Examining the Relation between Race and Student Evaluations of Faculty Members: A Literature Review

DANA A. WILLIAMS

The assertion that scholarship is limited on the relation between ethnicity and student evaluations of faculty members is perhaps an understatement. While there is a wealth of scholarship on the relation between gender and student evaluations of faculty members, little has been published on how ethnicity (of both faculty members and students) informs students' rating of teaching effectiveness. Throughout research into issues specific to minority faculty members there are passing references to the sometimes unfair use of student evaluations to determine faculty tenure, promotion, and merit pay; but these references do not and cannot serve as pertinent scholarship on how a faculty member's ethnic background creates biases that reveal themselves in those evaluations.

In one of the few essays that address the relation between ethnicity and evaluation, Heidi J. Nast explores, among other things, "student resistances to multicultural teaching and faculty diversity [and] the risks that derive from problematic institutional deployment of student evaluations as a means of judging multicultural curricular and faculty success" (103). Nast's essay is especially revealing in the following articulations. First, "students use evaluations to register anger and disapproval at having to negotiate topics and issues in a scholarly way which conflict with heretofore learned social values and assumptions." Second, the likelihood of negative evaluations increases when faculty members "curricularly address issues of homophobia, racism, classism, misogyny or heterosexism"

The author is associate professor of English at Howard University.

(104). Third, the problem of racially motivated negative evaluations does not end with racist students; it is an institutional problem, as colleagues and administrators are often unequipped and unwilling to recognize and to deal with racism. Nast speculates that the fear of negative evaluations and the lack of institutional support to combat racism lead faculty members to assume defensive postures with regard to evaluations. Those who fear, she argues, are more likely to be lenient with students, to administer evaluations strategically, and on occasion to plead with students before an evaluation. All these actions, especially the last two, compromise the accuracy of the evaluations. The most effective strategies for decreasing the potential for racially specific negative faculty evaluation, she argues, must begin at the institutional level, so that sensitivity to evaluative biases can be properly addressed.[1] One strategy that faculty members of color might find useful is to structure supplemental evaluations that are connected to course content and that help students develop analytic skills that make them aware of their biases. An evaluation might have a prefatory section that, first, describes how the evaluation will be used and how it is valued institutionally and that, second, warns students about "the potential *emotional* impact of education" (109).[2]

Nast's argument is not based on quantitative research, nor is its framework theoretical. It presents qualitative comments from anonymous sources. These comments are revealing but not particularly useful in theorizing ways to improve the evaluative process. Similarly, Roxanna Harlow's essay offers revealing qualitative comments that speak to her examination of "how and to what degree race shapes professors' perceptions and experiences in the undergraduate classroom" (348). While the article does not speak directly to evaluations, it does identify specific ways that race compels black faculty members, first, to negotiate a devalued racial status and, subsequently, to engage in "extensive emotion management"; both strategies affect students' evaluation of faculty performance and competence.[3]

Unlike Nast's and Harlow's essays, Kristin J. Anderson and Gabriel Smith's offers an important balance among a qualitative, a quantitative, and a theoretical approach to the relation between ethnicity and students' evaluations of faculty members. In their attempt to address the lack of studies that examine students' perceptions of ethnic minority faculty members and the interaction of such perceptions with "course content and the impact of these perceptions on student evaluation of instruction," the authors examine "the interactions of gender, ethnicity, and teaching style on students' perceptions of professors teaching a politically charged social science course" (185).[4] Their findings are based on an experimental

comparison of student ratings of Latino and Anglo professors.[5] The strength of their study lies in the use of multiple variables to hypothesize the efficacy of student evaluations and in the quantitative results that the authors present. For instance, they investigated the ways the combination of course content, gender, and ethnicity affect students' perceptions of a course and a professor; and they conducted three five-way between-subjects analyses of variance (ANOVAs) to measure their findings. The most revealing findings are: "[R]atings of professor warmth and availability for Latino professors appear to be contingent on their teaching style, whereas the rating of Anglo professors' warmth is less contingent upon teaching style" (196); "Among women professors with strict teaching styles, Anglo women were rated as more capable than Latinas with the same teaching style. . . . Lenient Latinos were viewed as more capable than strict Latinos, whereas Anglo professors' capability was not contingent on teaching style" (197); Anglo men, more so than any other group of students, judge women professors and Latino professors as having a political agenda (198); and "Latino students rated Anglo women professors as more capable than Latino professors, and they saw Anglo women professors as more capable than did African American students. Also, Latino students saw Anglo women as more capable than Anglo men" (199).

