

Student Evaluations of Faculty Performance: The Role of Gender Expectations in Differential Evaluations

Marilyn S. Chamberlin, & JoAnn S. Hickey, Western Carolina University

Past research has shown how gender contributes to different styles of communication, management, parenting and teaching. Gender roles influence the expectations people have for job performance within various settings. These expectations in turn influence performance evaluations. The purpose of the current study is to identify gendered expectations, holding occupation constant, based on evaluations of instructors' performance in the classroom. Male and female instructors from the same disciplines were evaluated using the same questionnaire. Questions were selected based on gendered expectations. The results of this study indicate that male and female instructors are evaluated differently on aspects of presentation and classroom structure. These findings have implications for promotion and advancement of males and females in traditional gender fields and non traditional gender fields. In addition the results indicate the need to address such differences both within the classroom and within academic institutions.

Evaluations of instructors are influenced by many factors. When these evaluations are used to determine promotion, salary and tenure, identification of these factors becomes critical. If evaluation techniques and tools have inherent biases then the decisions based on these tools and techniques will carry similar biases. Ascribed statuses such as race, age, and sex are potential sources for differential evaluations. The focus of this study is to isolate the effect of sex on the course evaluations of faculty by students.

SEX OF THE INSTRUCTOR

Past research on the influence of the sex of an instructor on evaluations by students has generated conflicting results. Some studies have found no differences between male and female faculty on overall evaluations (Feldman 1993; Dukes and Victoria 1989). Some of the reasons for the lack of general differences can be attributed to methodological factors. Dukes and Victoria (1989) used scenarios rather than actual faculty teaching in front of the students. The students filled in any details that may have been omitted in these scenarios. Even using fictional situations, specific teaching characteristics were found to be influential. Enthusiasm, rapport, knowledge of the material, and organization were all found to influence effectiveness of teaching. The analysis failed to examine the gender differences in these teaching characteristics. Other research has found that these teaching characteristics are related to sex. Female faculty are expected to be warm, friendly (Martin 1984), and encouraging of questions (Feldman 1993). These characteristics form teacher enthusiasm. Characteristics that are related to rapport include being supportive, deferential, and open to others' opinions. These attributes have been identified as expectations or

behaviors of female faculty (Feldman 1993; Martin 1984). Male faculty have been rated higher on clarity and knowledge of subject matter (Feldman 1993) and identified by knowledge and organization in Dukes and Victoria (1984). If teaching characteristics reflect gender stereotypes then the differences of sex may be masked by these characteristics. For example a woman is supposed to have better rapport, so the influence of rapport may be a sex difference not a style difference.

One study that indicates sex of the instructor may be a mediating factor in style and presentation influences examines course content and sex differences of the instructor when there is student resistance to the topic. Moore (1997) examined how students react to a male instructor presenting feminist views and a female instructor presenting the same views. She found that students described the male as "more scientific,

TABLE I
Attributes that form the sex-stereotype

MALE ATTRIBUTES	AUTHOR	FEMALE ATTRIBUTES
Aggressive	Feldman et al. (1989)	Tactful
Independent	Sabin (1978)	Cauter
		Quieter
Competence	Breverman et al. (1972)	Warmth
Rationality		Expressiveness
Independent	Deaux & Lewis (1984)	Emotional
Active		Able to devote self completely to others
Can make decisions easily		Cauter
Competitive		Kind
Never gives up easily		Aware of the feelings of others
Self-confident		Understanding of others
Stands up well under pressure		Wary in relation with others
Feels superior		Helpful to others

more objective, more qualified, and less political." The two were presenting the same view in a course on family issues, that included information on gender. Moore's study is primarily anecdotal but provides

the basis for a broad analysis of differences in evaluations due to sex of the instructor for similar course content.

Overriding in the evaluations of faculty are sex-stereotypes. Sex stereotypes influence both the expectations and evaluations of the faculty members. Identifying attributions within these stereotypes becomes an important step in understanding the differential evaluations of male and female faculty.

