

The Existence and Nature of Racial Bias in Supervisory Ratings

Joseph M. Stauffer
Texas A&M University

M. Ronald Buckley
University of Oklahoma

The purpose of this article is to facilitate a reconsideration of what the authors consider to be a mistaken belief among personnel psychologists—the belief that supervisory ratings of job performance are not biased on the basis of race. In this article, the authors reviewed the current literature, reexamined the data, and concluded that the research on which this mistaken belief is based clearly demonstrates that racial bias may indeed exist and is significant, both in statistical and practical terms.

The development of criteria that are free from bias is an important pursuit for psychologists interested in selection procedures. We find it interesting that except for a few meta-analyses, a dearth of research studies concerning criterion bias have been published (Martocchio & Whitener, 1992). From a moral, ethical, legal, and fairness standpoint, there is a need to know whether this criterion of great interest among researchers (supervisory ratings of performance) is contaminated by a factor (racial bias) unrelated to the ultimate criterion. We suggest that the time has come to revisit this issue.

The most explicit refutation of racial bias in supervisory ratings comes from a review of research on personnel selection by Landy, Shankster, and Kohler (1994). Landy et al. examined the issue of whether supervisory ratings exhibited any kind of racial or gender bias. They concluded that there was no bias present in supervisory ratings and opined the following: “No one seems to have found direct gender or race effects, at least in large-scale field studies” (p. 283).

In reviewing this same literature, Latham and Wexley (1994) reached a similar conclusion, stating that: “[W]hen the appraiser uses behaviorally based appraisal scales, rater characteristics, such as age, race, and sex, have a negligible effect on the resulting performance appraisal” (p. 152).

We believe that this is the general understanding that most personnel psychologists have regarding this issue. Other authors, although not dismissing claims of racial bias as explicitly as Landy et al. (1994) and Latham and Wexley (1994), seem to either implicitly endorse the belief that there is no bias in supervisory ratings or present the idea in such a manner that it leaves one with the impression that such bias does not exist. In fact, Waldman and Avolio (1991) reported that there was no racial bias in performance evaluation in their large sample, and furthermore, they suggested that there was a need to more closely examine the qualitative work experiences of both Black and White employees to explain performance differences between Blacks and Whites.

We believe that the data may reveal something quite different—there is racial bias in supervisory ratings. Furthermore, it may be found in the very large-scale field studies that Landy et al. (1994) cited to support their claim that supervisory ratings are free from bias. The bias that we believe exists is, in the language of analysis of variance, a rater–ratee race interaction effect. Specifically, although Black and White supervisors tend to agree that White workers’ mean performance is higher than that of Black workers, they disagree substantially on the size of that disparity. For Black supervisors, the difference is negligible, but for White supervisors, the difference is about one third of a standard deviation, before correcting for measurement error. The result is that for the average White worker, it makes no difference who does the rating, Black or White supervisor. However, for the average Black worker, it does matter. We cannot tell from these data whether Black raters are inflating the performance of Blacks, White raters are downgrading the performance of Blacks, or both phenomena are occurring. However, the issue remains that there may be bias, and it has numerous and profound consequences.

The important question is as follows: Why has no one reviewing this literature come to the conclusion that supervisory ratings are indeed biased? We think that there are three basic reasons why this might have occurred.

First, some researchers who have found a *statistically* significant interaction effect have dismissed the effect as not being *practically* significant. That is, they find that the size of the effect, when stated in terms of the percentage of variance accounted for, is trivial (e.g., Pulakos, White, Oppler, & Borman, 1989). However, translating effects into the percentage of variance accounted for improperly obfuscates the actual effect and may lead to “substantively erroneous” conclusions (Hunter & Schmidt, 1990, p. 199). We believe that the magnitude of the disagreement between Black and White supervisors when rating Black workers (about one third of a standard deviation) is practically as well as statistically significant.

The second reason we believe that others have failed to find this bias is that, even when the interaction is noted and not dismissed as trivial, there is a reluctance to call it bias without first understanding its source. For example, despite the fact that they found significant rater–ratee race effects (which normally would serve to indicate bias), Rotundo and Sackett (1999) dismissed these effects as evidence of bias because of the following:

It is not clear whether a rater–ratee race effect reflects inflated ratings by one group of raters, deflated ratings by another, or some combi-

Joseph M. Stauffer, Department of Management and Marketing, Texas A&M University; M. Ronald Buckley, Division of Management, University of Oklahoma.

