

Student Ratings of Women Faculty: Research & Strategies*

I. Objectives

The purpose of this document is to:

- list concerns about student ratings identified by some women faculty in traditionally male-dominated fields
- provide readers with an overview of research findings about interactions between instructor gender and student ratings (SRs) of teaching
- provide suggestions for responding to suspicion of gender bias

II. Concerns

Many women faculty are concerned about how students' perspectives of faculty authority and credibility will impact student ratings of their courses. Women faculty are also concerned about the impact of students' *expectations* on student ratings and about setting boundaries for appropriate levels of contact and accessibility. Also of concern to most faculty, not just female faculty, is how student ratings are interpreted and used by peers and administrators. All of these issues are of concern because at many institutions, students' ratings of instruction (also called student evaluations of teaching or SETs) are the only measure of teaching effectiveness considered in tenure and promotion decisions.

III. Student Ratings Research Comparing Male and Female Faculty

Why do we keep hearing that there is no significant difference between male and female faculty SR data, when many women faculty, particularly those teaching in STEM fields, are convinced that "something is going on" in their courses and their student ratings?

A. "No Significant Difference" Research Results

- Tremendous numbers of studies document that student ratings are reliable indicators of *student satisfaction with their learning experiences*.
- Substantial evidence exists that student ratings are positively, consistently, and significantly related to student learning.
- In studies that analyze large samples of courses from a variety of disciplines, the consistent result is that there are no significant differences in ratings due to systematic gender bias.
- Female and male faculty do not appear to be rated higher or lower by students as a result of their their gender.
- Male and female student ratings habits are not, on average, significantly different.

(Cashin 1995; Feldman 1992, 1993; Greenwald 1997; McKeachie 1987; Seldin 1999; Theall & Franklin 1990)

* This document originally created August 2003, by A. Linse for an NSF ADVANCE program workshop.

However:

- Large-scale studies best reflect methods that explicitly attempt to identify common behaviors and results, i.e. they report what the majority of cases suggest. (Theall, M., 5/2/2000, POD Archives <<http://listserv.nd.edu/archives/pod.html>>).
- As in all large studies, instances of variance can be overwhelmed by large samples (cf. the Central Limit Theorem). When female faculty constitute a small percentage of an academic unit's faculty, it may well be an "instance of variance" that results from bias, and which would be difficult to detect.
- Gender differences among instructors may be more related to teaching approaches than to instructor effectiveness and teaching approaches might be gender-related (Centra and Gaubatz 2000).
- "There is a low relationship between single general items and specific items, and...the single general items had a much higher relationship to descriptive variables (gender, status, required-vs-elective, etc.) than did the specific items" (Arreola 1995: 86).

1. Is gender bias detectable in student ratings?

Few, if any, standard student ratings forms elicit information about student expectations. Student ratings systems are designed to measure common teaching behaviors. They are not designed with the intention of detecting biases.

Differences in male and female students' ratings are not generally detectible in field studies because students' gender is generally not recorded (Bassow 1994; Bassow and Silberg 1987). Thus, in studies comparing student ratings of male and female faculty, differences in male and female *students'* ratings would average to "no significant difference." Consequently, statistical insignificance may obscure that different standards may be used for male and female faculty.

2. What could explain lower ratings for women faculty?

On an individual-by-individual basis, the numbers may reflect that students are less satisfied with the instructor and/or the course for a variety of reasons, including:

- 1) Instructor's Teaching Abilities (e.g. organization, presentation skills, approachability, testing/feedback, etc.; Cashin 1995; Greenwald 1997; McKeachie 1987; Seldin 1999).
- 2) Gender-based student expectations (Anderson and Miller 1997; & citations in next section).
- 3) Contextual Factors (e.g. class size, required/elective status, content, course level, experience, or discipline). For example, one study found that in some cases women faculty were disproportionately assigned required, lower division, large enrollment courses. Their average ratings were lower than the ratings of male faculty teaching upper level, elective seminars. This is an expected result for any group assigned to teach such courses. The ratings were reliable and in line with what the literature suggests, but due to *administrative gender bias* in course assignments, women were placed at a disadvantage and further, their average ratings could have been misused to verify misconceptions. This is gender bias, but its source is not student raters or the evaluation process (Franklin and Theall 1994).
- 4) Negative Instructor feedback. One study demonstrates students' evaluations of female instructors are more dependent on the grades they have received from them than are students' evaluations of male instructors (Sinclair & Kunda 2000). Other research has documented a small positive correlation between grades and student ratings of teaching (Greenwald and Gilmore 1997).

