *Contemp. Educ. Psychol.* 55, 30–41.
- Ogier, J., 2005. Evaluating the effect of a lecturer's language background on a student rating of teaching form. *Assess Eval. High Educ.* 30, 477–488.
- Peterson, D.A.M., Biederman, L.A., Andersen, D., Ditonto, T.M., Roe, K., 2019. Mitigating gender bias in students evaluation of teaching. *PLoS One* 14, e0216241.
- Potvin, G., Hazari, Z., 2016. Student evaluations of physics teachers: on the stability and persistence of gender bias. *Phys. Rev. Phys. Educ. Res.* 12, 020107.
- Protivínský, T., Münich, D., 2018. Gender bias in teachers grading: what is in the grade? *Stud. Educ. Eval.* 59, 141–149.
- Rajan, S., Morgan, S.P., 2018. Selective versus generalized gender bias in childhood health and nutrition: evidence from India. *Popul. Dev. Rev.* 44, 231–255.
- Sax, L.J., Lehman, K.J., Jacobs, J.A., Kanny, M.A., Lim, G., Monje-Paulson, L., Zimmerman, H.B., 2016. Anatomy of an enduring gender gap: the evolution of women's participation in computer science. *J. High Educ.* 88, 258–293.
- Shor, E., van de Rijt, A., Fotouhi, B., 2019. A large-scale test of gender bias in the media. *Sociol. Sci.* 6, 526–550.
- Singh, A., 2017. Gender based within-household inequality in childhood immunization in India: changes over time and across regions. *PLoS One* 12, e0173544.
- Sitzman, K., 2010. Student-preferred caring behaviors for online nursing education. *Nurs. Educ. Perspect.* 31, 171–178.
- Spooren, P., Christiaens, W., 2017. I liked your course because I believe in (the power of) student evaluations of teaching (SET). Students' perceptions of a teaching evaluation process and their relationships with SET scores. *Stud. Educ. Eval.* 54, 43–49.
- Sprague, J., Massoni, K., 2005. Student evaluations and gendered expectations: what we can't count can hurt us. *Sex Roles* 53, 779–793.
- Tang, F., Chou, S., Chiang, H., 2005. Students' perceptions of effective and ineffective clinical instructors. *J. Nurs. Educ.* 44, 187–192.
- Tatro, C.N., 1995. Gender effects on student evaluations of faculty. *J. Res. Dev. Educ.* 28, 169–173.

- Uttl, B., White, C.A., Wong Gonzalez, D., 2017. Meta-analysis of faculty's teaching effectiveness: student evaluation of teaching ratings and student learning are not related. *Stud. Educ. Eval.* 54, 22–42.
- Valencia, E., 2019. Acquiescence, instructors gender bias and validity of student evaluation of teaching. *Assess Eval. High Educ.*
- Valian, V., 1999. *Why So Slow? the Advancement of Women*. MIT Press, Cambridge, MA.
- Vidal-Correa, F., 2020. Media coverage of campaigns: a multilevel study of Mexican women running for office. *Commun. Soc.* 33, 167–186.
- Villanueva, A., 2019, January 22. Por la igualdad de género; se une Tec a la campaña mundial HeforShe. *Conecta*. Retrieved from: <https://tec.mx/es/noticias/nacional/institucion/por-la-igualdad-de-genero-se-une-tec-la-campana-mundial-heforshe>.
- Wagner, N., Rieger, M., Voorvelt, K., 2016. Gender, ethnicity, and teaching evaluations: evidence from mixed teaching teams. *Econ. Educ. Rev.* 54, 79–94.
- Wang, M.T., Degal, J.L., 2017. Gender gap in Science, Technology, Engineering, and Mathematics (STEM): current knowledge, implications for practice, policy, and future directions. *Educ. Psychol. Rev.* 29, 119–140.
- Wiley, C., 2019. Standardized module evaluation surveys in UK higher education: establishing students' perspective. *Stud. Educ. Eval.* 61, 55–65.