Correspondence concerning this article should be addressed to M. Ronald Buckley, Division of Management, University of Oklahoma, Norman, OK 73019-0450. E-mail: mbuckley@ou.edu

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

nation of the two. Thus, there is no definitive way of determining whether the criterion used in a validity study is biased. (p. 816)

In fact, Rotundo and Sackett (1999) claimed that “there is no current method of establishing whether there is bias in performance ratings” (p. 816). So even when an interaction effect is found to be significant, personnel psychologists appear reluctant to conclude that the interaction reflects bias. In this study, we used a repeated-measures design to examine the rater–ratee race interaction effect.

Third, in perhaps the most important and influential study, the finding of a rater–ratee race effect was obscured by the study’s focus on a more restricted question—a question that prompted the study. Sackett and DuBois (1991) were specifically challenging the idea that supervisors rated members of their own race more favorably than those of another race. The more general interaction effect (the one in which we find bias) was not the explicit focus of the investigation and, though detected, was never properly highlighted or identified as bias. The authors and subsequent reviewers all overlooked its potential significance. We would like to reexamine Sackett and DuBois’s study to elucidate the existence and nature of racial bias in supervisory ratings.

Background of Sackett and DuBois’s (1991) Study

In 1989, the National Research Council (NRC) issued the report on its investigation of test bias in the U.S. Employment Service’s General Aptitude Test Battery (GATB) (Hartigan & Wigdor, 1989). In general, the NRC found the GATB to be unbiased. The NRC did, however, issue an important caveat. It observed that the criterion measure, supervisory ratings, was supplied mainly by White raters, and there existed evidence at that time that raters tended to rate members of their own race more favorably than members of other races. The possibility that such a criterion bias could obfuscate bias in the predictor made the NRC’s findings tentative.

When issuing their caveat, the NRC cited the work of Kraiger and Ford (1985). In a meta-analysis of rater–ratee race effects, Kraiger and Ford found that White raters gave White ratees higher ratings than they did Black ratees. Conversely, Black raters were found to give Black ratees higher ratings than they gave to White ratees. Sackett and DuBois (1991) challenged Kraiger and Ford’s finding that raters favor members of their own race. They pointed out a fundamental shortcoming in previous research by claiming that all but one study used a between-groups design in which each participant was evaluated by a single rater. Only the Pulakos et al. (1989) study incorporated a repeated-measures design, having both a Black and a White rater evaluate the same individual. Therefore, the differences in mean ratings for Blacks and Whites observed in all other studies, including those in Kraiger and Ford’s meta-analysis, could have been due to true mean performance differences.

To overcome this limitation, Sackett and DuBois (1991) also conducted repeated-measures analyses. The repeated measures consisted of performance scores of individuals rated by both a Black supervisor and a White supervisor and were obtained from the U.S. Army’s Project A (Campbell, 1990) and the U.S. Employment Service’s GATB validity study database. These were the large-scale field studies to which Landy et al. (1994) referred. In

these studies, the researchers used repeated-measures analyses to control for true performance differences that used the same workers (ratees) for both the Black supervisor treatment and the White supervisor treatment. The results obtained from repeated-measures analyses differed substantially from those of the meta-analysis. They found that, on average, both Black and White supervisors gave Whites higher ratings than Blacks. This finding supports previous assertions that a true difference exists in mean group job performance between Blacks and Whites (Schmidt, 1988).

Although such a finding is important, it does not address the larger question of whether racial bias exists and is not sufficient to dismiss the concerns of the NRC. If there were a true mean performance difference between Blacks and Whites, and there existed no racial bias in the performance ratings, then Black and White supervisors should not only agree on the direction of the difference in mean job performance but they should also agree on the magnitude of the difference. A reexamination of the large-scale field study data reported in Sackett and DuBois’s (1991) study shows that this is not the case.

Meaning of Race Effects

Table 1 shows the typical rater–ratee race effects study design. There are four cells corresponding to the four possible rating combinations: Whites rating Whites, Blacks rating Whites, Whites rating Blacks, and Blacks rating Blacks. There are three effects associated with this design.

The difference between the row marginal means, $M_{.W} - M_{.B}$, defines the worker race, or between-groups, effect. This is the only effect that is not considered evidence of racial bias. The worker race effect reflects the mean difference in job performance between Blacks and Whites as judged by all supervisors, Black and White.

The supervisor race effect, unlike the worker race effect, reflects a source of bias and is defined by $M_{.W} - M_{.B}$, the difference between the column marginal means. In the between-groups design, in which each supervisor treatment (Black and White) rates a separate sample of workers, the worker race effect is also a between-groups effect. The difference $M_{.W} - M_{.B}$ reflects the mean performance difference between the group (sample) of Black and White workers rated by Whites and the group of Black and White workers rated by Blacks. Such a difference could confound multiple effects in that it could represent true differences in performance that might result from differences between the two worker samples in ability, education, experience, resources, and so forth.

