



**AN EXAMINATION OF RACE AND SEX-BASED BIASES IN PROFESSIONAL
EMPLOYEE PERFORMANCE EVALUATIONS**

Journal:	<i>Academy of Management Journal</i>
Manuscript ID:	draft
Manuscript Type:	Original Manuscript
Keyword:	Gender and Diversity in Organizations < Topic Areas, Social cognition theory < Theoretical Perspectives, Performance evaluation & management < Human Resource Management and Industrial Relations < Topic Areas, Health Care Administration < Topic Areas, Diversity management < Human Resource Management and Industrial Relations < Topic Areas



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

AN EXAMINATION OF RACE AND SEX-BASED BIASES IN PROFESSIONAL EMPLOYEE PERFORMANCE EVALUATIONS

ABSTRACT

This paper examines whether the relationship between objective and subjective performance is stronger for professional employees who belong to certain demographic groups. We hypothesize and find that customer ratings of employee performance quality are influenced by the combination of observable employee physical characteristics (e.g., race, sex) and the degree to which employees display behaviors that benefit customers. For white or male employees, the display of customer centered behaviors increased customers' ratings of employee performance quality. However, for non-white or female employees, customer-centered behaviors exerted a negligible, or in some cases negative, influence on perceptions of employee performance quality.

1
2
3 As the U.S. economy matures, the manufacturing industry is shrinking and the service
4 industry continues to expand. Changes in the labor force mirror these macroeconomic shifts.
5
6
7
8 The fastest growing segment of the labor force is professional workers¹ such as doctors, lawyers,
9 and professors. Moreover, according to the Bureau of Labor Statistics, professional occupations
10 will continue to be the fastest growing segment of the U.S. workforce for the foreseeable future
11 (Hecker, 2005). Professionals already comprise over 20% of the workforce and are expected to
12 increase by another 5.3 million by 2014 (Hecker, 2005).
13
14
15
16
17
18

19
20 The increase in professional jobs is also changing the nature of performance evaluation.
21
22 In non-professional jobs, organizations typically rely on supervisors to assess employee
23 performance because they are viewed as the most qualified and knowledgeable source for
24 evaluating the employee. However, managers may have difficulty evaluating the skills of
25 professional employees because their jobs involve complex skills that include pleasing
26 customers. Pleasing customers is important for a service-oriented organization because it is
27 assumed that positive service experiences should yield greater profitability. For this reason
28 customer ratings have become an increasingly popular source of performance feedback for
29 professionals (Hagan, Konopaske, Bernardin, & Tyler, 2006). For example between 1998 to
30 2003, the percentage of service organizations giving their employees feedback from customers
31 increased from 10 to 30 percent (Heathfield, 2006).
32
33
34
35
36
37
38
39
40
41
42
43
44

45
46 The willingness of organizations to place greater weight on customer satisfaction as a
47 way of assessing employee performance can be found in organizations ranging from airlines
48
49

50
51 ¹ We use the Bureau of Labor Statistics definition of professional to mean any person engaged in work
52 predominantly intellectual and varied in character as opposed to routine mental, manual, mechanical, or physical
53 work; involving the consistent exercise of discretion and judgment in its performance; of such a character that the
54 output produced or the result accomplished cannot be standardized in relation to a given period of time; requiring
55 knowledge of an advanced type in a field of science or learning customarily acquired by a prolonged course of
56 specialized intellectual instruction and study in an institution of higher learning or a hospital, as distinguished from a
57 general academic education or from an apprenticeship or from training in the performance of routine mental,
58 manual, or physical processes.
59
60

1
2
3 (Jenkins, 1992) to universities (Davis & Davis, 1999). Some business leaders (Bracken, Church,
4 & Timmreck, 2001) and researchers (Salam, Cox, & Sims, 1997) have applauded the use of
5
6 customer ratings because they believe that aggregated evaluations are highly reliable measures of
7
8 employee performance quality. Other observers have questioned their use, pointing to research
9
10 showing that biases based on racial, gender, or other social categorization processes are
11
12 unavoidable when using subjective evaluations of performance, particularly when they come
13
14 from naïve and inexperienced raters (Pulakos, White, Oppler, & Borman, 1989; Wilkinson &
15
16 Fontaine, 2002; Woehr & Roch, 1996).
17
18
19
20
21

22 The attempt to detect racial and gender biases in performance evaluation in general has
23
24 proven to be a theoretical and methodological challenge because of one major problem:
25
26 employees' rated performance scores can be viewed as capturing employees' true performance
27
28 scores. As Rotundo and Sackett (1999) write, "There is no definitive way of determining
29
30 whether the criterion used in a validity study is biased. Thus, there is no current method of
31
32 establishing whether there is bias in performance ratings (Rotundo & Sackett, 1999: 816)." The
33
34 results of the studies themselves have also been inconclusive. For example, in their
35
36 comprehensive review of research on personnel selection Landy and colleagues concluded that:
37
38 "No one seems to have found direct gender or race effects, at least in large-scale field studies
39
40 (Landy, Shankster, & Kohler, 1994: 283)." Reviewing this same literature, Latham and Wexley
41
42 (1994) reached a similar conclusion, stating that: "[W]hen the appraiser uses behaviorally based
43
44 appraisal scales, ratee characteristics, such as age, race, and sex, have a negligible effect on the
45
46 resulting performance appraisal" (p. 152). Indeed, researchers have found it difficult to reach
47
48 definitive conclusions about how demographic characteristics might influence the performance
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 evaluation process (e.g., Barrett & Morris, 1993; Ellis, Ilgen, & Hollenbeck, 2006; Sackett &
4
5 DuBois, 1991; Vecchio, 2002).
6
7

8 Despite the failure to demonstrate that racial and gender biases systematically influence
9
10 performance evaluations, there are *theoretical* reasons to expect that in jobs where the
11
12 assessment of performance has unavoidable subjective elements and where consensus on what
13
14 constitutes excellent performance is difficult to reach – as is the case in most professional
15
16 occupations – stereotypes and other unconscious attitudes and beliefs may sometimes be used by
17
18 evaluators to “fill in the blanks.” As a result, we hypothesize that a significant amount of
19
20 variance in performance evaluation may be attributable to these cognitive processes. The
21
22 theoretical bases for this hypothesis are drawn from several converging streams of research
23
24 showing that while overt expressions of prejudice and discrimination in U.S. society are now
25
26 relatively rare, negative beliefs, stereotypes, and attitudes towards women and ethnic minorities
27
28 persist in more subtle and often unconscious forms (e.g., Gaertner & Dovidio, 1986; Kinder &
29
30 Sears, 1981). For example, studies show that many people hold a male leader prototype (Banaji
31
32 & Greenwald, 1995; Eagly, 1987) and a generalized belief that men are superior and deserve to
33
34 control and receive more resources than women (Jost & Banaji, 1994). Similarly, people who
35
36 report low levels of overt prejudice against blacks have nevertheless been shown to recommend
37
38 hiring a white compared to a black applicant when the two candidate’s qualifications were
39
40 ambiguous (Dovidio & Gaertner, 2000), indicating the operation of what some scholars (e.g.,
41
42 Gaertner & Dovidio, 1986) refer to as “aversive racism.” What these studies suggest is that
43
44 potential biases in the performance evaluation process may be motivated not by what people say
45
46 they believe, but by stereotypic beliefs or attitudes that are not expressed or that may be
47
48 unknown even to the raters.
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 If stereotypes and unconscious biases influence social cognition, as many studies over the
4 years have shown, then the theoretical argument that the demographic characteristics of
5 employees may influence how other people perceive, interpret, and evaluate their performance
6 seems tenable. The purpose of our study was to test this possibility by examining whether
7 customer service-oriented behaviors exhibited by highly skilled, professional employees and
8 customer ratings of those same employees' service quality vary, either in strength or direction,
9 depending on the employee's race and sex. We argue that if this relationship does in fact vary as
10 function of the employee's race and sex it could indicate a possible source of customer
11 evaluation bias arising from negative stereotypes and beliefs they hold about certain
12 demographic groups.
13
14
15
16
17
18
19
20
21
22
23
24
25
26

