Gender Bias in Faculty Teaching Evaluations: Key Journal Articles and Additional Resources

March 13, 2015

Original Literature Review: Journal Articles

1. Abel, M. H., & Meltzer, A. L. (2007). Student ratings of a male and female professors' lecture on sex discrimination in the workforce. *Sex Roles, 57,* 173-180.

   Male (n = 41) and female (n = 46) undergraduates evaluated an identical written lecture that they were told was by a male professor or a female professor. The lecture focused on pay disparities between men and women in the workforce and was intended to instill a perception of sex discrimination in the workforce. Both male and female students rated the male professor and his lecture more positively and less sexist than they rated the female professor and her lecture.

2. Arbuckle, J., & Williams, B. D. (2003). Students' perceptions of expressiveness: Age and gender effects on teacher evaluations. *Sex Roles, 49*(9/10), 507-516.

   This study investigated the relationship between students' perceptions of the expressiveness of professors and implicit age and gender stereotypes. Male and female undergraduates (*N* = 352) watched slides of an age- and gender-neutral stick figure and listened to a neutral voice presenting a lecture. Then, they completed teacher evaluation forms that indicated 1 of 4 different age and gender conditions (male, female, "old," and "young"). Students rated the "young" male professor higher than they rated the "young" female, "old" male, and "old" female professors on speaking enthusiastically and using a meaningful voice tone during the lecture.

3. Bachen, C. M., McLoughlin, M. M., & Garcia, S. S. (1999). Assessing the role of gender in college students' evaluations of faculty. *Communication Education, 48*(3), 193-210.

   In this study, researchers surveyed nearly 500 university students about their perceptions of male and female faculty, in order to investigate whether students' assessments of male and female professors are influenced by traditional gender schema. Female students rated female faculty especially high across five teaching dimensions and male faculty comparatively lower. By contrast, male students did not evaluate male and female professors as significantly different. Faculty assessments were further influenced by the strength of students' gender schema.

4. Baldwin, T., & Blattner, N. (2003). Guarding against potential bias in student evaluations: What every faculty member needs to know. *College Teaching*, *51*(1), 27-32.

   This article describes the influences of course evaluations on promotion, tenure, and merit decisions. It describes how faculty can combat biases to ensure that their teaching effectiveness is accurately portrayed. It reviews alternative evaluation methods, including portfolios, peer feedback sessions, and informal student surveys.

5. Basow, S. A. (1995). Student evaluations of college professors: When gender matters. *Journal of Educational Psychology*, *87*(4), 656-665.

   In this study, researchers analyzed student evaluations completed over a four-year period at a private liberal arts college for the effects of teacher gender, student gender, and divisional affiliation. Overall,

the ratings of male professors appeared to be unaffected by student gender. By contrast, female professors tended to receive their highest ratings from female students and their lowest ratings from male students. The interaction between student gender and teacher gender generally remained when possible confounding factors (e.g., teacher rank) were partialed out. The mean ratings that female and male professors received also varied as a function of the divisional affiliation of the course.

6. Basow, S. A. (2000). Best and worst professors: Gender patterns in students' choices. *Sex Roles, 43*(5/6), 407-417.

   In this study, researchers asked over 100 students to describe their best and worst professor. For "best" professors, female professors were chosen more by their female students and less by their male students than expected. The researchers found no gender differences in students' choice of "worst" professors. Students criticized similar qualities for their male and female professors (mainly a lack of organization and clarity). Students most often described "best" professors as caring and knowledgeable, but other qualities varied as a function of professor or student gender.

7. Basow, S. A., & Martin, J. L. (2012). Bias in student evaluations. In M. E. Kite (Ed.), *Effective evaluation of teaching: A guide for faculty and administrators* (pp. 40-49). Retrieved from the Society for the Teaching of Psychology web site: http://teachpsych.org/ebooks/evals2012/index.php.

   The authors describe and examine potential sources of bias in student evaluations of faculty teaching, such as professor gender, professor race/ethnicity, professor attractiveness, professor age, course difficulty, and expected grade.

8. Basow, S. A., & Montgomery, S. (2005). Student ratings and professor self-ratings of college teaching: Effects of gender and divisional affiliation. *Journal of Personnel Evaluation in Education, 18*, 91-106.

   At a liberal arts college, 20 female professors, 23 male professors, and 803 undergraduates participated in a questionnaire study of the effects of professor gender, student gender, and divisional affiliation on student ratings of professors and professor self-ratings. On student ratings, the researchers found main effects for both professor gender (students rated female professors higher than male professors on the two interpersonal factors) and division (students rated natural science courses lowest on most factors). Significant interactions existed between professor gender and division. Few significant correlations existed between professor self-ratings and students' ratings.