Table 1
Typical Rater × Ratee Race Effects Study Design

Ratee	White rater	Black rater	
White	M_{WW}	M_{WB}	$M_{.W}$
Black	M_{BW}	M_{BB}	$M_{.B}$
	$M_{.W}$	$M_{.B}$	

Note. The first letter indicates the mean; the first letter of the subscript denotes ratee race (W = White, B = Black); the second letter of the subscript denotes rater race. Marginal mean values are indicated with a dot.

Repeated-measures designs, however, use identical worker samples for both the Black and White supervisor treatments, making $M_W - M_B$ a within-groups (i.e., worker group) effect with an expected value of zero. The presence of a nonzero supervisor race effect means that one group of supervisors (Black or White) systematically rates everyone (both Black and White workers), on average, higher than does the other group of supervisors. Even though this type of rating bias does not actually alter the worker race (between-groups) effect, it could prove inequitable in the workplace and troublesome in most research settings in which it is rare that an individual receives performance ratings from both a Black supervisor and a White supervisor.

Even though the rater race main effect is technically a biasing effect, it is the interaction effect that defines what is typically considered racial bias. A significant interaction suggests that the two groups of supervisors (Black and White) do not agree on the magnitude of any performance difference between the two groups of workers. This disagreement can take a variety of forms, but in graphical terms, the worker or supervisor race effects are nonparallel. In Kraiger and Ford's (1985) meta-analysis, for example, the finding was that both groups of supervisors found a mean performance difference, but it was reversed. The Black supervisors saw the difference in favor of the Black worker group; the White supervisors saw the difference in favor of the White worker group. Graphically, the effects were not only nonparallel but they crossed.

A repeated-measures design is better suited to detect this interaction. In such a design, the expected performance difference between the repeated-measures is zero, and the true difference between group performance means is equal from one supervisor treatment to the other. Even with a main supervisor race effect, the two simple within-groups effects should be parallel. A significant interaction would, therefore, indicate a systematic distortion of true performance relationships by one or both rater groups.

Method

Results that were obtained from Sackett and DuBois's (1991) study, which used repeated-measures designs, were reexamined for evidence of interaction. The results from the between-groups analyses were not examined for reasons previously stated. The study reported a sample of 561 Black and 1,259 White soldiers and a sample of 331 Black and 286 White civilian workers, each rated by a Black and a White supervisor.

Two forms of evidence were obtained by computing standardized difference statistics (d values) from the repeated-measures designs. The statistical tests were not reexamined because, as we have noted, we do not dispute the fact that the interaction in question is statistically significant. Sackett and DuBois (1991) presented d values for worker race effects from the between-groups analyses of both studies. It would be more useful to have the d values from the repeated-measures analyses, particularly for supervisor race effects, because the expected value of these effects is zero. Fortunately, Sackett and DuBois (1991) presented standardized means of supervisory ratings of job performance from which these d values could be constructed (p. 875). (Recall that in the repeated-measures design, worker race is the between-groups effect, and supervisor race is the within-groups effect.) Hence, the first form of evidence derived from comparing the d values for Black supervisors with those for White supervisors served to determine whether the between-groups effects were parallel. The second form of evidence derived from comparing the d values for the Black worker group with those of the White worker group served to assess differences in the within-groups effects.

There were three different performance dimensions used in the military studies: technical skill and effort, personal discipline, and military bearing.

Sackett and DuBois (1991) expressed the opinion that the first dimension, technical skill and effort, was directly comparable with the overall performance measure used in the civilian studies. As a result, we confined our examination of military ratings to those made for technical skill and effort.

Results

The reexamination revealed clear evidence of significant rater-ratee race interaction. The between-groups and within-groups d values are presented in Tables 2 and 3, respectively. Table 2 shows that, for both the civilian data and the military data, between-groups d values were positive for both the White supervisor and the Black supervisor treatments. This indicated that both Black and White supervisors reported higher mean group performance for Whites than they did Blacks. There were, however, sizable differences between these d values, suggesting an interaction effect. In the civilian data, the d values for the White supervisor treatment were approximately 9 times larger than those for the Black supervisor treatment. This disparity was smaller in the military data but still relatively large (nearly 7 times). For the civilian data, the standardized difference for the White supervisor treatment was .333 (uncorrected for unreliability), with a standardized difference of .037 for the Black supervisor treatment. Similarly, for the military data, the standardized difference for the White rater treatment was .300, compared with .044 for the Black rater treatment.