27 In our study, the professional service employees were physicians working for a large
28 health maintenance organization; the customers were their patients. By shifting the focus of
29 subjective performance evaluations to customer rather than supervisor ratings, our study
30 overcomes a major methodological limitation of past research on performance evaluation
31 because the customer-raters had no knowledge of what types of physician behaviors that
32 organization valued and rewarded. Hence, customers' evaluations should have been based solely
33 on what they observed and experienced in their encounters with physicians. Moreover, our
34 design is an improvement over earlier studies of performance evaluation which measured
35 employee performance with a single subjective rating. In those types of studies, there is no
36 standard by which to determine if the rating is biased (Rotundo & Sackett, 1999). Importantly,
37 both the objective (service-oriented behaviors) and subjective (customer ratings of professional
38 quality) performance measures were used by the focal organization to determine employee
39 salaries which meant that any systematic biases in these measures could translate directly into
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 differences in pay and the generation of gender and racial inequality in other career outcomes.
4
5 The following sections present theoretical arguments for why we expect the same customer-
6
7 oriented behaviors performed by physicians to be differentially related to patients' evaluations of
8
9 physician performance as a function of the latter's race and sex.
10
11

12 **THEORETICAL BACKGROUND AND HYPOTHESES**

13
14
15 The classification of other people into categories based on ascribed characteristics like
16
17 race or sex appears to be a universal human response to social information (Devine, 1989; Fiske
18
19 & Neuberg, 1990; Hewstone, Hantzi, & Johnston, 1991). Social categorization, like other
20
21 cognitive heuristics, is useful and adaptive for simplifying information processing and bringing
22
23 order to the world. But categorization can also lead to biased or inaccurate judgments about those
24
25 whom we assign to a given social category and, importantly, categorization often occurs with
26
27 little conscious reflection (Hewstone et al., 1991). It has been shown that easily detectable
28
29 physical characteristics like race and sex are among the most reliably observed bases for
30
31 spontaneous social categorization (Stangor, Lynch, Duan, & Glass, 1992) and so it seems likely
32
33 that judgments of employees by customer-raters can be at least partly influenced by these
34
35 automatic categorizations.
36
37
38
39

40
41 One way the operation of social categorization biases in a performance evaluation
42
43 process might be revealed is when *identical* behaviors exhibited by employees show a different
44
45 relationship to customers' subjective evaluations of service quality as a function of the rated
46
47 employees' race or sex. Although we do not expect the correlation between objective behaviors
48
49 displayed by professionals and subjective ratings of service quality to approach unity, it seems
50
51 reasonable to expect a significant, positive correlation, assuming the behaviors do in fact have a
52
53 positive impact on customer welfare. The more crucial question though is whether this
54
55
56
57
58
59
60

1
2
3 relationship is stronger or weaker as a function of the race and sex of the professional being
4 evaluated. Two theories of how group-based stereotypes influence social judgment provide an
5 identical answer to this question. We review each of them beginning with status characteristics
6 theory.
7
8
9
10
11

12 *Status characteristics theory*

15 Status characteristics theory asserts that individuals have preconceived, status-based
16 notions about their own and each other's competence (Berger, Conner, & Fisek, 1974; Berger,
17 Fiske, Norman, & Zelditch, 1977). In the United States, whites and men are considered by most
18 people to be members of a high status social group relative to women and ethnic minorities (see
19 Ridgeway, 1991 for a review). This perception is not surprising given that members of these
20 groups control most of the resources and occupy most of the powerful positions in U.S. society
21 (Sidanius & Pratto, 1999). One of the benefits of belonging to a high status social group is that
22 observers are more likely to make favorable inferences about one's competence, normality, and
23 legitimacy (Aquino & Bommer, 2003; Giannopoulos, Conway, & Mendelson, 2005; Sidanius &
24 Pratto, 1999). In contrast, members of devalued groups are subject to negative stereotypes and
25 attributions concerning their work-related competencies (Fernandez, 1981; O'Leary & Ickovics,
26 1992). One implication of these differences in inferential judgment is that behaviors performed
27 by a high status group member may be interpreted more favorably by observers than the same
28 behavior performed by someone from a low status group. Evidence for this possibility is
29 provided by studies showing that women in leadership roles are disliked more than men in
30 similar roles (Eagly, Makhijani, & Klonsky, 1992), that ethnic minorities and women are
31 rewarded less than whites and males for exhibiting advice-giving or ingratiation
32 behaviors (Westphal & Stern, 2007), and that whites who exhibit high levels of organizational
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 citizenship behavior are less likely than blacks who exhibit the same level of citizenship to report
4 being victimized by the harmful actions of co-workers (Aquino & Bommer, 2003).
5
6
7

8 We propose that ascribed characteristics like sex and race can exert the strongest
9 influence on the performance evaluation process when subjective judgment is required. For
10 example, one study found that when job applicant resume quality was ambiguous, applicants
11 with African-American-sounding names (e.g. Aisha, Rasheed) were much less likely to be called
12 for a job interview than applicants with white sounding names (e.g. Kristin, Brad) (Bertrand &
13 Mullainathan, 2004). Likewise, when evaluators of orchestral positions could see the sex of the
14 applicant they were more likely to select men. When the applicant's sex could not be observed,
15 the number of women hired significantly increased (Goldin & Rouse, 2000). People tend to have
16 lower performance expectations for members of low-status groups and so it is also likely that
17 racial minorities and women may have to achieve higher levels of performance than whites and
18 men to be judged as having the same level of underlying ability (Biernat & Kobrynowicz, 1997;
19 Yarkin, Town, & Wallston, 1982). Compounding the adverse effects of status-based expectations
20 is the fact that people appear to resist information that contradicts them (Biernat & Kobrynowicz,
21 1997; Foschi, 2000). For example, individuals only acknowledge the competence of women if
22 there is explicit evidence (Shackelford, Wood, & Worchel, 1996). When no explicit, objective
23 evidence exists, people fit observable cues with their pre-conceived expectations. Based on status
24 characteristics theory, we therefore expect customers to evaluate the behavior of professionals
25 more favorably when that professional belongs to a high rather than low status demographic
26 group. The exemplar-based model of social judgment (Smith & Zárate, 1992) supports this
27 prediction.
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54

55 *The exemplar model of social judgment*
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55

According to the exemplar model, within any culture, particular person attributes come to be perceived as expected or default values. In Western society, white racial identity and male gender are treated as cultural expectations. Evidence for this “white male norm” hypothesis comes from experiments showing that membership in non-normative groups receives greater attention than membership in normative groups because of its incongruence (Smith & Zárate, 1990; Stroessner, 1996). Arguably, this normative expectation is exacerbated in the professional occupational category, which has historically been dominated by white males. For example, the 2006 Diversity Report by the Association of American Medical Colleges reports that only 24.5% of practicing physicians are women and 12.1% are non-whites.² If the exemplar model’s assumptions are valid, then customers of professionals should pay more scrupulous attention to the behavior of professionals who belong to groups that violate their normative expectation of the professional social category. As a result, customers might be more critical when forming subjective judgments of performance if the professional belongs to non-prototypical category. Some evidence for this possibility is provided by Westphal and Stern’s (2007) finding that women and ethnic minorities on corporate boards were more likely to be “punished” than their white male counterparts for carefully monitoring management initiatives and attempting to control management decision-making. In their study, being punished meant that they would have less chance of receiving a board appointment

46
47
48
49
50
51
52
53
54
55

Combining the idea that the actions of persons who do not fit within the social category elicit heightened attentiveness with the propositions of status characteristics theory, we theorize that any perceived deficiency in professional’s behavior toward a customer elicits a more negative evaluative response, and may perhaps be more easily recalled by customers, when the

56
57
58
59
60

² However, the Association of American Medical Colleges also reports that these numbers are changing dramatically as 44% of American medical school graduates in 2004 were women and 34% were non-whites.