Like Anderson and Smith, Jeannette M. Ludwig and John A. Meacham conducted an experimental study, creating fictitious instructors and syllabi to facilitate their analysis.[6] The purpose of their study was "to assess the impact of instructor gender and race on student evaluations of teaching effectiveness, particularly when courses include controversial content" (27). Interestingly, their findings did not support their main hypothesis, that students are apt to give low evaluations of teaching effectiveness for women and for minority instructors when the course content is controversial. But the findings did suggest that an instructor's characteristics (as revealed in a short biographical sketch) influence students' evaluations of which course material is considered controversial (36). In short, controversial course material alone does not appear to be detrimental to positive evaluations of teacher effectiveness. But Ludwig and Meacham point out that because their study was experimental, they could not factor in two key variables: the fact that a syllabus alone cannot include specific content that arises during the semester and the fact that only during the semester can students really feel the challenge to their beliefs of controversial course content (35). These findings and this literature review show how important it is for scholars to conduct additional research on this issue.

Yet even though the extent of scholarship on the relation between race and student evaluations of faculty members is limited, the little available

to us is both revealing and informative. It supports the following observa-
tions and recommendations:

Student evaluations are used primarily in four ways: to provide feedback to fac-
ulty members, to offer guidance to students in their course selections, to
assess teaching effectiveness for purposes of promotion and tenure, and to
justify or deny increases in pay based on merit.
Evaluations are in need of improvement. Those that claim to measure teach-
ing effectiveness must have some grasp of teaching and learning theories.
Measurement of evaluations must be quantitatively informed and sufficiently
sophisticated to be useful. Variables such as time of day, teaching style, in-
structor ethnicity and gender and sexual orientation (if the instructor has
made that orientation explicit to students), nature of the course (requirement
or elective), and course content should be factored in. Numbers alone reveal
little. An evaluation often tells more about a student's opinion of a professor
than about the professor's teaching effectiveness.
Department chairs should offer faculty members of color institutional support
that acknowledges and works to combat racism and discrimination by creating
an environment in which grievances can be articulated openly and addressed
seriously. Chairs should also ensure that student evaluations of faculty mem-
bers are constructed in such a manner that they reveal racially specific or
racially motivated bias. As a supplement, faculty members of color should con-
struct their own evaluations that are course-specific and designed to encour-
age students to reflect on the possibility of students' racially specific bias.
In addition to studying how race intersects with students' evaluations of faculty
members, researchers should construct normative questions that reveal ra-
cially specific bias and that departments and faculty members can use to limit
the impact of discriminatory student evaluations on the promotion and tenure
of faculty members of color.

Increasing the accuracy of student evaluations of faculty members of
color increases the integrity of the academic experience not only for faculty
members of color but for all faculty members and for students as well.

NOTES

1. Her argument moves along these lines: "institutions need to confront more
honestly the limitations of student evaluations" (108); department chairs must read
student evaluations carefully, paying careful attention to subtle clues to bias; "insti-
tutions can proactively prepare and educate students about meaningful assessment
procedures" and uses; and institutions should "conduct systematic, systemic studies
of racism, homophobia, and sexism" to understand the pressures on faculty members
who carry out multicultural or diversity projects (109).
2. Nast hypothesizes: "In some courses, controversial social issues such as rac-
ism, sexism, religious and ethnic conflict, and homophobia are discussed, which
may cause emotional discomfort because they carry with them difficult questions."

Sample questions include but are not limited to the following: "Were controversial topics raised in this class . . . ?" "Did you feel upset about the fact that the issues were raised?" "Do you feel that the controversial topic caused you to consider the course in a negative light?" (110). Note that all the comments of obvious racial bias that Nast cites are written, not revealed through multiple-choice selections.