SEX STEREOTYPES

Despite specific attributions, gender or sex stereotypes have been described as consisting of multiple components (Broverman et al. 1972). These components include traits, role behaviors, gender labels, physical appearance, and occupation. When one component is demonstrated, people infer the existence of other components. Deaux and Lewis (1984) studied these inferences by giving subjects one element and having them infer others. Specifically they provided some subjects with role behaviors and had them infer about gender labels and traits. Other subjects were given gender labels and traits and asked to infer about roles and occupations. Specific traits on role behaviors were found to be more important than gender labels, but physical appearance (tendency toward appearing masculine or feminine) was the most important predictor of inferences about performance (Deaux Lewis 1984). Relating to the importance of physical appearance is the study by Baker and Copp (1997). They found that when a professor entered the later stages of her pregnancy and began to have the physical signs of pregnancy her course evaluations changed. Students were more critical of her presentation of "feminist" views. The results of Baker and Copp's study indicate that physical traits and seeing an instructor rather than just reading a scenario describing the sex makes a difference in how students evaluate that person. Therefore physical cues are being used to determine the way in which material is perceived.

Similar to Dukes and Victoria's (1989) study, the subjects in Deaux and Lewis (1984) were only provided with one characteristic or characteristics separate from the gender identity, rather than simultaneously evaluating multiple characteristics. In addition subjects in both studies were given fictitious people to evaluate instead of people they knew. The question remains, when gender is known, occupation is the same, and task characteristics or roles are assumed to be the same, will subjects use sex stereotypes to describe the way the roles are performed by people with whom they have direct contact?

While most studies on student evaluations have focused on teaching characteristics, similar studies assessing the influence of sex stereotypes in the area of leadership and management provide evidence of evaluation bias based on sex rather than job specific characteristics. For example, researchers have found that stereotypical views influenced the evaluations of the leader of the group (McGlashan et al. 1995; Hartman et al. 1992). Furthermore, in terms of performing a task that has been established as either masculine or feminine, raters have imposed gender stereotypes in

their evaluations of performance. In other words if the task was masculine and the performer a female the evaluations were less positive (Hartman et al. 1992). Studies on the differential evaluation of faculty and group leaders have applied many stereotypical adjectives. The common terms have been summarized in Table 1.

While past research has focused on either the sex stereotype differences or the difference between males and females on evaluations, few have combined the two. In this study we have taken characteristics or aspects of teaching and asked students to identify the instructor's means of accomplishing the tasks. The means were either stereotypically masculine or feminine. Using past indicators from Table 1 and some from Bem's Sex Role Inventory (Bem 1981), we constructed these stereotypical dichotomies. Given the dichotomous sex stereotyped behaviors regarding faculty teaching style, will evaluators choose the characteristics that are consistent with the sex of the performers? For example grading is a teaching characteristic or aspect, applying impartiality or sensitivity to students are the sex attributes.

HYPOTHESIS I:

Male and female faculty will be differentially evaluated on teaching attributes. Male faculty will be viewed as consistent with masculine attributes and female faculty will be viewed as consistent with feminine attributes.

Sex of the Student

Another aspect of evaluations of faculty teaching style, is the interactions of the sex of the student and the sex of the instructor. As Feldman (1993) noted this is one of the most tested demographic variables of students. His meta-analysis indicates that the relationship between students' sex and their differential evaluations of faculty based on the instructor's sex seems to be significant. The majority of the studies Feldman reviewed found that male students rate male professors more positively and female students rate female professors more positively. As with the studies that tested the relationship between sex of the instructor and course evaluations, many of these studies used global evaluations and did not specify specific attributes of teaching. One study that has some information on the effects of students' sex on evaluations was Kay (1979). In her study, Kay found that males were more likely to give female faculty they defined as "hard graders" harsher evaluations. Although this is only one indicator, female instructors that perform in non-stereotypical ways seem to be more harshly evaluated for those deviations.

