This paper examines the relationship between instructor gender (as an external status characteristic) and student rating of instructor (both as teacher and advisor). Current research on the relationship between gender and evaluation is summarized. Three paradigms are suggested based on research findings: (1) no action be taken as sex of the professor is not an important variable in determining how students rate instruction; (2) individual faculty members must make the necessary behavioral adaptations in order to receive high ratings; and (3) a campus- and curriculum-wide sharing of responsibility to confront sexism. Since the everyday art of teaching is reflected in and guided by the feedback faculty receive from students, the role gender plays in teacher evaluations is clearly important. Throughout this paper, the term "sex" and "gender" are used interchangeably, as they are frequently used in the current literature. (Contains 38 references.) (EH)
He’s Tough But She’s Mean:
Gender Differences in Student Evaluations

Jean Garrity Hopefl
Cheri Frey-Hartel
Cardinal Stritch College

Presented at the UW System
Women’s Studies Conference
October 28, 1995

BEST COPY AVAILABLE
Interest in the impact of instructor gender on student evaluations grew out of Berger's Expectation States Theory (as presented by Andrews, 1992). According to Berger, external status characteristics such as gender, race, age, ethnicity, social class and sex impact how an individual is perceived, interpreted and evaluated. Higher status (as perceived by external characteristics) is typically associated with predictions of higher performance. This paper examines the relationship between instructor gender (as an external status characteristic) and student rating of instructor (both as teacher and advisor). Current research on the relationship between gender and evaluation is presented, and three paradigms for action reflecting the various research findings are suggested. Since the everyday art of teaching is reflected in and guided by the feedback faculty receive from students, the role gender plays in teaching evaluations is clearly important. Throughout this paper, the terms "sex" and "gender" will be used interchangeably, as they are frequently used in the current literature.

Teaching effectiveness has been identified as one of if not the primary criterion for promotion of faculty. According to the Carnegie Foundation for the Advancement of Teaching's 1989 National Survey of Faculty, 62% of all faculty respondents from research, doctorate-granting, comprehensive, liberal arts, and two year institutions agreed that teaching effectiveness should be the primary criterion for promotion of faculty (Boyer, 1990, p. 32). Percentages of faculty identifying teaching effectiveness as most important ranged from 21% at research institutions to 76% and 92% at liberal arts and two year institutions (Boyer, 1990, p. 32).

A number of different methods have been used to assess teaching effectiveness: self-assessment, peer assessment, and student assessment. Of these, student assessment has been identified as most important in the granting of tenure for all types of institutions except two year where observations of teaching by colleagues and/or administrators was identified as more important (Boyer, 1990, p. 30). The importance of student evaluations of courses taught for granting tenure is identified as particularly significant in liberal arts colleges where 90% of faculty identified student evaluations as either very important (45%) or fairly important (45%) for granting of tenure (Boyer, 1990, Table A-6).

For these liberal arts schools student evaluations emerge as the most important criterion in tenure decisions because they were identified as "very important" by 45% of the faculty responding followed by recommendations by other faculty within my institution (38%), recommendations from current or former students (30%), observations of teaching by colleagues and/or administrators (29%) and service within the university community (27%).
grouping of criterion was ranked as "very important" by a significantly smaller percentage of the faculty. Recommendations from outside scholars was identified as "very important" by only 16%, followed by academic advisement (15%), syllabi for courses taught (14%), service within the scholar's discipline (11%) and research grants received by the scholar (9%) (Boyer, 1990, p. 30).

Research has identified that male and female faculty may view the importance of student evaluations differently. Utilizing the findings of 250 completed questionnaires from faculty teaching at the four campuses of the University of Colorado, it was found that male faculty identified feedback from students as more helpful in improving of teaching than did female faculty. (Goodwin and Steven, 1993, p. 178).

A review of recent literature has revealed three distinct lines of thought on the relationship between instructor gender and student evaluation of instructor. One group of researchers has concluded that female faculty are evaluated lower than males because of their gender. A second group contends that an instructor's gender in itself is not consistently a predictor in evaluations, and that there are other variables that must be considered. A final group has concluded that it is actually females who receive higher evaluations than males.

First, some research indicates that female faculty are evaluated lower than their male counterparts because they are female. Goldberg’s (1968) study appears to be the foundation for this school of thought. In his work, he concluded that a bias exists, where subjects preferred a male author over a female author. He also found that art produced by a male artist was rated higher than a piece signed by a female artist. However, when subjects were told that the art work had won a contest, the effects for gender were nullified.