1
2
3 professional is female rather than male or an ethnic minority rather than a white. By extension,
4
5 any positive behaviors directed toward the customer by the professional will elicit amore
6
7 favorable evaluation when exhibited by a white or male than when exhibited by an ethnic
8
9 minority or female. As a result of these differences in evaluative judgments, we expect the
10
11 relationship between objective professional performance and subjectively measured performance
12
13 to differ depending on a professional's sex and race. Together, status characteristics theory and
14
15 exemplar-based models of social judgment lead to the same conclusion, stated formally in the
16
17 following hypotheses:
18
19

20
21
22 *Hypothesis 1: The positive relationship between professional employee's objective*
23
24 *performance, indicated by behavior directed toward the customer to improve customer*
25
26 *welfare, and subjective evaluations of performance quality, indicated by customer*
27
28 *ratings, is weaker for female than for male professionals.*

29
30
31 *Hypothesis 2: The positive relationship between a professional employee's objective*
32
33 *performance, as indicated by behaviors directed toward the customer to improve*
34
35 *customer welfare, and subjective evaluations of performance quality, indicated by*
36
37 *customer ratings, is weaker for ethnic minority than for white professionals.*
38
39

40 41 42 **METHOD**

43 44 **Sample**

45
46 Our sample was drawn from a large health maintenance organization, hereafter referred
47
48 to as Prohealth (a pseudonym). Prohealth was founded in 1947 and is non-profit and consumer-
49
50 governed. Today, it provides coverage and healthcare for about 350,000 people in the Pacific
51
52 Northwest region of the United States. It directly employs approximately 800 healthcare
53
54 providers (both general practitioners and specialists) to care for its members. Our sample
55
56
57
58
59
60

1
2
3 consisted of all 113 primary-care physicians (i.e., family practitioners) at Prohealth Within our
4
5 sample, 38.4% were women, 11.5% were ethnic minorities, and all had a medical degree.
6
7

8 **Measures**

9
10 Prohealth routinely collected patient ratings to assess the quality of physician
11
12 performance as well as objective behavioral indicators of physician performance that were
13
14 assumed to have a direct, positive impact on patient health and well-being. This feature of our
15
16 data represents a methodological improvement over many earlier studies of performance
17
18 evaluation bias because in most of these studies the only measure of employee performance was
19
20 a single subjective rating. In those types of studies, there is no standard by which to determine if
21
22 the rating is biased (Rotundo & Sackett, 1999). The dependent variable in our study was
23
24 physician performance quality as rated by patients. The independent variables were physician
25
26 demographics (race and gender) and three types of patient-centered behaviors.
27
28
29
30

31
32 **Customer rating of professional performance quality.** Like other health plans,
33
34 Prohealth measures customer satisfaction with the Consumer Assessment of Health Plans
35
36 Survey. A survey was mailed to a percentage of each physician's patients, following doctor
37
38 visits, selecting the patients from a stratified sample based on the patients' visit utilization rates
39
40 (to avoid a bias toward those patients with frequent appointments). Patients completed and
41
42 returned a total of 12,091 surveys for a response rate of 52%, so that each physician was rated by
43
44 an average of 107 patients. Patients rated each of the three items targeting their physician on a 5-
45
46 point Likert-scale (1=very poor; 5=excellent) "How would you rate..." (1) the attention the
47
48 provider paid to you, (2) this provider's thoroughness and competence, and (3) your opportunity
49
50 to ask questions of this provider. The three items were highly correlated (average correlation is
51
52 .93) so the organization combined them to create a composite patient satisfaction variable. The
53
54
55
56
57
58
59
60

1
2
3 organization did not provide us with access to raw patient-level surveys. Instead, they provided
4
5 us with data indicating what percentage of each physician's patients rated the physician as
6
7 "excellent." Thus, the range on this measure for each physician was from 0 to 100 percent. This
8
9 measure was collected in the same quarter as all other variables.
10
11

12 **Professional race.** Prohealth identified each physician's race, and we coded whites "0"
13
14 and ethnic minorities as "1." Of the 113 physicians in the sample, 10 were Asian or Pacific
15
16 Islanders, two were black, and one was Native American. The percentage of ethnic minority
17
18 physicians in our sample is consistent with the national average of 12.1%.
19
20
21

22 **Professional sex.** We also obtained this variable from Prohealth records and coded
23
24 males "0" and females "1." Forty-three of the physicians were female Thirty -eight percent of
25
26 physicians in our sample were women, which is slightly higher than the national average of
27
28 24.5%.
29
30
31

32 **Customer-centered behavior.** Patients' welfare has always been the ultimate driving
33
34 force in medicine; however, physicians have historically pursued patients' welfare via a
35
36 predominantly physician-centered perspective (Daghio et al., 2003; Laine & Davidoff, 1996).
37
38 The patient's point of view and concerns have historically been considered irrelevant or a
39
40 hindrance to understanding the core issue—the patient's disease. This physician-centered
41
42 approach to medical care involves focusing on doctors' convenience, rather than focusing on
43
44 patients' convenience. For example, patients historically have been expected to wait patiently
45
46 for doctors, and to never keep the doctor waiting (Mishler & Waxler, 1963).
47
48
49

50 However, with the growing prevalence of health maintenance organizations (HMOs) and
51
52 the increasing corporatization of medicine (Feinglass & Salmon, 1990), patients are increasingly
53
54 being viewed by organizational administrators and physicians as customers. Therefore,
55
56
57
58
59
60

1
2
3 physicians are increasingly being rewarded for engaging in *customer-centered*, rather than
4
5 physician-centered behaviors (Laine & Davidoff, 1996; Stewart et al., 2000). We used the
6
7 customer-centered behaviors identified by Prohealth as our indicator of objective physician
8
9 performance.
10
11

12 Prohealth measures customer-centered behaviors along three dimensions. The first is
13
14 physician productivity, which is the number of health procedures performed and issues discussed
15
16 in a given time period. The second is the physician's accessibility to customers measured by the
17
18 number of secure emails that doctors send to customers. The third is the physician's level of
19
20 quality measured by the standardized prescription rates of particular medications for customers
21
22 that possess precise disease criteria. All three dimensions reflect behaviors that benefit
23
24 customers by reducing the amount of time and money customers spend receiving medical care.
25
26 For all metrics, physicians are shown how they compare to both the organizational goal and the
27
28 organizational average. Physician compensation is tied to each of these customer-centered
29
30 behaviors. Physicians who exceed the 40th percentile are given a bonus, while those below the
31
32 40th percentile are not given a bonus.
33
34
35
36
37
38

39 All three customer-centered behaviors are presumed to benefit customers. More
40
41 productive physicians are able to treat more customer problems per visit, which saves customers
42
43 time and additional trips to the doctor. Higher physician accessibility saves customers
44
45 unnecessary visits to the doctor by allowing customers to simply email any medical questions to
46
47 their physician. Higher physician quality benefits customer by preventing costly and deadly
48
49 health events such as strokes and heart attacks.
50
51

52
53 *Physician Productivity.* The average number of patients seen, medical issues discussed,
54
55 and medical procedures performed by each doctor in a standardized 8-hour day was recorded by
56
57
58
59
60

1
2
3 the organization's scheduling software. Prohealth physicians have a great deal of control over
4 the amount of work that they do in a day as they can control the intensity of each visit (e.g., the
5 number of procedures performed and patient health issues addressed per visit). The number of
6 patients that physicians see each day is controlled by organizational administrators.
7
8
9
10
11

12 This variable was the composite of average face-to-face visits and phone visits adjusted
13 by the average intensity of each visit. Intensity was measured by Relative Value Units (RVUs),
14 which are coded by physicians at the end of each visit according to national coding guidelines.
15 RVUs capture the amount of time involved, the required physical and mental effort, the required
16 judgment and technical skill, and the psychological stress entailed (Hsaio, Braun, Becker, &
17 Thomas, 1988; Hsaio, Braun, Dunn, & Becker, 1988). According to quarterly audits by
18 administrators, Prohealth physicians accurately record RVU's in 90 percent of patient visits.
19 Coding errors resulting from physicians coding too many or too few RVU's are normally and
20 equally distributed. The raw measure of productivity was standardized based on the full-time
21 status of the physician and then multiplied by each physician's average visit intensity to obtain
22 the quarterly *average RVU-adjusted patient encounters per day*.
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

38 *Physician Accessibility.* The average daily number of emails that physicians sent to
39 patients for the quarter was used to measure another type of patient-centered behavior. Physician
40 emails benefit patients because patients value the ability to easily contact their physician.
41 Indeed, a Harris poll has shown that 90% of Americans who are online want the ability to e-mail
42 their physicians, and 37% are even willing to pay for it (Taylor, 2002). Prohealth patients can
43 send unlimited emails to their physician at no cost and Prohealth physicians are expected to reply
44 to each email within 24 hours. Patients primarily contact their physician via email to ask basic
45 health-related questions, to request prescription refills and to schedule follow up appointments.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Prohealth administrators encourage Prohealth physicians to contact patients via email because
4 they think that an email might save patients a trip to the doctor, thereby saving patients' time and
5 money. In general, physicians do not think that email improves the quality of patient care, but
6 rather that email increases convenience for patients (Kleiner, Akers, Burke, & Werner, 2002).
7
8
9
10
11