9. Basow, S. A., & Silberg, N. T. (1987). Student evaluations of college professors: Are female and male professors rated differently? *Journal of Educational Psychology, 79*(3), 308-314.

   More than 1,000 male and female college students of 16 male and female professors (matched for course division, years of teaching, and tenure status) evaluated their instructors on teaching effectiveness and sex-typed characteristics. Male students gave female professors significantly poorer ratings than male professors on the six teaching evaluation measures. Male students' ratings of female professors were poorer than those of female students on four of the six measures. On three measures, female students evaluated female professors less favorably. Student perceptions of a professor's instrumental/active and expressive/nurturant traits accounted for only a few of the gender-related effects.

10. Bennett, S. K. (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), 170-179.

In this study, 253 students in non-science introductory courses at a liberal arts college completed a course evaluation questionnaire. Women instructors were perceived as warmer and more potent individuals. This largely accounted for their higher formal student ratings in specific areas of teaching performance. However, students required women instructors to offer greater interpersonal support and judged them more closely than male instructors in providing it.

11. Benton, S. L., & Cashin, W. E. (2012). Student ratings of teaching: A summary of research and literature (IDEA Paper #50). Retrieved from http://ideaedu.org/sites/default/files/idea-paper_50.pdf.

    In this paper, the authors attempt to summarize the conclusions of the major reviews of the student ratings research and literature from the 1970s to 2010.

12. Centra, J. A., & Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching? *The Journal of Higher Education, 71*(1), 17-33.

    This study investigated gender differences in student evaluation of teaching through two analyses, based upon data from 741 college classes. In the first analysis, researchers compared female and male student ratings in the same classes for female and male instructors. In the second analysis, researchers examined how student ratings differed for male and female instructors. Female students gave higher ratings to female instructors on three of eight scales for all disciplines combined. Male students gave higher ratings to male instructors on only one scale. Male and female students showed no difference in their rankings of male teachers. For the total sample of classes, when students gave more favorable ratings, such ratings were largely given by female students to female instructors.

13. Chamberlin, M. S., & Hickey, J. S. (2001). Student evaluations of faculty performance: The role of gender expectations in differential evaluations. *Educational Research Quarterly, 25*(2), 3-14.

    This study was designed to identify gendered expectations, holding occupation constant, based on evaluations of instructors' performance in the classroom. Undergraduate students (N = 198) evaluated male and female instructors from the same disciplines, using the same questionnaire. Questions were selected based on gendered expectations. The researchers found that male and female instructors are evaluated differently on aspects of presentation and classroom structure.

14. Freeman, H. 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), 627-630.

    In the first experiment, researchers examined the effects of course type, student gender, and instructor gender and gender role on student evaluations of instructor effectiveness. In the second experiment, researchers explored students' perceptions of the importance of various gender role characteristics in instructors of different course types. Results suggested that instructor gender role is more important than instructor gender in affecting student evaluations. Female and male students preferred instructors (science instructors, in particular) who had both feminine and masculine characteristics, regardless of the instructor gender.

15. Kierstead, D., D'Agostino, P., & Dill, H. (1988). Sex role stereotyping of college professors: Bias in students' ratings of instructors. *Journal of Educational Psychology, 80*(3), 342-344.

    Researchers investigated the impact of three variables on students' ratings of instruction: social contact between instructor and students (present vs. absent), the instructor's facial expression (smiling

vs. neutral), and the instructor's sex. They presented subjects with hypothetical scenarios in which these factors varied while behaviors directly related to teaching were held constant. Behaviors indicative of friendliness toward students elevated students' ratings of instruction for female instructors but not for male instructors. Subjects rated the male professors as more effective than female professors. The findings support the view that students expect female instructors to excel in both stereotypically masculine domains (e.g., competence) and stereotypically feminine domains (e.g., warmth).

16. Laube, H., Massoni, K., Sprague, J., & Ferber, A. (2007). The impact of gender on the evaluation of teaching: What we know and what we can do. *NWSA Journal, 19*(3), 87-104.

    Some scholars have found gender to have no (or very little) influence on evaluations of teaching, whereas other scholars have found gender to affect evaluations significantly. The authors draw on insights from sociological scholarship on gender and evaluation and argue that this apparent inconsistency is itself an artifact of the way that quantitative measures can mask underlying gender bias. The authors offer concrete strategies for faculty, researchers, and administrators to improve the efficacy of the system of evaluation.

