

FACULTY AND STUDENT GENDER EFFECTS ON TEACHER EVALUATIONS

*Therese A. Louie, College of Business, San José State University, One Washington Square, San José, CA
95192, 408-924-3454, tlprofsj@cob.sjsu.edu*

ABSTRACT

Teacher evaluations can affect both faculty promotion decisions and students' course selections. This exploratory research examines if both faculty and student gender influence teacher ratings. A dataset of official undergraduate course evaluations at a large public university was augmented with information about each course's instructor gender. The findings contrast with past research—conducted when females were less prevalent on campuses—that showed a preference for male faculty. Examination of significant main and interaction effects reveals that ratings were highest when female students evaluated female faculty, and lowest when male students evaluated male faculty. Discussion focuses on future research.

INTRODUCTION

To achieve promotion and tenure professors must market successfully their contributions to learning. That information is often conveyed via teacher evaluations. Many factors can potentially influence how students rate their professors. This study focuses on the influence of both instructor and student gender, and examines if there are male-to-male and female-to-female matching effects.

Studies exploring “gender matching” using real world data are scarce largely due to the challenge of obtaining data with demographic information about both the faculty and students. Over two decades ago, Sidanius and Crane [6] analyzed teacher evaluations using both student and faculty variables. They found effects for the latter, such that male faculty received higher evaluations than female faculty on “global teacher effectiveness” and “academic competence.” Basow and Silberg [2] also found a preference for male faculty when surveying students, as well as some evidence of student-gender effects. Specifically, male students gave male professors more favorable ratings than their female counterparts on all measures. Female students rated professors more similarly, although they did rate female faculty lower on some behaviors. When the above-mentioned studies were conducted, there were noticeably more male than female professors on college campuses [3], and more male than female students [5].

This brings up the question of how past findings were affected by the proportion of male and female professors in the environments in which the studies were conducted. Since that time, diversity on campuses has increased substantially. Between 1976 and 1995, the percentage of women academics increased by 114% [3]. In addition, since 1991 the proportion of young women enrolled in college has exceeded that of young men, with the former comprising 54% of young adult college enrollment in 2005 [5].

Changing campus demographics provide an opportunity for student perceptions to shift such that the inclination towards male faculty expressed in past work may not be as prevalent. For example, when researching teacher evaluations in the 1990s Basow [1] found that female professors received their

highest ratings from female students, and their lowest ratings from male students. This suggests the possibility that students have a higher appreciation of instructors who share their gender. The one-directional preference for male faculty may have evolved with the changing norms prompted by increased exposure to both genders in college settings.

It is proposed that male students will more favorably evaluate male than female instructors. In addition, female students will more favorably evaluate female than male instructors. In sum, this research reinvestigates if male and female students have different reactions to faculty gender, and if there are matching effects.

METHOD

Teacher Evaluations Dataset

Students enrolled in undergraduate classes at a large public university anonymously completed official end-of-the-term evaluations. The dataset contains information from one semester, which includes 55,718 completed evaluation forms from all departments across campus that provided undergraduate lecture courses.

Faculty and Student Gender

In the dataset, which was devoid of faculty and student names, each course had an identification number that in a second file was listed with instructors' names. Research assistants used lists from human resources to ascertain faculty gender, which was added to the original teacher evaluation file. Note that demographic information was not available for a small percentage of courses taught by instructors whose positions were not processed through the typical faculty hiring procedure (e.g., graduate students, last-minute hires, and coaches). The 1.8% of the data that fell into this group was excluded from the dataset.

The last part of the teacher evaluation form asked students to note if they were male or female. For 1.56% of the returned materials the question was unanswered. In all, 54,729 entries contained the faculty and student gender information necessary for this investigation.

Dependent Variable

The dependent variable is a global measure used most frequently in the university's faculty promotion decisions. Students responded to the statement, "Overall, this instructor's teaching was..." with ratings ranging from one (1 = very ineffective) to five (5 = very effective). In addition, past research indicates that students' expected course grades influence heavily teacher evaluations [4]. Therefore, responses to the teacher evaluation form question, "What is your current estimate of your expected overall grade in this course?" were used as a covariate in the 2 x 2 gender of faculty by gender of student data analysis.

RESULTS

As anticipated, the results are contrary to those found in past work. Instead of the prior preference for male faculty [6] ANOVA reveals a significant main effect of faculty gender wherein the mean rating is higher for female than for male instructors ($M = 4.27$, and $M = 4.20$, respectively, $F(1, 53,555) = 79.08$, $p < .0001$.) In addition, a significant main effect for student gender reveals that females provided higher ratings than males ($M = 4.26$ and $M = 4.22$, respectively, $F(1, 53,555) = 24.863$, $p < .0001$).