12 The Prohealth computer server automatically recorded the number of emails that each
13 medical professional sent to his or her patients. Medical professionals had a great deal of control
14 over how many emails they sent for two reasons: (a) they could choose whether or not to inform
15 their patients of their email address, and (b) they could choose whether or not to personally
16 respond to their patients' emails. We calculated the number of emails sent per day, taking into
17 account the number of full working days that physicians were in clinic during the time-period of
18 this study. To enhance the normality of the variable, we used an inverse transformation and then
19 reflected these values such that higher values represented greater use (Tabachnick & Fidell,
20 2003).
21
22
23
24
25
26
27
28
29
30
31
32
33

34 *Physician Quality.* Every Prohealth primary care physician is responsible for a panel of
35 member-patients. Of the thousands of possible treatments, prescriptions, and procedures that
36 physicians can perform to benefit patients, one of the most important is each physician's
37 prescription rate of statins and angiotensin-converting-enzyme (ACE) inhibitors to patients with
38 cardiovascular disease. Treatment of cardiovascular events such as strokes, clots, and heart
39 attacks is the biggest healthcare cost for patients in the U.S. (Willerson & Cohn, 2000), and
40 these drugs prevent cardiovascular events over patients' lifetimes (Gerstein et al., 2000).
41
42
43
44
45
46
47
48
49

50 Nationally, only 50% of all cardiovascular disease patients that should be treated with
51 statins and ACE inhibitors are currently taking such medication (Dubois et al., 2002). According
52 to Prohealth guidelines, all patients with cardiovascular disease should be regularly taking ACE
53
54
55
56
57
58
59
60

1
2
3 inhibitors and some form of a statin. ACE inhibitors lower blood pressure, and statins lower
4 cholesterol. These drugs significantly lower the immediate risk of a cardiovascular event (e.g.,
5 stroke, heart attack) for all individuals, regardless of sex or previous history of cardiovascular
6 disease (LaRosa, He, & Vupputuri, 1999; Yusuf et al., 2000). To promote a higher prescription
7 rate, Prohealth administrators send emails to physicians to remind doctors to prescribe such
8 treatment. While prescription of these medications benefits patients by helping patients avoid
9 death and reduce healthcare expenses, individual physicians do not generally see their utility on a
10 daily basis. Indeed, physicians often forget to prescribe these medications (Isles, 2002).

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
This variable is the composite of the percent of cardiovascular disease patients 18 years
and older who were dispensed the equivalent of a 90-day supply for ACE inhibitors and statins at
any time within the quarterly reporting period. The component variables approached normality
and were added together. The resulting variable was each physician's overall *prescription rate*
of statins and ACE inhibitors for cardiovascular disease patients. The average prescription rate
at Prohealth is 50%, similar to the national average.

Control Variables

We controlled for several variables that were not of direct interest for testing our
hypotheses but that could be theoretically related to the dependent variable and might provide
plausible alternative explanations for our findings.

Average Practice Busyness. Patients who have to wait long periods of time to see their
physician may be less satisfied than those who can see their physician immediately so we
controlled for the busyness of each physician's practice. This variable is automatically calculated
by the Prohealth computer server at the end of each business day and represents the average time
until each physician's third available appointment for the quarter.

1
2
3 **Physician full-time status.** Patients might be more satisfied with physicians who work
4 more hours because these patients might be more likely to schedule a convenient appointment
5 time with these physicians. Thus, we controlled for the number of hours a physician worked.
6 We collected this variable from the archival records of Prohealth. Physicians ranged from
7 working 30 to 100 percent of a full-time position.
8
9

10 **Number of patients in panel.** Each full-time physician is assigned to care for a group
11 (i.e. panel) of roughly 2,000 patients. Physicians with large panels may have high levels of
12 patient demand, which could influence patient satisfaction. We controlled for the total number
13 of patients in each physician's panel divided by the full-time status of the physician over the
14 most recent 12-month period. This data was obtained from Prohealth's human resource
15 department.
16
17

18 **Average patient age.** Physicians whose patients are older may be more or less
19 demanding than physicians whose patients are younger. We used the average patient age for
20 each physician's panel for the most recent quarter.
21
22

23 **Average chronic sickness of panel.** Patient panels that are more chronically ill may be
24 more or less likely to be satisfied than patient panels that are less chronically ill so we controlled
25 from the chronic sickness of each physician's panel. This variable was calculated by Prohealth
26 for each physician and captures the chronic illness of each physician's panel (e.g., diabetes,
27 cardiovascular disease) so that higher values represent more chronically ill patient populations.
28 This variable excludes acute episodic illness (e.g., broken bones, injuries due to accidents).
29
30

31 **Physician Age and Tenure.** Physicians who are older or who have been employed by
32 Prohealth for more years may have more loyal, satisfied patients than younger or recently hired
33 physicians.
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

RESULTS

Table 1 reports the means, standard deviations, and correlation coefficients between the dependent, independent, and control variables.

Insert Table 1 here

We used hierarchical moderated regression models to examine the hypothesized interaction effects (Aiken & West, 1991). We centered all variables involved in the interaction terms to minimize multicollinearity between the interaction terms and their individual components (Aiken & West, 1991). We entered all of the control and independent variables in Model 1, in Model 2 we entered all hypothesized two-way interactions involving sex and the three dimensions of objective performance, and in Model 3 we entered all two-way interactions involving race and the objective performance measures. The full model includes all interaction effects. Table 2 presents the results of this analysis.

Insert Table 2 here

Hypothesis 1 states that the positive relationship between professional performance measured by objective, customer-centered behaviors and subjective evaluations of professional performance will be weaker for female than for male professionals. The two-way sex X objective performance interactions as a set explained a significant amount of incremental variance in the dependent variable ($R^2 = .08, p < .01$) providing preliminary support for this hypothesis. Inspection of the individual regression weights showed that the accessibility X sex

1
2
3 and quality X sex interactions were significant ($p < .05$). We probed the pattern of the
4
5 interaction by examining the simple slope of the objective performance measures for male and
6
7 female professionals (Aiken & West, 1991). The results of this analysis are shown graphically in
8
9 Figures 1 and 2.
10
11

12
13 -----
14
15 Insert Figures 1 and 2 here
16
17 -----
18
19

20
21
22 Both figures show a stronger positive relationship between customer-centered behaviors
23 and performance ratings for men than for women. We calculated the significance of the simple
24 slopes for interactions (Aiken & West, 1991). For male professionals the coefficient of the
25 simple slopes of quality on subjective performance was significant and positive ($\underline{b}=.03, p < .01$)
26 and accessibility behaviors on subjective performance was significant and marginally positive
27 ($\underline{b}=.13, p < .10$). However, for female professionals the coefficient of the simple slope of quality
28 behaviors on subjective evaluations of performance is negligible ($\underline{b} = -.01, n.s.$) and for
29 accessibility behaviors is marginally negative ($\underline{b} = -.13, p < .10$). These results support
30 Hypothesis 1 for two of the objective-subjective performance relationships.
31
32
33
34
35
36
37
38
39
40
41
42
43

44 Hypothesis 2 states that the positive relationship between professional performance of
45 customer-centered behaviors and subjective evaluations of professional performance is weaker
46 for non-whites as compared to whites. The two-way race X objective performance interactions
47 as a set explained a significant amount of incremental variance in the dependent variable (R^2
48 $=.06, p < .05$) providing preliminary support for our hypothesis. Inspection of the individual
49 regression weights showed that the productivity X race and quality X race interactions were
50
51
52
53
54
55
56
57
58
59
60

1
2
3 significant ($p < .05$). The forms of the interactions are shown graphically in Figures 3 and 4.
4
5

6
7 -----
8
9 Insert Figures 3 and 4 here
10
11 -----
12

13 Simple slope analysis reveals that for white professionals the coefficient of the simple
14 slope of quality on subjective performance is significant and positive ($b = .02$, $p < .01$) and
15 productivity on subjective performance is marginally significant and positive ($b = .02$, $p < .10$).
16
17 However, for non-white professionals the simple slope of quality behaviors on subjective
18 evaluations of performance is marginally significant and negative ($b = -.02$, $p < .10$) and
19 significant and negative for productivity behaviors ($b = -.03$, $p < .01$). These results support
20 Hypothesis 2 for two of the objective-subjective performance relationships.
21
22
23
24
25
26
27
28