B. Gender-Effects Research Results

Student ratings have been studied under the assumption that the underlying determinants of end-of-term ratings for female and male instructors are similar. Based on these studies, tenure and promotion decisions are being made with the assurance that there is no significant gender bias present in the student ratings system. But what if the assumption of equivalent determinants is not valid? (Kardia et al. 2001)

When the generalization of 'no bias' seems to be violated, fair practice demands that we investigate other possible reasons for the results. The effort required is more complex and demanding than simply looking at average scores for men and women and it is an absolute necessity if we want fair and unbiased decisions about faculty performance. (Theall 1999)

This section provides a compilation of the gender-effects research findings. The findings are grouped into broad categories (general, norms & stereotypes, student expectations, and disciplinary climate). A complete list of references cited is provided at the end of the document.

General

- There are gender-effects, but they are neither uniform, nor simple. However, even in studies that document significant gender-effects, teaching effectiveness criteria have the greatest effect on student ratings (Bachen et al. 1999; Hancock et al. 1993).
- When gender-interactions are present, typically it appears to reflect a same sex preference (Bachen et al. 1999; Bassow 1994; Centra and Gaubatz 2000; Hancock et al. 1993).
 - Male students rate female faculty lower (perhaps because of a mismatch in gender-related expectations; perhaps because male student learning preferences better match male faculty teaching styles). Effects may be more pronounced for males in male-dominated fields or in fields with a large percentage of males who have more traditional stereotypes of women. The results, however, are not consistent across all studies.
 - Female students rate female faculty higher (perhaps because of a match between preferred female teaching & learning styles and/or greater emphasis by female faculty on teaching styles that engage students and less dependence on traditional lecturing).
 - Male and female students generally do not rate male faculty significantly different (i.e. women and men students rate male faculty the same, but they do not rate women faculty the same). *This indicates a gender-effect that women faculty are perceived differently than male faculty.*

Norms & Stereotypes

- Male faculty are the "norm" in the academy, particularly in Science, Engineering, and Technology courses, i.e. they fit the prototype of faculty while females do not. Thus, female faculty are "marked for gender in ways males are not" (Bassow 1994). Male faculty are seen as "faculty," while women faculty are "women faculty" (Bassow 1992; Street et al 1996).
- Stereotypical expectations of women (e.g. to be nurturing and warm) overlap very little with stereotypical expectations of professors (e.g. to be knowledgeable and competent), but are aligned with stereotypical expectations of men (Bassow 1994; Kardia et al. 2001).
- Students expect female professors to excel both in stereotypically masculine and feminine traits (Freeman 1994; Kierstead et al. 1988).

- Faculty, especially women faculty, that adhere to 'gender-appropriate' models may be rewarded with higher evaluations (Freeman 1994; Martin 1984; Statham et al. 1991).
- While warmth and interpersonal contact are important for all faculty, only female faculty evaluations were influenced by these perceptions (Kierstead et al. 1988; Bennett 1982; Downs and Downs 1993).
- Students attribute the success of male professors to being effective and powerful and that of females to being concerned and likeable (Kaschak 1981).
- Male instructors are significantly more likely to be viewed as intelligent; female instructors needed to *prove* their intelligence, especially when seen as a problematic teacher. Problematic teaching behaviors are significantly more likely to be excused in male instructors than in female instructors (Kardia et al. 2001).
- Students who have received low grades from female faculty are more likely to invoke negative stereotypes of women as a means of discrediting them (Sinclair & Kunda 2000).

Student Expectations

- Research based on focus group interviews with faculty and students, supports what many women faculty sense, that male and female students have different expectations of them (Bachen et al 1999; Cook et al. 2000, Kardia et al. 2001).
- Gender appears to influence student evaluations when the professor's behavior in class somehow contradicts student expectations of male and female behaviors (Bassow and Silberg 1987; Bennett 1982; Freeman 1994).
- Female professors may be judged negatively if they are not perceived as more interested in and available to students relative to male professors (Bachen et al 1999; Bennett 1982; Statham et al. 1991). Even when women faculty are available and perceived as interested, they generally do not receive higher average ratings.
- When women faculty display more caring and warm behavior, students may interpret this behavior as weak or less valuable (Bassow and Silberg 1987; Sandler 1991).
- In a large study that included classroom observations, student evaluations, and interviews with professors (Statham et al. 1991), students rated male and female professors as equally effective. Both male and female faculty may be "penalized" for not meeting expectations.