These main effects are qualified by a significant interaction effect $F(1, 53,555) = 3.843, p = .05$. Post hoc tests ($p < .05$) reveal mixed support for the gender matching hypotheses. It was anticipated that male students would provide higher mean ratings for male than for female instructors; instead, the effect was significant in the opposite direction ($M = 4.19$ and $M = 4.24$, respectively). In contrast, although the male undergraduates did not show positive gender matching effects, the females did. Female students provided significantly higher mean ratings for female than for male instructors ($M = 4.30$ and $M = 4.21$, respectively).

DISCUSSION

Although it was anticipated that students would prefer instructors who share their gender, the results of this study suggest that gender matching was only favorable for females. The highest mean teacher evaluation was obtained when female students evaluated female professors. Surprisingly, the lowest was obtained when males evaluated male professors. The results are in contrast to past research that showed an overall preference for male instructors [2] [6]. Although student preferences for male and female instructors have changed with increased faculty and student gender diversity, the shifts were not quite as anticipated.

It is worthwhile to consider implications of these findings, to review limitations of this work, and to suggest research extensions. In terms of applying the findings, any existing gender preferences may have broad implications for faculty members' merit and promotion decisions. Yet, information about student and instructor gender is often not formally analyzed or even measured. Such data might be diagnostic for all faculty members. It might make departments aware of systematic preferences and biases for female or male instructors. On an individual level, it could reveal to instructors potential difficulties in reaching student gender groups.

The inconsistency between this study's findings and past work suggests that changes in cultural norms have evolved with shifts in the proportion of female faculty, as well as with more general shifts in gender diversity in professional settings. Although the results of this study suggest that female-to-female preferences exist, more research is needed to establish conditions under which they, as well as the traditional male-to-male effects, occur.

The dataset provided a unique opportunity to examine gender matching outside of laboratory settings; the use of actual teacher ratings instead of surveys enhances the study's external validity. However, there are many limitations to this research that must be addressed, as well as future research ideas that are worth pursuing. It is important when interpreting data to consider construct validity. Was it faculty gender that influenced the results, or perhaps another variable related to gender? One possibility is that students relate to newer instructors who, in turn, have much higher pressures to perform favorably in the classroom than their tenured and fully promoted counterparts. If there is a higher percentage of untenured female than male professors, or a higher percentage of tenured male than female professors, then the results could be influenced by professional status as well as gender. Future research can account for this factor by including information about faculty rank in the dataset.

Following typical campus procedures the teacher evaluation forms were completed voluntarily. The students who declined to participate may have withheld information relevant to gender-matching effects. On the other hand, even without their responses the dataset is large. So an opposite critique of the study is that the significant effects may be attributed in part to the large sample, and hence the high statistical

power. Conducting this study at different colleges would test how generalizable the findings are, as well as potentially provide insights into factors that foster both female and male success.

Having reviewed limitations of the study, it is worthwhile to reiterate that the dataset contained official teacher evaluations used for formal faculty promotion decisions. Even if the information gathering process was difficult to control, the data are consequential. Hence, the female instructor preference in this study warrants further attention, as does the higher ratings by female than male students. Future research can explore under what conditions these effects are likely to occur. In the meantime, it is hoped that the real world findings from this work provide new insights into student/instructor, as well as more general professional, gender effects.

REFERENCES

- [1] Basow, S.A. Student evaluations of college professors: When gender matters. *Journal of Educational Psychology*, 1995, 87(4), 656-665.
- [2] Basow, S.A. & Silberg, N.T. Student evaluations of college professors: Are female and male professors rated differently? *Journal of Educational Psychology*, 1987, 79(3), 308-314.
- [3] Clery, S. Faculty in academe. *NEA Higher Education Research Center Update*, 1998, 4(4), 1-8.
- [4] Ginexi, E.M. General Psychology course evaluations: Differential survey response by expected grade. *Teaching of Psychology*, 2003, 30(3), 248-251.
- [5] Mather, M. & Adams, D. The crossover in female-male college enrollment rates. *Popular Reference Bureau*, 2007, February, accessed January 13, 2009 at <http://www.prb.org/Articles/2007/CrossoverinFemaleMaleCollegeEnrollmentRates.aspx>.
- [6] Sidanius, J. & Crane, M. Job evaluation and gender: The case of university faculty. *Journal of Applied Social Psychology*, 1989, 19(2), 174-197.