Minnow (1988) concluded that students tend to favor male teachers over females. Several reasons for this tendency have surfaced in the literature. First, Basow and Silberg (1987) stated that females who do not demonstrate typically feminine behaviors are ranked lower than those who do. Martin (1984) found the same to be true for males demonstrating stereotypical masculine behaviors. However, the difference is pointed out in Martin (1976), who concluded that stereotypical masculine behaviors are generally seen as healthier and superior to stereotypical feminine behaviors. Given the datedness of Martin’s study, however, one can only hope that these attitudes are changing, particularly on college campuses.

Further evidence does lend support to the hypothesis that stereotypical feminine behaviors are highly valued in female teachers. For example, women are expected to spend more free time with students than are men. Bennett (1982) found that females are
evaluated low if they do not meet student expectations of time and attention, yet the Bennett study also indicated that students do not appreciate greater time and attention from men. In addition, society holds an expectation that women are supposed to be nice.

While women are expected to demonstrate behaviors such as friendliness, support and deference, males may be perceived as good or exceptional teachers without demonstrating any degree of warmth toward their students (Minner, 1988). Difficult female graders were perceived more negatively than easy female graders in Unger’s (1979) study, but grades did not surface as a variable for men. Kierstead, D’Agostino, and Dill (1988) stated that whether an instructor smiled or was sociable affected females’ evaluations, but did not affect males’ evaluations.

Women are often expected to be better teachers overall. Students tend to be more tolerant of what is described as a lack of "formal professionalism" in males than in females, and students expect more formal preparation and organization from females (Bennett, 1982). Females have been ranked lower than males on such items as clarity of presentation (Wilson and Doyle, 1976), organization (Goebel and Cashen, 1979), and on presenting a stimulating course (Centra, 1979). Sandler (1991) explains that even when women include more participation and interaction in their classes, they are evaluated as less competent instructors than men who utilize a more traditional, less interactive format.

Although this research is somewhat alarming for female faculty, a closer examination of the methodology utilized in some of the above mentioned studies reveals several problems. Dukes and Victoria (1989) have criticized Bennett’s work and others since their studies relied on survey data, which did not untangle the intercorrelations between factors other than genders. Bennett, for example, labeled other factors that surfaced in her work as "intervening variable" and did not treat them concurrently with gender.

A second line of thought, then is that gender is not consistently a predictor in student evaluations of faculty, and that in fact, when examining gender alone, ratings for males and females tend to be equal. There must be variables other than gender then, or perhaps, an interaction between gender and other variables. For example, some research indicates that on the college level, there is less gender bias in studies on evaluation. Students at the college level, because of their experience with the educational system, tend to be competent judges of what they perceive their product to be. Also, society tends to attribute high status to college professors, thereby eliminating the impact of feminine or masculine stereotypes (Dukes and Victoria, 1989). There also appears to be a relationship between gender of student evaluator and gender of instructor. Rankin (1981) found that male students rated professors of both genders equally, but female
students rated female faculty members more highly. Ferber and Huber (1975) found that students rated same-sex faculty members more highly.

Standardized teaching evaluations are affected by a number of factors beyond gender of instructor. Marsh (1980) found that course characteristics, characteristics of students and characteristics of teachers (other than gender) account for 12-14 percent of the variance of ratings of instructors. Years of teaching experience also have an impact on evaluations. Those instructors with from two to twelve years of experience tend to receive higher evaluations than those with less or more experience (Dukes and Victoria, 1989). Martin (1984) found that the highest rankings given by students went to those instructors who combined feminine traits such as friendliness with masculine traits such as preparedness. Gundykunst (1982) and Elmore and LaPointe (1975) found that the overall relationship between instructor and student was the best indicator of student rating of instruction.

Dukes and Victoria (1989) divided effective teaching into four main characteristics: organization, enthusiasm, knowledge, and rapport. Although gender of professor, gender of subject, and whether the professor was a chair had little impact on the overall evaluations of instruction (only about one percent), organization accounted for 51% of teaching effectiveness, enthusiasm accounted for 72%, knowledge accounted for 72% and rapport accounted for 51% of overall teaching effectiveness.

A final line of thought is that females actually receive higher evaluations than do males, primarily because of the interaction of gender with other variables. Bennett (1982) found that female teachers are rated more highly on overall teaching effectiveness and favor more personal time and attention to students. Dukes and Victoria (1989) also supported this with their findings that females score higher than males in effectiveness (when neither are department chairs). Kierstead, et. al (1988) also stated that females are perceived as warmer than males. Also, when isolating the variables of confidence and decisiveness, females ranked high, but these same variables had no impact for males.