Disciplinary Climate

- Students in certain fields may require women faculty to meet more stringent credibility criteria by virtue of their sex (Bassow and Silberg 1987; Feldman 1992, 1993).
- Female professors are judged more highly on "feminine" attributes when they were in feminine stereotyped fields, while men were not rated differentially as a function of their field (Kaschak 1981).
- Female professors had lower competency ratings as well as lower "global" ratings, than male professors even while controlling for student's sex, GPA, expected grade, discipline, and course size (Basow and Silberg 1987; Sidanius and Crane 1989).
- In one of the rare examinations of gender-effects by study area, both the colleges of Engineering and Science & Mathematics show significant effects. In Engineering, effects were reported for instructor gender ($p = .025$), student gender ($p = .002$), and instructor x student gender multivariate interaction ($p = .021$). In Science and Mathematics, effects were reported for instructor gender ($p < .001$), student gender ($p < .001$), but not for instructor x student gender ($p = .083$).

IV. What should I do if I suspect a gender effect in my student ratings?

- 1) **Rule out that other factors are not an issue.** Student Ratings should *never* be the sole source of data used in performance appraisal (Anderson and Miller 1997). Consider the following:
 - Do evaluations include criteria beyond basic presentation and delivery skills? These skills are associated with traditional lecturing—a style that may be more favorably associated with male faculty.
 - Do evaluations include review of course plans, course design, and instructional development activities?
 - Do evaluations focus on learning outcomes and supporting assessment data?
 - Do evaluations consider faculty contact with and mentoring of students?
- 2) Be deliberate in your selection of student ratings forms or questions (if possible); choose forms that include questions about the teaching methods you use and that reflect characteristics that are important to you.
- 3) Be responsible for interpreting your own student ratings. Do not leave it to anyone else to decide the meaning of your students' ratings and comments. You were in class with them, so you are more likely to be able to interpret what the scores and comments indicate.
- 4) Ask other women faculty for strategies to compensate for differential expectations. Strategies used by male faculty may not be applicable or equally effective for women faculty.
- 5) Identify and align faculty and student expectations on the first day of class.
- 6) Consult with experts on instruction; e.g. consult with your campus teaching center, seek information posted on the internet by other campus teaching centers and other ADVANCE programs. If your campus does not have a teaching center, consider seeking advice from colleagues in Education.

References Cited

- Anderson, Kristi, and Elizabeth D. Miller (1997) "Gender and student evaluations of teaching," *Political Science & Politics*, 30(2): 216-219, June 1997.
- Arreola, Raoul A. (2000) *Developing a Comprehensive Faculty Evaluation System: A Handbook for College Faculty and Administrators on Designing and Operating a Comprehensive Faculty Evaluation System*, 2nd edition. Bolton Massachusetts: Anker.
- Bachen, Christine M., Moira M. McLoughlin, and Sara S. Garcia (1999) "Assessing the Role of Gender in College Students' Evaluations of Faculty," *Communication Education*, 48(3): 193-210, July 1999.
- Basow, Susan A. (1994) Student Ratings of Professors are Not Gender Blind. *AWM Newsletter*, 24(5), Sept.-Oct. 1994. <http://www.awm-math.org/newsletter/199409/basow.html>
- Basow, Susan A. (1992) *Gender Stereotypes and Roles*, 3rd edition, Pacific Grove, CA: Brooks/Cole Publishing.
- Basow, Susan A. and Nancy T. Silberg (1987) "Student Evaluations of College Professors: Are Female and Male Professors Rated Differently?" *Journal of Educational Psychology*, 79(3): 308-14, September 1987.
- Bennett, Sheila Kishler (1982) "Student Perceptions of and Expectations for Male and Female Instructors: Evidence Relating to the Question of Gender Bias in Teaching Evaluation," *Journal of Educational Psychology*, 74(2): 170-79, April 1982.
- Cashin, William E. (1995) Student Ratings of Teaching: The Research Revisited. IDEA Paper No. 32. Center for Faculty Education and Development. IDEA Center, Kansas State University. <<http://www.idea.ksu.edu/products/Papers.html>>
- Centra, John A., and Noreen B. Gaubatz (2000) "Is There Gender Bias in Student Evaluations of Teaching?" *Journal of Higher Education*, 71(1): 17-33, January-February 2000.