3. Harlow conducted fifty-eight in-depth interviews with twenty-nine white and twenty-nine African American faculty members at a predominately white (91%) state university. The faculty members were then matched by similar gender, rank, and department or area of study and questioned about "their anxiety on the first day of class, their teaching style, the level of students' energy in their classes, students' opinions of the respondents, how they would like to be viewed by students, and other subjects" (350). Interestingly, Harlow's findings suggest that although black faculty members were clear about racially specific stereotypes and felt that racial inequality was still an issue in society today, almost half were "reluctant to claim that their race mattered to students, or that race influenced their classroom experience in any negative way" (351). They seemed to acknowledge race on the macro level but were less inclined to acknowledge it personally. This notwithstanding, Harlow's research revealed that while only 7% of white professors felt that students questioned their qualifications, 76% of black professors felt that students questioned their qualifications. White professors seldom considered that race might influence students' professional evaluation of them. White women, when asked how race or gender affected students' views of them, tended to focus on gender alone (354). For further commentary on the relation between race and students' perceptions of professor credibility, see Hendrix, who effectively engages this issue as it relates specifically to African American faculty members.

4. See Coren on how course content influences students' perceptions of instructors' biases. Coren is more interested in student evaluation of an instructor's racism than in the influence of race on students' evaluation of professors. But Coren analyzes "fundamental attribution error" and the "halo effect" as two basic psychological errors that weaken the credibility of standardized student course evaluation forms that are developed specifically to assess an instructor's racism, sexism, and sensitivity to multicultural issues. Students cannot separate the message from the messenger when dealing with course content that is politically unpopular, and "negative opinions about an unpopular instructor who does not teach well are likely to be generalized." The result is the attribution of racist and sexist views to that instructor (201). Anderson and Smith characterize teaching styles as strict or lenient. They note that "[t]eaching style varied according to the language each professor used on the syllabus. The styles were modeled after the coding scheme used by Eagly, Makhijani, and Klonsky (1992) in their meta-analysis of studies on gender bias and leadership" (190).

5. They "examined the impact of professor and student characteristics on students' perceptions of professors. The characteristics of the professor examined . . . were teaching style . . . , professor gender, and professor ethnicity (Latino and Anglo). Student characteristics examined were student gender and ethnicity (Latino, African American, and Anglo). A syllabus was constructed for a social science course called *Race, Gender, and Inequality*, and versions of the syllabus varied according to teaching style, gender, and ethnicity. [They] asked undergraduate respondents to read the syllabus and rate the course and the instructor on dimensions such as warmth, availability, knowledge of the topic, preparedness and capability, and lack of objectivity and

political bias" (188). By using the experimental method, the writer-researchers were able to control the course—the content, lectures, requirements, and time of day; the gender and ethnicity of the professor; and the professor's teaching style.

6. Students were given booklets that presented fictitious instructors who were proposing new courses. The cover page of the booklet suggested that the professor was proposing a new course for undergraduate students and that a committee needed students' assistance to evaluate the course before it was taught. A brief description of the professor was offered, and students were asked to examine the biographical sketch closely. The descriptions identified the professor by name, race, rank, place of education, specializations, and publications. The package also included a sample syllabus, which was characterized either as "high-controversy" (presenting topics that concerned racism and sexism) or as "low-controversy" (presenting topics that concerned more general social problems).

WORKS CITED

Anderson, Kristin J., and Gabriel Smith. "Students' Preconceptions of Professors: Benefits and Barriers according to Ethnicity and Gender." *Hispanic Journal of Behavioral Sciences* 27 (2005): 184–201.

Coren, Stanley. "Student Evaluations of an Instructor's Racism and Sexism: Truth or Expedience?" *Ethics and Behavior* 8.3 (1998): 201–13.

Eagly, A. H., M. G. Makhijani, and B. G. Klonsky. "Gender and the Evaluation of Leaders: A Meta-analysis." *Psychological Bulletin* 111 (1992): 3–22.

Harlow, Roxanna. "'Race Doesn't Matter, but . . .': The Effect of Race on Professors' Experiences and Emotion Management in the Undergraduate College Classroom." *Social Psychology Quarterly* 66 (2003): 348–63.

Hendrix, Katherine Grace. "Student Perceptions of the Influence of Race on Professor Credibility." *Journal of Black Studies* 28 (1998): 738–63.

Ludwig, Jeanette M., and John A. Meacham. "Teaching Controversial Courses: Student Evaluations of Instructors and Content." *Educational Research Quarterly* 21.1 (1997): 27–38.

Nast, Heidi J. "'Sex,' 'Race,' and Multiculturalism: Critical Consumption and the Politics of Course Evaluations." *Journal of Geography in Higher Education* 23.1 (1999): 102–15.