Physically attractive instructors overall rank higher than unattractive instructors, and attractive females who also demonstrate authoritarianism receive the highest evaluations (Buck and Tiene, 1989), possibly because attractiveness overrides perceptions of authoritarianism. Although, unattractive humanitarian females are evaluated equally positively as attractive authoritarian females (Buck and Tiene, 1989). Expressive instructors get very positive evaluations, with the exception of expressive males, who are evaluated the lowest (Basow, 1987). Interestingly, expressiveness seems to enhance masculinity ratings of females, but decreases masculinity ratings of males.
Socializing with students outside of class has had a positive impact on females' evaluations, but had no impact on males' evaluations (Basow, 1987). Females who return work promptly are also given high evaluations (Buck and Tiene, 1989). Finally, females who perform beyond expectations are rated higher than a male who performs equally well (Abramson, Goldberg and Greenberg, 1977).

How can the major differences in the research in this area be accounted for? A general problem with the majority of research in this area is that it tends to be strictly case study and survey data. Little research has been conducted in classroom settings. An additional problem is that the research has focused primarily on first impressions and the data may change substantially if gathered at a later point in an actual course, when a student-instructor relationship has developed.

Gender of the faculty advisor may also play a role in how students evaluate the advising they receive. Do discrepancies exist between the time, effort, and skill expended by female and male professors in academic advising and student evaluation of the two groups? Nadler and Nadler's (1992) exploration of this question indicated that while students with female faculty advisors report seeing their advisors more frequently, view them as more empathic, and are more likely to recommend them to their friends than are students with male advisors, they do not rate their advisors more favorably than do students with male advisors.

Exploration of similar issues was undertaken in a different kind of midwestern university, using results of the university's annual academic advising survey. No difference was found in female and male students' perceptions of female and male advisors' dependability, responsiveness, credibility, respectfulness, and overall usefulness to students, or in reported frequency of contact and desired frequency and preparation in advising sessions. Institutional context is offered as a possible explanation of results. This institution gave awards for excellence in advising (Meyers, Ferguson and Mack, 1993).

Sandler (1991) contends that both male and female students may also expect women faculty members to be more available for discussion of student's personal problems and to be more supportive listeners than male faculty members. Therefore women faculty members have a much higher workload of informal advising, especially with women students.

What action plans, if any, follow from these often contradictory research findings? Three general paradigms are suggested. The first advocates no action be taken as the sex of the professor is not an important variable in determining how students rate instruction. The second and third paradigms share the assumption that "sex of the instructor" plays a significant
role, whether it is "sex" as the primary variable or "sex" interrelated with other key variables. They are also grounded in the assumptions that teaching evaluations indicate teaching effectiveness or at least student perceptions and reactions to the classroom experience and that positive evaluations are necessary for promotion and tenure. Although both paradigms may share the goals of improving the classroom environment and increasing the opportunity for female faculty to earn/receive positive evaluation, they differ profoundly in the placement of responsibility for the achievement of these goals. The second paradigm identifies individual faculty members as the primary actors and the third paradigm establishes a campus and curriculum wide sharing of responsibility.

The first paradigm is in response to the research findings of Dukes and Victoria (1989). They found that Teaching Effectiveness was strongly correlated with Organization, 51 percent; Enthusiasm, 72 percent; Knowledge, 72 percent; and Rapport, 51 percent, and that "gender of professor," "gender subject," and "chair" explained only about one percent of the variation in Effectiveness (455). Their findings indicate that quality of teaching overwhelmed the other variables. They also contend that as student raters become more competent, their ratings of performance become more objective and gender bias diminishes. Their identification of "gender of professor" as explanation for less than one percent of variation in effectiveness is well within the estimates provided by Marsh (1980), who contends that 12 - 14 percent in the variance of ratings of instructors is explained by non-teaching effectiveness factors such as "characteristics of teacher."

If these findings are accepted one could argue that the "sex of professor" plays a very insignificant role in the rating of instructors by students and thus sexism is not present in the evaluation process. If sexism is not present than no action for remedy is required. However even Dukes and Victoria end their article questioning how their findings should be interpreted.

Does this finding mean that gender bias does not affect teaching evaluations? It does not appear to do so in a simple way; exploration of more complex relationships must wait for future research (455).