29 30 **DISCUSSION**

31
32 We set out to explore whether professional's demographic characteristics moderate the
33 relationship between objectively measured customer-centered behaviors and subjective ratings of
34 performance quality. We found that customer-centered behaviors were positively related to
35 subjective evaluations of performance quality, but *only* for whites and males. Indeed, the direct
36 relationships between customer-centered behaviors and subjective evaluations were not
37 significant indicating that these relationships were fully qualified by professional employee
38 demographics. We also found that certain types of customer-centered behaviors were *negatively*
39 related to customer evaluations of performance for women and non-whites. This was an even
40 stronger result than we anticipated because our prediction was that the relationship between
41 customer-centered behaviors and subjective performance would be weaker, but still positive, for
42 women and non-whites compared to men and whites.
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 We suggest that the observed pattern of relationships between objective, customer-
4 centered behaviors and subjective performance ratings by customers may indicate the operation
5 of a stereotypic bias that leads customers to ignore or discount positive behaviors when they are
6 performed by professionals who do not fit within the exemplar category of physician. It is also
7 possible, based on status characteristics theory, that beliefs about the competence and legitimacy
8 of professionals who belong to a high status group (i.e., whites, males) can lead customers to
9 elevate subjective performance ratings in response to customer-centered behaviors relative to
10 when these same behaviors are performed by a professional from a lower status and socially
11 devalued group.
12
13
14
15
16
17
18
19
20
21
22
23

24 One way that social categorization processes might influence customer ratings of
25 professional service quality is by leading them to draw erroneous inferences of underlying traits
26 and abilities based on an employee's group membership (Biernat & Kobrynowicz, 1997). For
27 example, Biernat and Kobrynowicz (1997) showed that when judging a fictional employee,
28 participants required blacks and women to achieve higher levels of performance than whites and
29 men in order to make the same inferences of underlying ability. Consequently, even if female or
30 ethnic minority professionals behave the same way as their male and white counterparts,
31 customers may not have the same level of confidence in their ability and competence because the
32 standards for demonstrating the underlying traits are higher. This stereotype-driven bias is
33 particularly likely to occur in cases where judgments are highly subjective and performance
34 cannot be objectively assessed (Glick, Zion, & Neslon, 1988).
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

50 Traditionally, organizations have relied on supervisors to assess employee performance
51 because they were viewed as the most qualified and knowledgeable source for evaluating the
52 employee. More recently, many organizations have supplemented the standard supervisory
53
54
55
56
57
58
59
60

1
2
3 appraisals with feedback from other sources like customers (e.g., Miles & Snow, 1995). Our
4
5 results suggest that if customer evaluations become widely and uncritically used to determine
6
7 pay and promotion opportunities, the job outcomes of women and ethnic minorities could be
8
9 adversely impacted.
10

11
12 Our method of testing for possible bias in performance evaluations was a significant
13
14 improvement over past studies. First, we used an objective performance standard so that we
15
16 could compare subjective ratings to this standard and therefore determine whether the later
17
18 ratings might be influenced by race and sex. Second, our subjective ratings were based on
19
20 aggregated judgments from a large number of customers rather than relying on the judgments of
21
22 a single supervisor. This is important because the large number of raters provides a highly
23
24 reliable subjective performance rating for each professional. Finally, we controlled for several
25
26 variables that could provide alternative explanations for our results, such as the baseline
27
28 busyness level of each professional's practice, and each professional's tenure with the
29
30 organization.
31
32
33
34
35

36
37 Given these methodological strengths of our study, it is unsettling to find that even for
38
39 highly skilled employees who possess similar levels of human capital and occupational prestige,
40
41 positive, customer-centered behaviors exhibited by those who belong to low status demographic
42
43 groups may not translate into increased customer satisfaction ratings relative to their prototypic,
44
45 high status counterparts. At Prohealth, women and non-white physicians can in fact be
46
47 economically harmed by subjective performance evaluations because physicians who fail to
48
49 achieve the target level of performance quality (i.e. the 40th percentile) do not receive a salary
50
51 bonus. The practical implications of our results become more apparent when we when examine
52
53 the effect sizes in our sample. The interactions between sex, race and objective performance
54
55
56
57
58
59
60

1
2
3 explained 14.5% of the variance in subjective evaluations of physician performance. Cohen
4
5 (1988) provides ballpark descriptors of effect sizes based on R-squared values—"large" ($R^2 =$
6
7
8 .25), "medium" ($R^2 = .09$), and "small" ($R^2 = .01$). Therefore, the observed effect size of the
9
10 interactive effects of race, sex, and objective performance on subjective evaluations in our
11
12 sample is between medium and large.
13
14

15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

There is a potential alternative explanation for our findings that we must reckon with before drawing firm conclusions about what we have shown. Our predictions implicitly assume that customer-centered behaviors influence customers' subjective ratings, although we are cautious to avoid making any causal statements because our design is cross-sectional. It is possible, though, that it is subjective ratings of performance that influence professionals' customer-centered behaviors. The potential for incorrect conclusions based on the cross-sectional design is minimized in our study because we tested interaction rather than main effects (Bowen & Wiersema, 1999). Nevertheless, the reverse-causal argument could be made and so we wanted to test this possibility statistically even though we recognize that a stringent test requires a longitudinal design.

We tested an alternative model implying reverse causation by creating two interaction terms (i.e. sex by subjective performance and race by subjective performance) and regressing each of our three objective performance measures on these variables along with the main effects and controls. When we analyzed our data this way, none of the six interactions involving the subjective ratings and physician race and sex were significant ($p > .10$). This does not rule out the possibility of reverse causality, but it does show that the model we specified explains our data better than an alternative model with different interaction terms.

1
2
3 In many ways, physicians are an ideal sample for testing our research question of whether
4 objective service behaviors are related to subjective customer evaluations of service quality.
5
6 Physicians provide a service that can literally impact their customer's very survival. Thus,
7
8 patients should be highly motivated to attend to and scrutinize the behavior of their physicians.
9
10 The work physicians do also involves many subtle, rich, and complex features that are not easily
11
12 amenable to quantification. Physicians not only dispense medical knowledge, but may also
13
14 provide socio-emotional and psychological support during times of illness and distress. And as
15
16 anyone who has had a thorough medical examination knows, the level of physical contact they
17
18 have with a patient's body is generally surpassed only in the most intimate of social
19
20 relationships. For these reasons, patients' assessments of physician performance are likely to be
21
22 influenced by cognitive and emotional processes that are intuitive, automatic, and perhaps
23
24 largely unconscious. This does not mean that no aspects of physician performance can be
25
26 assessed objectively or that patients do not deliberate about their physician's performance. On
27
28 the contrary, the focal organization identified several objective, behavioral criteria of
29
30 performance that were assumed to directly benefit patients. If these behavioral criteria do indeed
31
32 contribute to patient welfare, and if patients recognize this, then it is logical to expect that
33
34 physicians who exhibit more of these behaviors should be evaluated more favorably by their
35
36 patients than those who exhibit less of them. However, this was not the case for all physicians in
37
38 the organization we studied.
39
40
41
42
43
44
45
46
47

48 We believe that the results we found among physicians can generalize to other highly
49
50 skilled, professionals because the theoretical bases for our hypotheses are not specific to a
51
52 health-care context. We would expect our findings to be most strongly replicated in jobs
53
54 historically held by white males like the upper echelons of organizations (e.g. Westphal and
55
56
57
58
59
60

1
2
3 Stern, 2007) or in other prestigious professions (e.g. lawyers, professors). In these fields, what
4
5 constitutes exceptional performance is open to multiple interpretations and so the evaluation of
6
7 performance would seem highly vulnerable to stereotyping and unconscious biases. Of course,
8
9 whether our findings are upheld in such samples is a question for future research.
10
11

