**Empirical Research on the Impact of Race & Gender**

**in the Evaluation of Teaching**

**A report compiled by Therese Huston, Ph.D.**

*Director of the Center for Excellence in Teaching & Learning*

*Seattle University*

**October 5, 2005**

**Introduction**

This report was prompted by a faculty member’s request to summarize the research on the impact of race and gender in the evaluation of university teaching.

The unfortunate news is that there is relatively little in the way of empirical, quantitative research on the intersection of race and gender as they pertain to course evaluations. Numerous studies have examined the possibility of a gender bias in student evaluations of teaching, but the possibility of a race bias is a relatively new area of inquiry, as evidenced by the fact that most of the relevant articles are published in 2005. The silver lining, of course, is that the findings are recent and are likely to reflect current institutional trends.

This document summarizes the key empirical findings. Further information about most of the studies listed here can be obtained through CETL.

**Findings by sub-topic**

INTERACTION OF RACE & GENDER

* Hamermesh & Parker (2005) found lower final course evaluation ratings for female minority faculty members, but not for male minority instructors. In their study, women instructors received significantly lower course evaluations than male instructors (nearly ½ standard deviation lower) and faculty of color received lower course evaluations than white faculty. There was also an interaction of race and gender such that female faculty of color received particularly low course evaluations.
* In the same study, there was a different pattern of interaction between gender and language. Hamermesh & Parker (2005) found that male instructors who were non-native English speakers were judged more harshly on course evaluations than female instructors who were non-native English speakers. There was also a main effect of language such that non-native English speakers had significantly lower course evaluations overall, but male faculty had lower scores than female faculty within this category.

* Two recent studies have shown that faculty of color receive either the most favorable or the least favorable ratings when compared to Anglo faculty (Anderson & Smith, 2005; Smith & Anderson, 2005). Both studies only compared Latino/a and Anglo faculty and did not include African-American instructors. In one study of a politically charged social science course, there was an interaction between ethnicity, gender, and teaching style (Anderson & Smith, 2005). Female Latina professors who had a lenient teaching style were viewed as more warm than female Anglo professors with a lenient teaching style. Likewise, female Latina professors who had a strict teaching style were viewed as more strict than their Anglo female peers with a comparable style. In summary, perceptions of warmth and strictness were magnified for female faculty of color. (These summaries are based on the abstract for these articles – the full articles have been requested through Inter-Library Loan.)

RACE

* Hamermesh & Parker (2005) – as mentioned above, there was a main effect of race such that faculty of color received lower course evaluations than their white peers.
* Researchers have looked at other measures besides course evaluations to gauge racial bias in the classroom. For example, using a survey separate from the final course evaluation, Rubin (1998) found that students rate Asian-American instructors as less credible and intelligible than white instructors. In a series of semi-structured interviews, Hendrix (1998) found that students in a predominantly White university did not believe that a professor’s race influenced their perceptions of that instructor’s credibility, yet the students simultaneously described a different set of criteria for evaluating the credibility of their Black instructors (relative to the criteria applied to their White instructors) for courses on certain topics. Students’ comments revealed that Black instructors had more credibility when they taught courses that had an ethnic or racial focus, and students reported that they would more readily question and challenge the credibility of Black instructors for courses that lacked an ethnic / racial component to them.
* There is also literature that demonstrates a feedback bias for minorities, but these studies have been done outside of instructional settings. One of the interesting dimensions of this research is that the direction of the bias depends on whether the feedback is perceived as going directly to a person of color or whether the feedback is perceived as going to some other third party. When Whites rate the performance of a person of color with the understanding that their judgments would be communicated to a third party for the purposes of evaluation, Whites consistently rate performance negatively (e.g. Henderson-King & Nisbett, 1997; Lambert, Cronen, Chasteen & Lickel, 1996). When Whites rate the performance of a person of color for the purposes of giving feedback to that person directly, however, the Black person being evaluated actually receives significantly more favorable marks than the White person being evaluated (Harber, 1998). In light of this research, it may be helpful to consider whether students view their evaluations for administrative, summative evaluation purposes or for the instructor to improve his or her teaching.

GENDER

* The issue of gender bias in teaching evaluations is a huge area of controversy in the literature. Some studies, such as Hamermesh & Parker (2005; see also Etaugh & Riley, 1983; Helgeson, 1994), found that female faculty members receive lower course evaluation ratings, whereas other studies report that male faculty members receive lower course evaluation scores (Costa, Terracciano, & McCrae, 2001; Kierstead, D’Agostino, & Dill, 1988). However, in the more comprehensive and controlled studies that look for gender bias in student evaluations across multiple disciplines, researchers have not found an overall bias favoring male or female instructors (Feldman, 1993; Centra & Gaubatz, 2000). After reading a variety of studies on the subject, the following conclusions appear to receive the most support:
  + There are same-gender preferences, such that a) female students tend to give higher ratings to female instructors, and b) male students tend to give higher ratings to male instructors (Feldman, 1993; Centra & Gaubatz, 2000). Such same-gender preferences have been attributed to theories of gender differences in “ways of knowing,” such that both women instructors and women students prefer teaching methods that emphasize “connected classrooms” (Belenky, Clinchy, Goldberger, & Tarule, 1986) that value “understanding and acceptance over evaluation and collaboration over debate” (Centra & Gaubatz, p. 31, 2000).
  + In science classrooms, which are traditionally male dominated, both males and female students give higher ratings to female instructors (Centra & Gaubatz, 2000). Male instructors were more likely teach using lecture methods in these disciplines whereas female instructors used more discussion in the sciences which could account for the gender difference in the course evaluations (Centra & Gaubatz, 2000).
  + The question still remains as to why there could still be such dramatic differences in the research findings, such that one steady line of research has female faculty at a disadvantage and another line has male faculty at a disadvantage. Not surprisingly, the underlying issue might concern the types of questions that are asked on the student evaluations.
    - If some of the questions on the course evaluation focus on warmth, such as the instructor’s ability to understand or respond to students’ needs (qualities usually associated with feminine expressiveness) then female instructors have higher scores (Costa, Terracciano, & McCrae, 2001; Kierstead et al., 1988).
    - If however, some of the questions ask students to rate instructors’ enthusiasm or eagerness in the classroom (qualities usually associated with masculine expressiveness), then male instructors have higher scores (Helgeson, 1994; Mulac & Lundell, 1982).