The second paradigm is based upon an acceptance that "sex of professor" does play a significant role in how the instructor is rated by students. This paradigm contends that individual faculty members must make the necessary behavioral adaptations in order to receive high ratings. Reviewing the literature on student evaluation of instructors provides some guidelines for behaviors female faculty may need to enact in order to improve their evaluations:
--smile a lot (Kierstead et al. 1988; Hall et al. 1982)
--be friendly with students (Kierstead et al. 1988)
--be as attractive as possible (Goebel and Cashen, 1979)
--utilize formal presentations (Bennett, 1982); and reduce class participation (Macke, Richardson and Cook, 1980)
--be very organized (Bennett, 1982; Goebel and Cashen, 1979)
--return work promptly (Buck and Tiene, 1989)
--socialize with students outside of class (Kierstead et al. 1988)
--be expressive (Basow and Silberg, 1987)
--demonstrate confidence and decisiveness (Bennett, 1982)
--display stereotypic female patterns of behavior (Basow and Silberg, 1987; Wheeless and Potori, 1989; Schein, 1975) and speech patterns (Papalewis, 1990)
--be non-dominant (Norton, 1977)
--be understanding and caring (Rubin, 1981)
--be an easy grader (Unger, 1979)

The above suggestions are summarized nicely by Kierstead et al. (1988).

Female instructors who want to obtain high student ratings must not only be highly competent with regard to factors directly related to teaching but also careful to act in accordance with traditional sex role expectations. Male and female instructors will earn equal SRI's for equal professional work only if the women also display stereotypical feminine behavior (344).

There are a number of weaknesses with this paradigm. The first being that the aforementioned suggested behaviors, if enacted, do not guarantee high evaluations. The second is a potential that engaging in these behaviors may reinforce sexism. This weakness follows from some research findings establishing that students rated same sex faculty members more highly (Ferber and Huber, 1975) and that female students rated female faculty members more highly (Ranking, 1981). If female students do provide female faculty with higher rankings than their male student counterparts are, might the female faculty end up pandering to the male student and masculine preferences by adjusting their behaviors for higher evaluations.

A third weakness has been labeled the "double bind." Although students may expect more caring and warm behavior from a woman faculty member, they may not evaluate this behavior, if demonstrated, positively because it is viewed as "too feminine"
(Zikmund, 1988). However, should she not demonstrate this expected feminine behavior and instead should engage in behavior comparable to that of her male colleagues she may be viewed as "too masculine" and once again evaluated negatively. This "double bind" reflects the ambivalence many students feel for female faculty members who are expected to demonstrate stereotypical feminine characteristics of warmth, friendliness, supportiveness and deference, yet as professionals are to be objective, authoritarian, and critical (Martin, 1984; Hall et. al, 1982; Minner, 1988).

A fourth weakness is that female faculty will be evaluated less positively than male counterparts no matter what behaviors they engage in because as women they are devalued. Whether it is that women are devalued or that men are idealized (Sandler, 1991) their behavior is seen as superior to and healthier than females (Harris, 1976) or that male teachers are favored (Minner, 1988; Kaschak, 1978) men receive more attention, eye contact and direct praise from students. Students accept and forgive demanding male teachers and are hyper-critical of the same behavior in their female teachers (Sandler, 1991).

The third paradigm advocates a confrontation of sexism not only within the classroom but throughout the curriculum and the academic culture. Sandler (1991) gives language to the importance and need for this institution-wide action.

The differential experience of women faculty in the classroom is not just a private problem for the individuals affected and to be solved by individual faculty members. Women faculty members alone, can at least respond to some of these problems but they cannot solve them. These problems are institutional and public. Extensive policy and institutional strategies are even more critical than the specific strategies for individual women faculty members which help them deal with student behaviors that create a chilly climate.

(Sandler, 1991, 8)

Sandler provides both general recommendations and specific recommendations for the identifying and challenging of sexism in schools. Enactment of these policies and practices will hopefully promote more awareness and reduction of sexism.

GENERAL RECOMMENDATIONS

- Incorporate climate issues into the mission of the institution.
- Recognize that change is the responsibility of everyone on campus, not only women, faculty, or administration. Recognize that change is an
ongoing process and that dealing with the issue once (or several times) will not "solve" the problem.

- Work actively to create an atmosphere where women’s issues are viewed as institutional issues and where issues of differential treatment in the classroom and elsewhere can be discussed.
- Educate all members of the academic community—from the trustees to students and staff—about differential treatment, the forms it takes, and the institution’s commitment to make the climate more equitable.
- Adopt a non-sexist language policy for institutional communications. A number of institutions, such as the University of New Hampshire, have done so.
- Insure that efforts to improve the classroom climate recognize the experiences and concerns of women of color.
- Insure that efforts to improve the classroom climate cover teaching faculty, teaching assistants and students.
- Use existing offices and structures to evaluate the classroom climate for women faculty members, such as faculty development programs, committees or commissions on the status of women, women’s studies coordinating committees, and graduate student organizations.
- Monitor these offices to insure that they are knowledgeable and incorporate climate issues into their ongoing activities.