12 **Conclusions**

13
14
15 Over the last 40 years, women and ethnic minorities have entered occupations and
16
17 attained positions of leadership in business and government that at one time were almost entirely
18
19 the province of white men (Stroh, Brett, & Reilly, 1992). To some observers these trends show
20
21 that U.S. society has become more egalitarian, fairer, and less discriminatory since the Civil
22
23 Rights Movement of the 1960s (Gross & Hook, 1978; Lynch, 1991). Yet women and ethnic
24
25 minorities are still not well represented among the ranks of highly paid managers and
26
27 professionals in U.S. corporations and in prestigious occupations like law and medicine (e.g.,
28
29 Baldi & McBrier, 1997; Eagly & Karau, 2002; Wilson, Sakura-Lemessy, & West, 1999) and
30
31 their earnings lag behind those of white men within the same occupational category (Dreher &
32
33 Cox, 1996; Petersen & Morgan, 1995). Some writers have suggested that women and ethnic
34
35 minorities should simply try harder (Wenneras & Wold, 1997; Yarkin et al., 1982). Yet the
36
37 findings of our study suggest that it is an open question whether working longer hours, choosing
38
39 demanding work assignments, or going above and beyond the minimal expectations of their jobs
40
41 will necessarily help women and minorities achieve parity with their white male counterparts in
42
43 the absence of other interventions that directly attempt to counteract the operation of stereotypic
44
45 biases and unconscious beliefs that lead others' to discount or ignore these efforts.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- 1
2
3
4
5
6
7 Aiken, L. S. & West, S. G. 1991. *Multiple regression: Testing and interpreting interactions*.
8 Newbury Park, CA: Sage Publications.
9
- 10 Aquino, K. & Bommer, W. H. 2003. Preferential mistreatment: How victim status moderates the
11 relationship between organizational citizenship behavior and workplace victimization.
12 *Organization Science*, 4: 274-285.
13
14
- 15 Baldi, S. & McBrier, D. B. 1997. Do the Determinants of Promotion Differ for Blacks and
16 Whites? Evidence from the U.S. Labor Market. *Work and Occupations*, 24: 478-497.
17
18
- 19 Banaji, M. R. & Greenwald, A. G. 1995. Implicit gender stereotyping in judgments of fame.
20 *Journal of Personality and Social Psychology*, 68: 181-198.
21
22
- 23 Barrett, G. V. & Morris, S. B. 1993. The American Psychological Association's amicus curiae
24 brief in Price Waterhouse v. Hopkins: The values of science versus the values of the law.
25 *Law and Human Behavior*, 17: 201-215.
26
- 27 Berger, J., Conner, T. L., & Fiske, M. H. 1974. *Expectation States Theory: A Theoretical*
28 *Research Program*. Cambridge, MA: Winthrop.
29
- 30 Berger, J., Fiske, M. H., Norman, R. Z., & Zelditch, M., Jr. 1977. *Status Characteristics and*
31 *Social Interaction: An Expectation States Approach* New York: Elsevier.
32
33
- 34 Bertrand, M. & Mullainathan, S. 2004. Are Emily and Greg More Employable than Lakisha
35 and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic*
36 *Review*, 94(4): 991-1013.
37
38
- 39 Biernat, M. & Kobrynowicz, D. 1997. Gender and race-based standards of competence: Lower
40 minimum standards but higher ability standards for devalued groups. *Journal of*
41 *Personality and Social Psychology*, 72: 544-557.
42
43
- 44 Bowen, H. P. & Wiersema, M. F. 1999. Matching method to paradigm in strategy research:
45 Limitations of cross-sectional analysis and some methodological alternatives. *Strategic*
46 *Management Journal*, 20(7): 625-636.
47
48
- 49 Bracken, D. W., Church, A. H., & Timmreck, C. W. 2001. *Handbook of Multisource Feedback:*
50 *The Comprehensive Resource for Designing and Implementing MSF Processes*. New
51 York, NY: Jossey-Bass.
52
- 53 Cohen, J. 1988. *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ:
54 Lawrence Erlbaum.
55
56
57
58
59
60

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- Daghio, M., Ciardullo, A., Cadioli, T., Delvecchio, C., Menna, A., Voci, C., Guidetti, P., Magrini, N., & Liberati, A. 2003. GPs' satisfaction with the doctor-patient encounter: Findings from a community-based survey. *Family Practice*, 20(3): 283-288.
- Davis, W. E. & Davis, D. R. 1999. The University Presidency: Do Evaluations Make a Difference? *Journal of Personnel Evaluation in Education*, 13(2): 119-140.
- Devine, P. G. 1989. Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56: 5-18.
- Dovidio, J. F. & Gaertner, S. L. 2000. Aversive racism and selection decisions: 1989 and 1999. *Psychological Science*, 11(4): 315-319.
- Dreher, G. F. & Cox, T. H. 1996. Race, gender, and opportunity: A study of compensation attainment and the establishment of mentoring relationships. *Journal of Applied Psychology*, 81: 297-308.
- Dubois, R. W., Alexander, C. M., Wade, S., Mosso, A., Markson, L., Lu, J. D., Nag, S., & Berger, M. L. 2002. Growth in use of lipid lowering therapies: Are we targeting the right patients? *American Journal of Managed Care*, 8(10): 81-86.
- Eagly, A. H. 1987. *Sex differences in social behavior: A social-role interpretation*. Hillsdale, NJ: Erlbaum.
- Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. 1992. Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin* 111(1): 3-22.
- Eagly, A. H. & Karau, S. J. 2002. Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109: 573-598.
- Ellis, A. P. J., Ilgen, D. R., & Hollenbeck, J. R. 2006. The Effects of Team Leader Race on Performance Evaluations. *Small Group Research* 37(3): 295 -322.
- Feinglass, J. & Salmon, J. W. 1990. Corporatization of medicine: the use of medical management information systems to increase the clinical productivity of physicians. *Journal of Health Services*, 20(2): 233-252.
- Fernandez, J. P. 1981. *Racism and sexism in corporate life*. Lexington, MA: Heath.
- Fiske, S. T. & Neuberg, S. L. 1990. A continuum model of impression formation: From category based to individuating processes as a function of information, motivation, and attention. In M. P. Zanna (Ed.), *Advances in experimental social psychology*, Vol. 23: 1-74. San Diego, CA: Academic.
- Foschi, M. 2000. Double standards for competence: Theory and research. *Annual Review of Sociology*, 26: 21-42.

- 1
2
3 Gaertner, S. L. & Dovidio, J. F. 1986. The aversive form of racism. In J. F. Dovidio & S. L.
4 Gaertner (Eds.), *Prejudice, Discrimination, and Racism*. Orlando, FL: Academic Press.
5
6
7 Gerstein, H. C., Yusuf, S., Mann, J. F. E., Hoogwerf, B., Zinman, B., Held, C., Fisher, M.,
8 Wolffenbittel, B., Bosch, J., Richardson, L., Pogue, J., & Halle, J. P. 2000. Effects of
9 ramipril on cardiovascular and microvascular outcomes in people with diabetes mellitus:
10 Results of the HOPE study and MICRO-HOPE substudy. *Lancet*, 355(9199): 205-212.
11
12
13 Giannopoulos, G., Conway, M., & Mendelson, M. 2005. The gender of status: The laypersons'
14 perception of status is gender-typed. *Sex Roles*, 53: 795-806.
15
16
17 Glick, P., Zion, C., & Neslon, C. 1988. What mediates sex discrimination in hiring decisions?
18 *Journal of Personality and Social Psychology*, 55: 178-186.
19
20
21 Goldin, C. & Rouse, C. 2000. Orchestrating Impartiality: The Impact of 'Blind' Auditions on
22 Female Musicians. *American Economic Review*, 90(4): 715-741.
23
24
25 Gross, B. R. & Hook, S. 1978. *Discrimination in Reverse: Is Turn-About Fair Play?* New
26 York, NY: New York University Press.
27
28 Hagan, C. M., Konopaske, R., Bernardin, H. J., & Tyler, C. L. 2006. Predicting assessment
29 center performance with 360-degree, top-down, and customer-based competency
30 assessments *Human Resource Management*, 45(3): 357-390.
31
32 Heathfield, S. M.; 360 Degree Feedback: The Good, the Bad, and the Ugly
33 http://humanresources.about.com/od/360feedback/a/360feedback_2.htm.
34
35
36 Hecker, D. E. 2005. Occupational employment projections to 2014. *Monthly Labor Review*,
37 128(11): 70-101.
38
39
40 Hewstone, M., Hantzi, M., & Johnston, L. 1991. Social categorization and person memory: The
41 pervasiveness of race as an organizing principle. *European Journal of Social*
42 *Psychology*, 21: 517-528.
43
44
45 Hsaio, W. C., Braun, P., Becker, E. R., & Thomas, S. R. 1988. The resource-based relative value
46 system. *Journal of the American Medical Association*, 258: 799-802.
47
48
49 Hsaio, W. C., Braun, R., Dunn, D., & Becker, E. R. 1988. Resource-based relative values: An
50 overview. *Journal of the American Medical Association*, 260: 2347-2353.
51
52
53 Isles, C. G. 2002. Patients with acute coronary syndrome should start a statin while still in
54 hospital *Heart*, 88: 5-6.
55
56
57
58
59
60 Jenkins, K. J. 1992. Service Quality in the Skies. *Business Quarterly*, 57(2): 13-18.