Students may be relying on either traditional views of gender performances or their own attributional processes when making the evaluations of others. Hartman et al. (1992) found that female evaluators of performance in the work place used more gender stereotypes than men. Women also tend to focus on external factors for their success and performance (Roberts & Hoeksema 1989) and have lower expectations for their own performance (Basow & Medcalf 1988). Based on these

Table 2
t-Test of Teaching Attributes for Male and Female Faculty

Teaching Attribute	# of Cases	Mean	Std. Dev.	Levdiff
Most Likely to Present Own Beliefs Clearly				
Male Faculty	33	4.37	1.53	3.31**
Female Faculty	114	3.69	1.46	
When Students Demand: Sense of Urgency				
Male Faculty	34	4.32	1.84	3.38**
Female Faculty	14	2.27	1.58	
Teaching Style More: Involved- Inflexible				
Male Faculty	34	3.39	1.24	5.27**
Female Faculty	113	2.25	1.27	
When Presenting Material: Assertive- Unassertive				
Male Faculty	33	2.55	1.44	4.53**
Female Faculty	113	1.71	1.06	
Emphasizing Relationships Among Topics: Own Beliefs- Adaptable				
Male Faculty	33	4.84	1.56	1.49
Female Faculty	112	4.51	1.54	
Feedback Provided: Sympathetic- Completely Objective				
Male Faculty	33	3.42	1.64	2.66*
Female Faculty	111	2.83	1.46	

* p < .05

** p < .001

studies and those on classroom performance, male and female students are expected to differ on their evaluations.

HYPOTHESIS II:

Male and female students will differ in their evaluation of faculty. Female students are expected to employ more gender stereotypes in their evaluations, but males are expected to rate female faculty more harshly.

Table 2 (cont.)
 T-Test of Teaching Attributes for Male and Female Faculty

Teaching Attributes	# of Courses	Mean	Std. Dev.	Level
Deals with Students: Without Regard-Nonstrictly				
Male Faculty	87	4.57	.56	2.94*
Female Faculty	113	5.26	.52	
Grades Students: Fairly/Impartially- With Compassion				
Male Faculty	87	3.38	1.57	1.60
Female Faculty	112	2.71	1.61	
When Student Has Problems: Helpful- Unconcerned				
Male Faculty	82	2.37	1.36	3.77**
Female Faculty	113	1.87	1.21	
When Student Has Diff. Interest/Abilities: Sensitive-Thawware				
Male Faculty	87	3.69	1.57	5.32**
Female Faculty	112	2.48	1.41	
Prevalent Classroom Atmosphere: Open- Dominated				
Male Faculty	84	2.42	1.39	
Female Faculty	113	1.87	1.33	1.51
Student Question/Opinions: Challenged- Openly Discussed				
Male Faculty	84	5.10	1.78	
Female Faculty	113	5.46	1.59	

* $p < .01$

** $p < .001$

Methods

The subjects in the study consisted of undergraduate students enrolled in 7 sections of general education classes in sociology and anthropology. The criteria for the selection of courses were that multiple sections of the course had to be offered during a given semester with both male and female instructors. The courses selected included 4 sections of introductory sociology (2 taught by a female Associate Professor and 2 taught by a male Associate Professor) and 3 sections of introductory anthropology (1 taught by a male Assistant Professor, 1 taught by a female Instructor, and 1 taught by a female Full Professor). Students in these 7 courses were asked to voluntarily fill out a questionnaire administered during the final month of the semester. A total of 198 students responded (84 males, 114 females).

Table 3
 Calculation of Sex of Faculty Member by Teaching Attributes
 Controlling for Sex of Student

Teaching Attribute	Chi-Square	Male Students	Female Students
Most Liked to Present Own Theories/Other's Beliefs	17.75*	17.75*	2.10
When Students are Bored: Stops If Easily, Is Unaware	18.26**	18.26**	53.6***
Teaching Style Makes Me Feel Involved, Inferior	5.18	5.18	39.94***
When Presenting Material: Is Anxious or Uncertain	4.49	4.49	20.25***
Emphasizing Relationships Among Topics: Derives Own Beliefs, Adaptable	5.6	5.6	2.16
Feedback Provided: Sympathetic Understanding, Completely Objective	3.32	3.32	18.89*
Deals with Students: Without Regard to Individual, Sensitively	1.04	1.04	25.30**
Grades Students: Fairly/Impartially, With Compassion	2.34	2.34	6.57
When Student Has Problems: Is Helpful, Unconcerned	14.25**	14.25**	13.40**
When Student has Diff. Interests/Abilities: Sensitizes, Informs	7.63	7.63	25.41**
Presented Classroom Atmosphere: Open to Discussion, Dominated by Instructor	2.43	2.43	9.73*
Student Questions/Opinions: Challenged Regularly, Openly Discussed	2.43	2.43	4.86