The course evaluation forms differ for each College and School at Seattle University, so it may be valuable to take a look at the questions used for each school. If one of the questions suggest that masculine qualities are valued in teaching a particular discipline, then such a question could create what is known as a “halo effect,” and bias how students answer other questions on the course evaluation (\*\*\*\*). The halo effect and this gender bias are somewhat complicated because they appear to affect questions that require a high degree of inference (such as “Overall, how was the quality of this instructor’s teaching?”) but they do not appear to affect objective questions that ask about specific competencies and more concrete behaviors (such as “Did the instructor answer questions effectively?”).

* Hamermesh & Parker (2005) also found an interaction between gender and course level (i.e. whether the course being evaluated was an upper or lower division course). Female faculty teaching upper-division courses received course evaluations that were about average for the sample, but female faculty teaching lower-division courses received course evaluations that were far below average.

**Summary of Key Findings**

The bulleted list below summarizes the key findings on race and gender bias in student evaluations of teaching more succinctly.

* There are very few studies that examine race and gender in the evaluation of teaching. One study indicates that relative to their colleagues, female faculty of color receive lower course evaluations, and two other studies indicate that students perceive faculty of color in positive or negative extremes.
* Students are able to simultaneously deny any impact of race on judgments of teaching and articulate the different criteria they use to judge Black and White faculty.
* In non-instructional settings, Whites are biased to evaluate the performance of a person of color more positively if they think they are providing direct feedback, but biased to evaluate performance more negatively if they think the evaluation is for a third-person’s evaluation.
* Many researchers have examined the issue of gender bias in teaching evaluations without controlling for race. Although individual studies can be found that support a gender bias favoring either male or female faculty, the reasonable conclusion based on the most comprehensive studies is that there is no general gender bias in course evaluations.
* There does tend to be a same-gender preference, with female students preferring female professors and male students preferring male professors.

**Limitations of the Research**

There are conceivably many limitations to the research on race and gender bias in students’ evaluations of teaching, particularly given the small collection of empirical studies. A few of these limitations are listed below.

* The issue of cognitive bias (either conscious or unconscious bias) has not been applied or addressed in these studies.
* Although one study shows that students confer more credibility to faculty of color who teach courses related to race and ethnicity, it is not clear which academic disciplines, if any, are subject to greater racial bias. Given that there is a gender bias within the sciences, there may be comparable racial biases in specific fields.
* Are white students more likely to perceive that they are reporting to what is typically a predominantly white administration when they evaluate instructors of a different race (relative to when they evaluate instructors of the same race)?

**Selected References\***

Anderson, K. J., & Smith, G. (2005). Students preconceptions of professors: Benefits and barriers according to ethnicity and gender. *Hispanic Journal of Behavioral Sciences, 27*, (2), 184-201.

Belenky, M. F., Clinchy, B. M., Goldberger, N. R., & Tarule, J. M. (1986). *Women’s Ways of Knowing*. New York: Basic Books.

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

Etaugh, C., & Riley, S. (1983). Evaluating competence of women and men: Effects of marital and parental status and occupational sex-typing. *Sex Roles*, *9,* 943-952.

Feldman, K. A. (1993). College students’ views of male and female faculty college teachers: Part II – Evidence from students’ evaluations of their classroom teachers. *Research in Higher Education*, *34*, 151-211.

Hamermesh, D. S. & Parker, A. M. (Forthcoming, 2005). Beauty in the classroom: Instructors’ pulchritude and putative pedagogical productivity. *Economics of Education Review*.

Harber, K. D. (1998). Feedback to minorities: Evidence of a positive bias. *Journal of Personality and Social Psychology*, *74*, (3), 622-628.

Henderson-King, E. I., & Nisbett, R. E. (1997). Anti-Black prejudice as a function of exposure to the negative behavior of a single Black person. *Journal of Personality and Social Psychology*, *71*, 654-664.

Hendrix, K. G. (1998). Student perceptions of the influence of race on professor credibility. *Journal of Black Studies*, *28*, 738-764.

Lambert, A. J., Cronen, S., Chasteen, A. L., & Lickel, B. (1996). Private vs. public expressions of prejudice. *Journal of Experimental Social Psychology*, *32*, 437-459.

Rubin, D. L. (1998). Help! My professor (or doctor or boss) doesn’t talk Englisgh. In J. N. Martin, T. K. Nakayama, L. A. Flores (Eds.), *Readings in Cultural Contexts* (pp. 149 – 160). Mountain View, CA: Mayfield Publishing Company.

Smith, G., & Anderson, K. J. (2005). Students’ ratings of professors: The teaching style contingency for Latino/a professors. *Journal of Latinos and Education, 4*, (2), 115-136.

*\*Full list of references available upon request*