17. MacNell, L., Driscoll, A., & Hunt, A. N. (2014, December 5). What's in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education*, 1-13.

    In this study, assistant instructors in an online class each operated under two different gender identities. Students rated the male identity significantly higher than the female identity, regardless of the instructor's actual gender. This study has been widely cited in the media, but it has a number of methodological flaws (e.g., the instructors were not blind to whether they were assigned a "male" or "female" identity).

18. McPherson, M. A., Jewell, R. T., & Kim, M. (2009). What determines student evaluation scores? A random effects analysis of undergraduate economics classes. *Eastern Economic Journal, 35*, 37-51.

    In this study, researchers applied a feasible generalized least squares model to a panel of data from undergraduate economics classes. They found evidence that instructors could ''buy'' better evaluation scores by inflating students' grade expectations. Class size and instructor experience were important determinants of evaluation scores in principles classes, but not in upper-level courses. Male instructors received better evaluation scores than females, and younger instructors were more popular than older instructors. Other factors were also important determinants of evaluation scores.

19. Sidanius, J., & Crane, M. (1989). Job evaluation and gender: The case of university faculty. *Journal of Applied Social Psychology, 19*(2), 174-197.

    In this study, researchers examined the effects of students' and professors' sex on student evaluations of professors' teaching effectiveness. Researchers analyzed the ratings of over 400 faculty made by over 9,000 students and controlled for many variables. Results showed that (a) students gave male faculty significantly higher evaluations on global teacher effectiveness and academic competence than they gave female faculty; (b) female faculty were not rated as more sensitive to student needs than male faculty; and (c) students seem to place more weight on academic competence for male faculty than for female faculty when making overall global judgments of faculty performance.

20. Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research, 83*(4), 598-642.

This article reviews recent literature on student evaluation of teaching (SET) in higher education. It is based on the SET meta-validation model, drawing upon research reports published in peer-reviewed journals since 2000.

21. Sprague, J., & Massoni, K. (2005). Student evaluations and gendered expectations: What we can't count can hurt us. *Sex Roles, 53*(11/12), 779-793.

    In this study, the authors focused on the question of whether a teacher's gender impacts students' evaluations. From their critical evaluation of the research literature, the authors concluded that the form gender bias takes may not be easily detectible by quantitative scales. The authors conducted a qualitative analysis of the words that 288 college students at two campuses used to describe their best-ever and worst-ever teachers. Results provided indications that students hold teachers accountable to certain gendered expectations. The burdens that these expectations place on women teachers are more labor-intensive than the burdens that they place on men teachers. The authors observed signs of much greater hostility toward women than toward men who do not meet students' gendered expectations.

22. Wright, S. L., & Jenkins-Guarnieri, M. A. (2012). Student evaluations of teaching: Combining the meta-analyses and demonstrating further evidence for effective use. *Assessment & Evaluation in Higher Education, 37*(6), 683-699.

    Given that there is not a single study that summarizes research on student evaluations of teaching (SETs) with regard to their validity, susceptibility to bias, practical use, and effective implementation, the authors conducted a comprehensive overview of SETs by combining nine prior meta-analyses (covering 193 studies) related to SETs. This yielded a small-to-medium overall weighted mean effect size between SETs and the variables studied. The authors' findings suggest that SETs appear to be valid, have practical use that is largely free from gender bias, and are most effective when implemented with consultation strategies.

23. Young, S., Rush, L., & Shaw, D. (2009). Evaluating gender bias in ratings of university instructors' teaching effectiveness. *International Journal for the Scholarship of Teaching and Learning, 3*(2), Article 19. Retrieved from http://digitalcommons.georgiasouthern.edu/ij-sotl/vol3/iss2/19.

    Undergraduate students and graduate students (n = 765) rated instructors on three factors: interpersonal characteristics, pedagogical characteristics, and course content characteristics. The authors analyzed group differences based on student gender, instructor gender, and student level. Ratings of pedagogical characteristics and course content characteristics yielded significant interactions between student gender and instructor gender. However, the researchers found no differences among groups on interpersonal characteristics.

Additional Literature Review: Journal Articles (focused on studies in which female instructors received better teaching evaluations than male instructors did)

1. Rowden, G. V., & Carlson, R. E. (1996). Gender issues and students' perceptions of instructors' immediacy and evaluation of teaching and course. *Psychological Reports, 78*, 835-839.