* $p < .05$ ** $p < .01$ *** $p < .001$

The data analyzed consisted of questions regarding student perception of the instructor's highest degree received, rank, and tenure status in addition to 12 questions regarding attributes of the instructor's teaching style. Each of the latter questions was designed with a 7-point thermometer response scale with bipolar anchor phrases placed over the 1 and 7 values. Students were instructed to circle the number of the response that most typically represented the instructor's teaching style. The teaching attribute questions were selected based on the characteristics deemed relevant in previous research assessing the effects of student/faculty sex composition of courses and students' evaluation of instruction.

Results

Two sets of preliminary analyses were run to ascertain whether the demographic characteristics of students differed in the classes for male

compared with female instructors and whether students perceived male and female instructor's expertise differently. Regarding the background characteristics of the students in classes taught by male versus female instructors, no significant differences were found for racial composition, age, year in school, overall grade point average, or the grade expected in that class. Male versus female faculty credentials and expertise were also nonsignificant on items assessing student perceptions of the highest degree received by the faculty member, the rank of the faculty member, and whether the faculty member was tenured. Thus, the researchers concluded that any differences found in teaching attributes were not related to differences in the student composition of classes or to perceived differences in the "legitimacy" of the instructor.

T-tests and 3-way crosstabulations (instructor's sex by evaluation of teaching attributes controlling for sex of student) were computed to evaluate the following general hypotheses:

1. Male and female faculty will be differentially evaluated on teaching attributes.
2. Male and female students will differ in their evaluation of faculty. Essentially, these hypotheses test the assumption that the teaching attributes demonstrated by male and female faculty are perceived differently and that male and female students evaluate faculty attributes differently.

To test the first general hypothesis regarding perceived differences in the evaluation of male and female faculty, t-tests were computed for each of the 12 questions on teaching style. Three of the twelve items were not statistically significant (See Table 2). These three items asked students to rate the instructor's tendency to defend his/her own beliefs rather than be adaptable, to evaluate how the instructors grade students (fairly and impartially or with compassion), and whether the instructor responds to differing opinions raised by students by challenging the opinions or openly discussing them. Although the differences for these three questions were not statistically significant, the mean values suggest that male faculty were perceived by students to be more adaptable and more compassionate in their grading than the female faculty, but less likely to openly discuss differing opinions or questions raised by students. Each of the remaining 9 items was significant at $p < .01$ providing support for Hypothesis I: Male and Female faculty will be differentially evaluated on teaching attributes.

In order to test the second hypothesis regarding the tendency for male and female students to evaluate faculty differently, crosstabulations were computed for sex of instructor by each of the twelve items, controlling for the sex of the student evaluating the faculty member (Table 3). The three items with nonsignificant t-tests (adaptability, grading, and responding to differing opinions) also resulted in nonsignificant chi squares when sex of student was controlled.

CROSSTABS SIGNIFICANT FOR BOTH MALE AND FEMALE STUDENT RATERS

Items assessing the instructor's ability to sense boredom and the instructor's helpfulness when students have a problem resulted in statistically significant chi squares for responses by both male and female students. Regarding the instructor's ability to sense when students are bored, both student groups evaluated female faculty as likely to sense the boredom easily (56.4% by male students; 75.8% by female students), while male faculty tended to be evaluated as unaware of student boredom (40.5% by male students; 50.4% by female students). Judging the instructor's response to students with problems, male and female students were more likely to rate female faculty as helpful rather than unconcerned (74.1% by male students; 77.6% by female students).

CROSSTABS SIGNIFICANT FOR MALE STUDENT RATERS ONLY

When asked whether the professor tends to present his/her own beliefs or the beliefs of others, the chi square was significant for male students ($p < .05$), but not for female students. The crosstab revealed that male students felt that male faculty members were more likely to present the beliefs of others (47.2%), while the comparable figure for female faculty was 16.4%.