- Cook, Constance, Mary Wright, Diana Kardia, and Leigh Bowinik (2000) "The Relevance of Gender to Authority for Women Faculty," Presented at the annual meeting of the Association for the Study of Higher Education, November 2000.
- Downs, V.C, and T.M. Downs (1993) An Exploratory and Descriptive Study Identifying Communicative Behaviors Associated with Effective College Teaching Paper presented at the annual meeting of the International Communication Association. May 1993, Washington D.C.
- Feldman, Kenneth A. (1992) College students' views of male and female college teachers: Part I--Evidence from the social laboratory and experiments. *Research in Higher Education*, 33(3): 317-375.
- Feldman, Kenneth A. (1993) College students' views of male and female college teachers: Part II--Evidence from students' evaluations of their classroom teachers. *Research in Higher Education*, 34(2): 151-211.
- Franklin, Jennifer and Michael Theall (1994) "Student Ratings of Instruction and Sex Differences Revisited." Paper presented at the 78th annual meeting of the American Educational Research Association, New Orleans. April 7, 1994.
- Freeman, Harvey R. (1994) "Student Evaluations of College Instructors: Effects of Type of Course Taught, Instructor Gender and Gender Role, and Student Gender," *Journal of Educational Psychology*, 86(4): 627-30, December 1994.
- Greenwald, A. G. (1997) Validity Concerns and Usefulness of Student Ratings of Instruction. *American Psychologist* 52(11): 1182-1186.
- Greenwald, A. G. and Gillmore, G. M. (1997) Grading Leniency is a Removable Contaminant of Student Ratings, *American Psychologist*, 53: 1209-1217.
- Hancock, Gregory R., David M. Shannon, and Landa L Trentham (1992) "Student and Teacher Gender in Ratings of University Faculty: Results from Five Colleges of Study," *Journal of Personnel Evaluation in Education*, 6(3): 235-248, February 1992.
- Kardia, Diana, Dulcey Simpkins, Janetta Lun, and Louise August (2001) "Does Faculty Gender Matter to Students? Reflections on the Controversy Relating to Gender and Undergraduate Student Ratings," Presentation at the Annual Conference of the American Educational Research Association, April 2001.
- Kaschak, Ellyn (1981) "Another Look at Sex Bias in Students' Evaluations of Professors: Do Winners Get the Recognition that They Have Been Given?" *Psychology of Women Quarterly*, 5(5): 767-772.
- Kierstead, Diane, Patti D'Agostino, and Heidi Dill (1988) "Sex Role Stereotyping of College Professors: Bias in Students' Ratings of Instructors," *Journal of Educational Psychology*, 80(3): 342-344, September 1988.
- Martin, Elaine (1984) "Power and Authority in the Classroom: Sexist Stereotypes in Teaching Evaluations," *Signs: Journal of Women in Culture and Society*, 9(3): 482-492, Spring 1984.
- McKeachie, W. (1990) Research on College Teaching: The Historical Background. *Journal of Educational Psychology* 82; 189-200.
- Sandler, Bernice R. (1991) Women Faculty at Work in the Classroom, or, Why it Still Hurts to be a Woman in Labor. *Communication Education*, 40: 6-15.
- Seldin, Peter, and Associates (1999) *Changing Practices in Evaluating Teaching: A Practical Guide to Improved Faculty Performance and Promotion/Tenure Decisions*. Bolton, Massachusetts: Anker.
- Sidanius, Jim and M. Crane (1989) Job Evaluation and Gender: The Case of University Faculty. *Journal of Applied Social Psychology* 19: 174-197.
- Lisa Sinclair and Ziva Kunda (2000) Motivated Stereotyping of Women: She's Fine if She Praised Me but Incompetent if She Criticized Me, *Personality and Social Psychology Bulletin*, 26(11):1329-1342.
- Statham, Anne, Laurel Richardson, and Judith A. Cook (1991) *Gender and University Teaching: A Negotiated Difference*. Albany: State University New York Press.
- Street, S., E. Kimmel, and J.D. Kromrey (1996) Gender Role Preferences and Perceptions of University Students, Faculty, and Administrators. *Research in Higher Education* 37:615-632.
- Theall, Michael and Jennifer Franklin (editors) (1990) *Student Ratings of Instruction: Issues for Improving Practice*. New Directions for Teaching and Learning, 43. San Francisco: Jossey-Bass.
- Theall, 1999, March 5, POD Archives <<http://listserv.nd.edu/archives/pod.html>>.