- 1
2
3 Jost, J. & Banaji, M. 1994. The role of stereotyping in system justification and the production of
4 false consciousness. *British Journal of Social Psychology Quarterly*, 33: 1-27.
5
6
7 Kinder, D. R. & Sears, D. O. 1981. Prejudice and politics: Symbolic racism versus racial threats
8 to the good life. *Journal of Personality and Social Psychology*, 40: 414-431.
9
10
11 Kleiner, K. D., Akers, R., Burke, B. L., & Werner, E. J. 2002. Parent and physician attitudes
12 regarding electronic communication in pediatric practices. *Pediatrics*, 109(5): 740-744.
13
14 Laine, C. & Davidoff, F. 1996. Patient-centered medicine: A professional evolution. *Journal of*
15 *the American Medical Association*, 275: 152-156.
16
17
18 Landy, F. J., Shankster, L. J., & Kohler, S. S. 1994. Personnel selection and placement. *Annual*
19 *Review of Psychology*, 45: 261-296.
20
21
22 LaRosa, J. C., He, J., & Vupputuri, S. 1999. Effect of statins on risk of coronary disease: a meta-
23 analysis of randomized controlled trials. *Journal of the American Medical Association*,
24 282: 2340-2346.
25
26
27 Latham, G. P. & Wexley, K. N. 1994. *Increasing productivity through performance appraisal*.
28 Reading, MA: Addison Wesley.
29
30
31 Lynch, F. R. 1991. *Invisible victims: White males and the crisis of affirmative action*. Westport,
32 CN: Praeger.
33
34
35 Miles, R. E. & Snow, C. C. 1995. The new network firm: A spherical structure built on a human
36 investment philosophy. *Organizational Dynamics*, 23: 4-18.
37
38
39 Mishler, E. G. & Waxler, N. E. 1963. Decision process in psychiatric hospitalization: Patients
40 referred, accepted and admitted to a psychiatric hospital. *American Sociological Review*,
41 28: 576-587.
42
43
44 O'Leary, V. E. & Ickovics, J. R. 1992. Cracking the glass ceiling: Overcoming isolation and
45 alienation. In U. Sekaran & F. Leong (Eds.), *Woman power: Managing in times of*
46 *demographic turbulence*: 7-30. Thousand Oaks, CA: Sage.
47
48
49 Petersen, T. & Morgan, L. 1995. Separate and Unequal: Occupation-Establishment Segregation
50 and the Gender Wage Gap. *American Journal of Sociology* 101(2): 329-365.
51
52
53 Pulakos, E. D., White, L. A., Oppler, S. H., & Borman, W. C. 1989. Examination of race and sex
54 effects on performance ratings. *Journal of Applied Psychology*, 74(5): 770-780.
55
56
57
58
59
60 Ridgeway, C. L. 1991. The Social Construction of Status Value: Gender and Other Nominal
Characteristics. *Social Forces*, 70(367-386).

- 1
2
3 Rotundo, M. & Sackett, P. S. 1999. Effect of rater race on conclusions regarding differential
4 prediction in cognitive ability tests. *Journal of Applied Psychology*, 84: 815-822.
5
6
7 Sackett, P. R. & DuBois, C. L. Z. 1991. Rater-ratee race effects on performance evaluation:
8 Challenging meta-analytic conclusions. *Journal of Applied Psychology*, 76: 873-877.
9
10
11 Salam, S., Cox, J. F., & Sims, H. P., Jr. 1997. In the eye of the beholder: How leadership relates
12 to 360-degree performance ratings. *Group and Organization Management*, 2: 185-209.
13
14 Shackelford, S., Wood, W., & Worchel, S. 1996. Behavioral styles and the influence of women
15 in mixed-sex groups. *Social Psychology Quarterly*, 59: 284-293.
16
17
18 Sidanius, J. & Pratto, F. 1999. *Social Dominance: An Intergroup Theory of Social Hierarchy*
19 *and Oppression*. Cambridge, U.K.: Cambridge University Press.
20
21 Smith, E. R. & Zárate, M. A. 1990. Exemplar and prototype use in social categorization. *Social*
22 *Cognition*, 8: 243-262.
23
24
25 Smith, E. R. & Zárate, M. A. 1992. Exemplar-based model of social judgment. *Psychological*
26 *Review*, 99: 3-21.
27
28
29 Stangor, C., Lynch, L., Duan, C., & Glass, B. 1992. Categorization of individuals on the basis of
30 multiple social features. *Journal of Personality and Social Psychology*, 62(2): 207-218.
31
32
33 Stewart, M., Brown, J. B., Donner, A., McWhinney, I. R., Oates, J., Weston, W. W., & Jordan, J.
34 2000. The impact of patient-centered care on outcomes. *Journal of Family Practice*,
35 49(9): 796-804.
36
37
38 Stroessner, S. 1996. Social categorization by race or sex: Effects of perceived non-normalcy on
39 response times. *Social Cognition*, 14(247-276).
40
41
42 Stroh, L. K., Brett, J. M., & Reilly, A. H. 1992. All the right stuff. *Journal of Applied*
43 *Psychology*, 77(3): 251-260.
44
45
46 Tabachnick, B. G. & Fidell, L. S. 2003. *Using Multivariate Statistics* (Fifth ed.). Boston: Allyn
47 and Bacon.
48
49
50 Taylor, H. 2002. Patient/physician online communication: Many patients want it, would pay for
51 it, and it would influence their choice of doctors and health plans. *Harris Interactive*
52 *Health Care News*, 2(8): 1-3.
53
54
55 Vecchio, R. P. 2002. Leadership and gender advantage *Leadership Quarterly*, 13(6): 643-671.
56
57
58
59
60 Wenneras, C. & Wold, A. 1997. Nepotism and Sexism in Peer Review. *Nature*, 387: 341-343.

- 1
2
3 Westphal, J. D. & Stern, I. 2007. Flattery will get you everywhere (especially if you are a male
4 caucasian): Ingratiation, boardroom behavior, and demographic minority status affect the
5 likelihood of gaining additional board appointments at U.S. companies. *Academy*
6 *Management Journal*, in press.
7
8
9 Wilkinson, T. J. & Fontaine, S. 2002. Patients' global ratings of student competence. Unreliable
10 contamination or gold standard? . *Medical Education*, 36(12): 1117-1121.
11
12 Willerson, J. T. & Cohn, J. N. (Eds.). 2000. *Cardiovascular medicine*. Philadelphia: Churchill
13 Livingstone.
14
15
16 Wilson, G., Sakura-Lemessy, I., & West, J. P. 1999. Reaching the Top: Racial Differences in
17 Mobility Paths to Upper-Tier Occupations. *Work and Occupations*, 26(2): 165-186.
18
19
20 Woehr, D. & Roch, S. G. 1996. Context effects in performance evaluation: The impact of rater
21 gender and performance level. *Organizational Behavior and Human Decision*
22 *Processes*, 66(1): 31-41.
23
24
25 Yarkin, K. L., Town, J. P., & Wallston, B. S. 1982. Blacks and women must try harder: Stimulus
26 person's race and sex attributions of causality. *Personality and Social Psychology*
27 *Bulletin*, 8: 21-24.
28
29
30 Yusuf, S., Sleight, P., Pogue, J., Bosch, J., Davies, R., & Dagenais, G. 2000. Effects of an
31 angiotensin-converting-enzyme inhibitor, ramipril, on cardiovascular events in high-risk
32 patients. *New England Journal of Medicine*, 342(3): 145-153.
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1
Means, Standard Deviations, and Correlations between Predictor Variables and Dependent Variable³