CROSSTABS SIGNIFICANT FOR FEMALE STUDENT RATERS ONLY

The remaining six items had statistically significant chi squares for female students evaluating instructors but resulted in nonsignificant crosstabulations when male student responses were evaluated. Female students rated female faculty as demonstrating a teaching style that makes students feel involved rather than inferior (75.4% for female instructors; 23.9% for male instructors) and a style of presentation that is more assertive than uncertain (89.6% for female instructors; 51.1% for male instructors).

Regarding the provision of feedback to students, female students noted significant differences in male and female instructors (48.2% of female faculty were judged sympathetic compared to 33.3% of male faculty). When asked whether faculty tended to deal with students without regard to the individual or with sensitivity to their needs, female students rated female faculty as significantly more likely to demonstrate sensitivity (60.3%) compared to male faculty (28.9%).

Finally, evaluations of instructor reactions to students with different interests and abilities suggest that female faculty are viewed as more sensitive to these differences (68.5%) compared to male faculty (20.0%) and the classroom atmosphere in courses with female instructors was rated as more open to discussion (82.7%). The comparable figure for male instructors was 63.1%.

Summary

The research reported in this paper was designed to assess the extent to which students' evaluation of male and female faculty conformed to sex stereotypes. Courses were selected controlling for course level (Introductory, general education, sociology and anthropology courses) and

fulfilling the criterion of multiple sections of the same course offered by both male and female instructors. Preliminary analyses revealed no significant differences in the student demographic composition of these classes and no significant differences in students' perception of male and female faculty education, rank, and tenure. Having established no manifest differences in the students across the seven course sections surveyed or in the perceived status of the faculty teaching these classes, subsequent analyses were conducted to test hypotheses regarding differences in male and female teaching attributes and differences in male and female student evaluations of faculty.

Based on the analyses measuring differential evaluations for teaching attributes demonstrated by male and female faculty, 9 of the 12 items showed statistically significant differences. No significant differences were noted between male and female faculty in terms of their adaptability while emphasizing relationships among various topics, their tendency to grade fairly and impartially, and their willingness to openly discuss student questions and differing opinions offered by students.

Both male and female students discerned differences in male instructors compared to female instructors when it came to sensing when students were bored and in their helpfulness when students have a problem. In each case, female faculty were perceived to be significantly more aware of student boredom and more likely to assist students with a problem. The only teaching attribute rated as significantly different for male and female faculty by male students, but not by female students, assessed the instructor's tendency to present the beliefs of others rather than their own beliefs. For this item, male students felt that male faculty were more likely to present the beliefs of others than female faculty were.

Analysis of the remaining six items revealed that female students rated male and female faculty as significantly different while male students did not differentiate between male and female faculty on these same items. The teaching attributes evaluated by female students as differentially exhibited by male and female instructors were the tendency for female instructors:

- to be more assertive rather than uncertain when presenting material;
- to be more sympathetic rather than objective in their feedback;
- to deal with students with sensitivity to their needs rather than ignoring the individual;
- to be helpful rather than unconcerned when students have a problem;
- to make students feel involved in the class rather than inferior;
- to have a classroom atmosphere more open to discussion rather than dominated by the instructor's beliefs.

Conclusions

When instructors begin to teach a subject they usually focus on the words selected, the amount of time spent on each topic, and the structure of the course. They also tend to find their own comfort zone with how much interaction they will have with the students and the form this interaction will take. Most instructors do not concentrate on the role sex

will play in determining the way that students will interpret their actions and words. Certain classes or topics, such as gender, may cause an instructor to pause and reflect on sex more closely but in survey courses and those with a majority of information presented in gender neutral topics, such as introduction to sociology, the instructor attends to material and presentation rather than how he or she will be perceived.

Unlike past research which indicated limited influences of sex on evaluations, the results presented above indicate that even the structure of the course, the general style of the instructor, and the level of interaction are perceived based on the sex of the instructor. Past studies indicated that there are differences in how an instructor is evaluated based on rapport and interaction types (Feldman 1993; Dukes & Victoria 1989), however the influence of sex on these factors was not examined. When sex is considered as an influence on the elements of teaching which form the evaluations of instructors, sex does emerge as an important influence. Students seem to be differentiating between male and female instructors when describing how they teach.