(Sandler, 1991, 8 & 9)

In addition Sandler provides specific recommendations geared for administrators, female faculty members and male faculty members. Some of the guidelines that relate directly to student evaluation of faculty include the following:

--Acknowledge and Publicize the Issue in Order to make Classroom Behavior a Topic for Campus-wide Discussion.

--Develop materials to be disseminated to students and/or for faculty to handout and discuss; if desired, as to what constitutes appropriate behavior toward other students and faculty members.

--Gather data by surveying faculty and students about these issues and make the results a matter of public discussion.
--Develop Appropriate Policies.

--Appoint a university-wide committee to explore the problem and develop recommendations.
--Develop or evaluate existing policies covering student behaviors toward faculty members.
--Develop standards for behavior toward faculty members, male and female, in the student code of behavior.
--Evaluate policies and their implementation against faculty.
--When women faculty members are charged with bias against men, insure that the procedure includes determining whether the faculty member was truly favoring men or just giving all her students, including women an equal chance.

--Respond Swiftly and Publicly When Sexist Incidents Occur.

--Support Institutional Research on Climate Issues.

--Provide Training for Faculty and Administrators.

--Conduct workshops for promotion and tenure committees to insure that members are aware of devaluation and its impact on student evaluations.
--If student evaluations are formally conducted by the institution try to counteract devaluation by listing specific behaviors such as:
  --Does this teacher help students think about issues or does the teacher present all the information as a given?
  --Asking only for generic ratings such as "Is this a good teacher?" may obscure specific traits that are associated with good teaching.
--Conduct workshops focused on this issue at retreats, general faculty meetings, college-wide programs, lecture series and departmental meetings.
--Include information about and strategies for dealing with student behaviors in faculty development programs and in teaching programs for graduate and teaching assistants.

--Recommendations for Women Faculty Members.

--Don’t be modest about your accomplishments.
--Develop a handout, or discuss during the first class session, what constitutes appropriate behavior toward other students and toward the faculty member.
--Decide how you want students to address you.
--Be aware of your style of speaking and how it might affect others. Choose the style most useful for different parts of your teaching so that you can be appropriately assertive and/or collaborative.
--Be aware of self-effacing comments, especially at the beginning of the semester.
--When problems occur, talk to other women faculty members for clarification of what is happening and exchange ideas about strategies.
--When instances of disrespectful, disruptive or sexist behavior occur, recognize that you need to deal with them as soon as possible, if not immediately.
--Do not be afraid to tell students when their behavior is unacceptable.
--If one or more students frowns when you are speaking, confront it openly.
--When students interrupt, keep talking and continue making your point.
--Consider using humor to handle some issues.
--One way to deflect aggressive questioning is to direct the question to the class, rather than trying to restate your position.
--Respond when students tell sexual jokes, make sexual innuendos or sexist remarks.
--Some instances of offensive behavior can be handled nonverbally rather than verbally.

--Recommendations for Men Faculty Members.

--Be a role model for male students in terms of how to treat women equitably.
--Make a concerted effort to be equitable in your own classes and to do more than merely being neutral.
--Avoid the generic "he" or other words that do not connote equity.
--Do not assume that because you treat women fairly that others do the same.

(Sandler, 1991, 9-13)

Since this paradigm is demanding that confrontation of sexism occur in all areas of the academic community and curriculum it is vital that a variety of sources be used to gain better understanding of the dynamics of sexism and its various manifestations in the classroom. Sources such as Grossman and Grossman’s text Gender Issues in Education (1994), although geared at elementary and pre-college secondary education, provides some excellent suggestions for eliminating stereotypical perceptions and behaviors in their chapter entitled "Reducing Gender-Stereotypical Behavior."

Learning about and confronting sexism is strongly bound to issues of multiculturalism presently being addressed in education. Many of the policies and procedures suggested by Sandler can be adopted to other issues of multiculturalism. Likewise action
presently being taken to address other forms of prejudice and discrimination in higher education may be adapted and used to confront sexism.


Zikmund, B.B. (1988, September 1). The well-being of academic women is still being sabotaged by colleagues, by students, and by themselves. *Chronicle of higher education*, A44.