	M	s.d.	1	2	3	4	5	6	7	8	9	10	11	12
1. Patient Satisfaction	0.51	0.11	-											
2. Practice Busyness	0.66	0.47	-.30	-										
3. Full time status	0.80	0.20	-.07	.11	-									
4. Number of patients in panel	1749.77	550.63	-.10	.26	.59	-								
5. Panel age	45.84	4.89	.07	-.07	.05	-.03	-							
6. Chronic sickness of panel	1.04	0.12	.13	-.12	-.15	-.13	.55	-						
7. Tenure with Prohealth (years)	14.81	8.51	.20	-.14	.14	.08	.33	-.20	-					
8. Age (years)	50.34	6.58	.09	-.09	.16	.21	.29	-.05	.69	-				
9. Race (0=white; 1=non-white)	.12	0.32	-.15	.02	.01	-.04	-.14	-.05	-.12	-.03	-			
10. Sex (0=male, 1=female)	.38	0.49	-.06	-.06	-.63	-.44	-.21	.04	-.25	-.31	.12	-		
11. Productivity	23.00	1.97	.05	.12	.22	.30	-.06	.22	-.25	-.01	-.04	-.15	-	
12. Quality	-0.01	1.55	.11	.03	.07	.08	.21	.05	.14	.11	.02	.04	-.03	-
13. Accessibility to Patients	0.16	0.15	.13	-.11	-.18	-.23	.05	-.06	-.04	-.16	-.11	.21	.05	.23

³ N = 113; All correlations larger than .15 are significant at p<.05

TABLE 2
Analysis Examining Moderating Effects of Physician Race, Sex and Objective Performance
on Patient Perceptions of Physician Performance Quality⁴

	Model 1	Model 2	Model 3
Constant	.60	.69	.67
Controls			
Practice Busyness	-.05**	-.04†	-.05*
Full time status	-.07	-.11	-.13†
Number of patients in panel	.00	.00	.00
Panel age	-.01†	-.01*	-.01*
Chronic sickness of panel	.24†	.27*	.22†
Tenure with Prohealth (years)	.01**	.01**	.01**
Age (years)	.00	.00	.00
Direct effects			
Race (0=white; 1=non-white)	-.03	-.03	-.05†
Sex (0=male, 1=female)	-.04†	-.05*	-.04†
Productivity	.01	.01	.01†
Quality	.01	.01*	.01†
Accessibility to Patients	.06	.09	.03
Interactions			
Productivity X Sex		.00	.00
Quality X Sex		-.03**	-.03**
Accessibility X Sex		-.26*	-.27*
Productivity X Race			-.03**
Quality X Race			-.04*
Accessibility X Race			-.36
R ²	.222**	.306**	.367*
ΔR ² from previous model		.084**	.061*

⁴ One tailed t-test used for hypothesized coefficients, two-tailed for non-hypothesized coefficients; †p<.10 * p<.05 ** p<.01 N = 113. All participants are medical doctors. The sample consisted of 100 whites, 10 Asian or Pacific Islanders, 2 blacks, and one Native American.

Figure 1. The influence of physician sex and performance of patient-centered behaviors on patient perceptions of physician performance quality

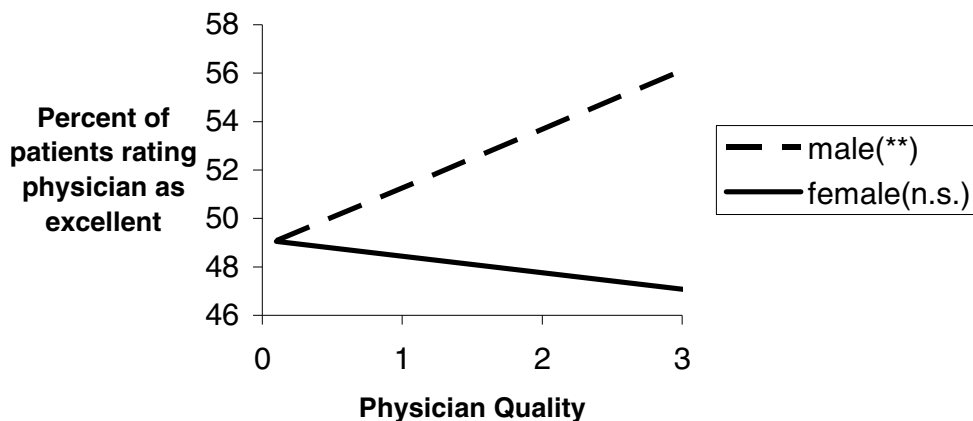
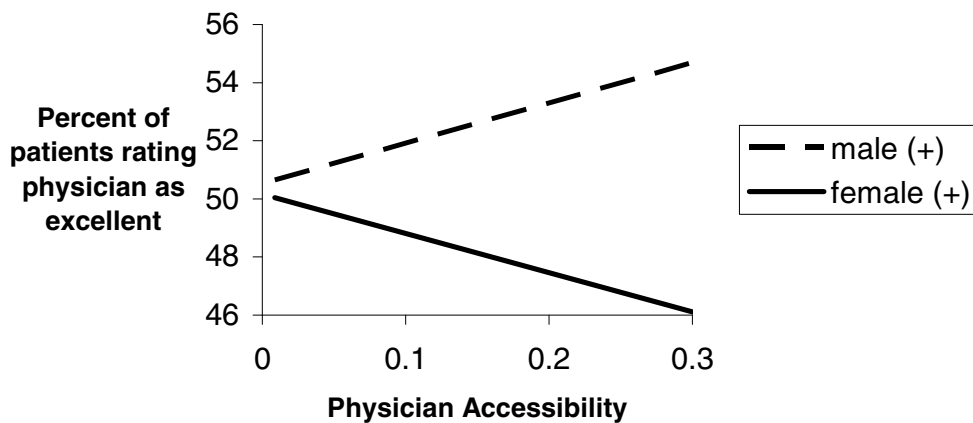


Figure 2. The influence of physician sex and performance of patient-centered behaviors on patient perceptions of physician performance quality



** p < .01
 * p < .05
 + p < .10
 n.s. p > .10

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 3. The influence of physician race and performance of patient-centered behaviors on patient perceptions of physician performance quality

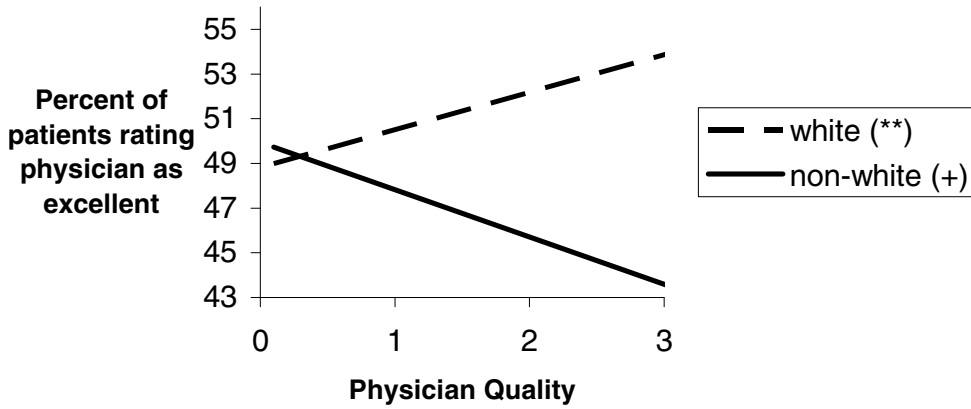
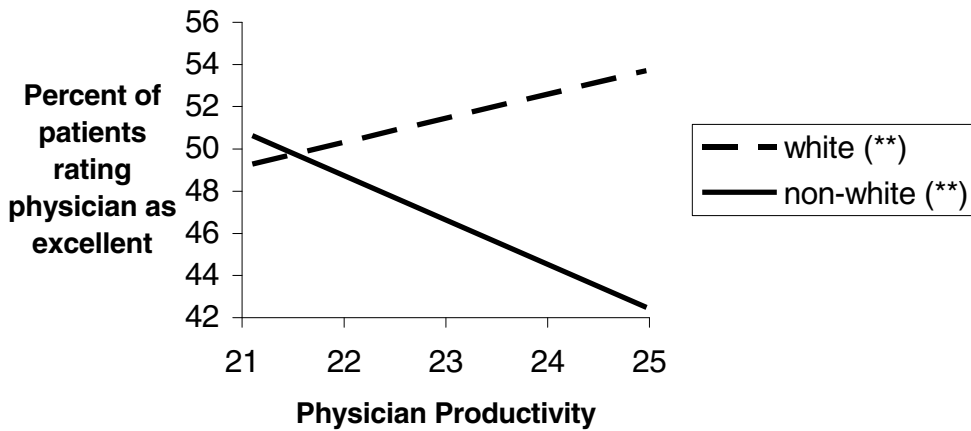


Figure 4. The influence of physician race and performance of patient-centered behaviors on patient perceptions of physician performance quality



** p < .01
 * p < .05
 + p < .10