Differences between male and female students' evaluations of faculty were found, however female students were more likely to evaluate male and female instructors differently than male students. These findings are consistent with information found in other areas such as the workplace. Females appear to be applying sex stereotypes and therefore evaluate based on sex (Hartman et al. 1992; Roberts & Hoeksema 1989) Whether these differential evaluations are due to females invoking gender stereotypes or preferring instructors that are more similar to themselves is indeterminable by the current analysis. The results of past research would suggest that female students appear to be applying stereotypes more than males students.

Since course evaluations contribute, in varying degrees, to the general evaluation of faculty for promotion, tenure, and salary the biases identified here have compensatory implications. If women are evaluated differently than men as a result of gender expectations then how do these differences translate into rank and salary differences?

Similar to the class discussions held by Moore (1998) regarding the different reactions students had to the male professor and herself, instructors may choose to open a dialogue about the differences in the styles of male and female instructors and how those styles can become translated into evaluations. The discussion can also be used as a springboard for discussing evaluation differences outside of academia. In addition to discussing these issues in the classroom, administrators need to become aware of these biases when using student evaluations to make career decisions.

References

- Baker, Phyllis and Martha Copp. (1997). "Gender Performance Matters Most: The Interaction of Gendered Expectations, Feminist Course Content and Pregnancy in Students' Course Evaluations." *Teaching Sociology* 25(1): 29-43.

- Basow, Susan A. and Kristi Medcalf (1988). "Academic Achievement and Attributions Among College Students: Effects of Gender and Sex Typing." *Sex Roles* 19:555-567.
- Bem, Sandra L. (1981). *Bem Sex-Role Inventory: Professional Manual*. Palo Alto, CA: Consulting Psychologists Press.
- Broverman, I. K. S R Vogel, D. M. Broverman, F.E. Clarkson and P.S. Rosenkrantz. (1972). "Sex-role stereotypes: A current appraisal." *Journal of Social Issues* 28(2): 59-78.
- Deaux, Kay and Laurie L. Lewis. (1984). "Structure of Gender Stereotypes: Interrelationships Among Components and Gender Label." *Journal of Personality and Social Psychology* 46(5): 991-1004.
- Dukes, Richard L. and Gay Victoria. (1989). "The Effects of Gender, Status, and Effective Teaching on the Evaluation of College Instruction." *Teaching Sociology* 17: 447-457.
- Feldman, Kenneth A. (1993). "College Students' Views of Male and Female College Teachers: Part I - Evidence from Students' Evaluations of Their Classroom Teachers." *Research in Higher Education* 34: 151-211.
- Hartman, Sandra J., Rodger W. Griffeth, Lynn Miller, and Angelo J. Kinicki. (1992). "The Impact of Occupation, Performance, and Sex on Sex Role Stereotyping." *The Journal of Social Psychology* 128(4): 451-463.
- Martin, Elaine. (1984). "Power and Authority in the Classroom: Sexist Stereotypes in Teaching Evaluations." *Signs: Journal of Women in Culture and Society* 24: 128-133.
- McGlashan, Kay E., Patrick M. Wright, and Blaine McCormick. (1995). "Preferential Selection and Stereotypes: Effects on Evaluation of Female Leader Performance, Subordinate Goal Commitment, and Task Performance." *Sex Roles* 33(9/10): 669-686.
- Moore, Melanie. (1997). "Student Resistance to Course Content: Reactions to Gender of the Messenger." *Teaching Sociology* 25: 128-133.
- Roberts, Tomi-Ann and Susan Nolen-Hoeksema. (1989). "Sex Differences in Reactions to Evaluative Feedback." *Sex Roles* 21(11/12): 725-747.
- Russell, Joyce E. A. and Michael C. Rush. (1987). "The Effects of Sex and Marital/ Parental Status on Performance Evaluations and Attributions." *Sex Roles* 17(3/4): 221-236.
- Schein, Virginia. E. (1973). "The Relationship between Sex-Role Stereotypes and Requisite Management Characteristics." *Journal of Applied Psychology* 57:95-100.