    This study found that for 197 undergraduates (105 women and 92 men), instructor immediacy, course evaluation, and evaluations of instructors were positively correlated. Female instructors received higher teacher ratings and course ratings than their male peers. In addition, female students with male instructors rated the course lower than any other gender grouping.

2. Whitworth, J. E., Price, B. A., & Randall, C.H. (2002). Factors that affect College of Business student opinion of teaching and learning. *Journal of Education for Business*, May/June, 282-289.

   In this study, the authors analyzed faculty evaluations from 12,153 students to investigate the effects of faculty gender, course type, and course level (graduate versus undergraduate) on the faculty evaluations. They found that female instructors received higher ratings than male instructors and that ratings differed significantly by course type and by students' perceived amount of learning.

3. Wigington, H., Tollefson, N., & Rodriguez, E. (1989). Students' ratings of instructors revisited: Interactions among class and instructor variables. *Research in Higher Education, 30*(3), 331-344.

   Students at a midwestern university completed a Curriculum and Instruction Survey (C&I) form during the 1985-86 academic year. The researchers identified four types of classes (lecture, lecture-discussion, discussion, and laboratory) in sufficient quantity to be analyzed. They obtained a total of 5,843 evaluations for 242 different classes. They studied a number of class and instructor variables, as well as the interactions among those variables. One finding in this study was that overall, female instructors received higher ratings than male instructors did.

Benchmarking: Work of Other NSF ADVANCE Institutions

In the literature, I did not find any research studies on gender bias in faculty teaching evaluations by other NSF ADVANCE institutions. However, I found the following examples of initiatives by other NSF ADVANCE institutions to reduce the potential effect of gender bias in faculty teaching evaluations:

1. Cornell University. (2011). Year-End Report for the Cornell University ADVANCE Institutional Transformation Grant. Year 5: August 1, 2010 – October 31, 2011. "Advancing Cornell's Commitment to Excellence and Leadership." Retrieved from http://advance.cornell.edu/documents/Year5-Annual_Report.pdf.

   This report describes how CU-ADVANCE, in collaboration with the Cornell Interactive Theater Ensemble, is developing a new training DVD and workshop about gender bias and other challenges in the tenure and promotion evaluation process. This includes the topic of gender bias in teaching evaluations (see p. 30 of PDF document).

2. Linse, A. R. University of Washington Center for Engineering Learning & Teaching. (2003). "Student Ratings of Women Faculty: Data and Strategies." Retrieved from https://advance.washington.edu/apps/resources/docs/20030513-student_ratings_ds.pdf.

   This document is intended to (a) list concerns about student ratings identified by some women faculty in SEM disciplines, (b) provide readers with research findings about interactions between instructor gender and student ratings of teaching, and (c) provide suggestions for responding to suspicion of gender bias.

3. Shipley, S. University of Illinois at Chicago. (March 4, 2013). "Implicit Bias Panel: NSF ADVANCE Program Workshop." Retrieved from http://www.portal.advance.vt.edu/index.php/authors/brooke-shipley.

   This is a summary of UIC's WISEST Program activities with regard to implicit bias.

4. University of Illinois at Chicago. (2009). WISEST Women in Science and Engineering System Transformation Quarterly Report, Year 4, First Quarter (August 1, 2009 – October 31, 2009). Retrieved from http://www.uic.edu/depts/oaa/wisest/docs/Yr%204%20Qtr%201%20report%20final.pdf.

   In this report, UIC describes its initiatives with regard to gender bias in faculty teaching evaluations.

5. University of Illinois at Chicago. (2010). WISEST Women in Science and Engineering System Transformation Quarterly Report, Year 4, Second Quarter (November 1, 2009 – January 31, 2010). Retrieved from http://www.uic.edu/depts/oaa/wisest/docs/FINAL%20NSF%20YR%204%20Q%202%20Jan%2031%202010%20submitted%20Jan%2029%20MK.pdf

   In this report, UIC describes its initiatives with regard to gender bias in faculty teaching evaluations.

6. University of Nebraska-Lincoln. ADVANCE Nebraska Faculty Committee 2012. (2012). "ADVANCE-Nebraska Annual Evaluation of Faculty Best Practices." Retrieved from http://advance.unl.edu/files/annualevalutationoffaculty3_2013.pdf.

   On page 1, in the "Ensure Transparency – Avoid implicit bias" section, this document states, "Student evaluations should not be a major consideration in the quality of a faculty member's teaching. Student assessment has been extensively documented to reflect implicit bias that negatively impacts specific ethnic, age and gender demographics. Furthermore, such data vary directly with course assignment."