A clearer picture emerged when the within-groups d values were examined. These are presented in Table 3 and represented graphically in Figures 1 and 2. Because the true performance difference between these repeated-measures is expected to be zero (the same performance is viewed by different supervisors), a nonzero value suggests systematic differences in the way Blacks and Whites rate members of that particular group of workers. This was the case for all worker groups (Black and White, military and civilian), although they might appear negligible for the White worker groups.

In the civilian data set, $d = .029$ (uncorrected) for the White worker group, indicating that White supervisors rated Whites slightly higher than did Black supervisors. An opposite and much larger effect was found for the Black worker group ($d = -.267$). Black supervisors rated Blacks higher than did White supervisors. In the military data, Black supervisors rated both Whites and Blacks higher than did White supervisors. Again, the contrast was much larger for Black workers than for White workers ($-.290$ vs. $-.034$, respectively).

Table 2
Standardized Ratee Race (Between-Groups) Effects Listed by Rater Race

Corrected set	White rater		Black rater	
	Uncorrected d	Corrected d^a	Uncorrected d	Corrected d
Civilian ^b	.333	.398	.037	.044
Military ^c	.300	.359	.044	.053

Note. Positive values mean that White ratees received higher ratings than Black ratees from that group of raters.

^a Corrected for attenuation with $r_{yy} = .70$, the value used by Sackett and DuBois (1991). ^b White ($n = 286$), Black ($n = 331$). ^c White ($n = 1,259$), Black ($n = 561$).

Table 3
Standardized Rater Race (Within-Groups) Effects Listed by
Ratee Race

Corrected data set	White ratees		Black ratees	
	Uncorrected <i>d</i>	Corrected <i>d</i> ^a	Uncorrected <i>d</i>	Corrected <i>d</i>
Civilian ^b	.029	.035	-.267	-.319
Military ^c	-.034	-.041	-.290	-.347

Note. Negative values mean that Black raters assigned higher ratings than White raters for that group of ratees.

^a Corrected for attenuation with $r_{yy} = .70$, the value used by Sackett and DuBois (1991). ^b White ($n = 286$), Black ($n = 331$). ^c White ($n = 1,259$), Black ($n = 561$).

Discussion

These within-groups *d* values, with expected values of zero, and the large differences in between-groups *d* values provide clear evidence for bias. Given an identical sample of Black and White workers, White supervisors gave significantly, both statistically and substantively, higher mean ratings to White workers relative to Black workers than did Black supervisors. It would be difficult to dispute that systematic bias exists in this sample of supervisory ratings of job performance. Both Black and White supervisors reported a mean group performance difference that favored Whites (see Table 2). They did not, however, agree on the size of this difference. The perhaps trivial difference between Black and White supervisors when assessing Whites, coupled with the large difference between these two groups of raters when assessing Blacks, suggests that the problem lies in assessing the job performance of Blacks. It is impossible to discern from these data whether the bias resulted from White supervisors understating the performance of Blacks, Black supervisors overstating it, or a combination of both.

It is also difficult to argue from these results that the bias in the performance ratings of Blacks is trivial, that is, not practically significant. First of all, if the interaction were practically insignificant to us, then we would be indifferent as to which supervisors did the rating, and this is clearly not the case. Black workers should certainly prefer Black supervisors. Second, if the approximate .300 standard deviation difference between Black and White ratings of the performance of Blacks was trivial (i.e., small enough to be considered zero), then we would have to declare the approximate .300 standard deviation difference in the mean job performance of Blacks and Whites found by the NRC to be essentially zero as well (Hartigan & Wigdor, 1989). Third, depending on the point where successful performance is defined, a .300 standard deviation increase in the performance ratings of Blacks could result in up to 12% more Blacks being considered successful performers. All of these suggest that the magnitude of the effect is not trivial.

This reanalysis is an important step in recognizing that racial bias may indeed exist in supervisory ratings of job performance. It is not sufficient to demonstrate agreement among Black and White supervisors that Whites perform at a higher level than Blacks. We must recognize what systematic bias may exist, try to understand how and why it occurs, and develop strategies toward eliminating bias. The practical implications of the results of this study are unmistakable: If you are a White ratee, then it does not matter whether your supervisor is Black or White. If you are a Black ratee, then it is important whether your supervisor is Black or White. Because organizations and researchers rely extensively on supervisory ratings of job performance, future research efforts should be directed at better understanding the nature of the disparity between supervisors of different races, particularly the disparity between Black and White supervisors when assessing the job performance of Blacks.

It would be wrong to report that these differences are due to some kind of racial prejudice. The data are insufficient to address

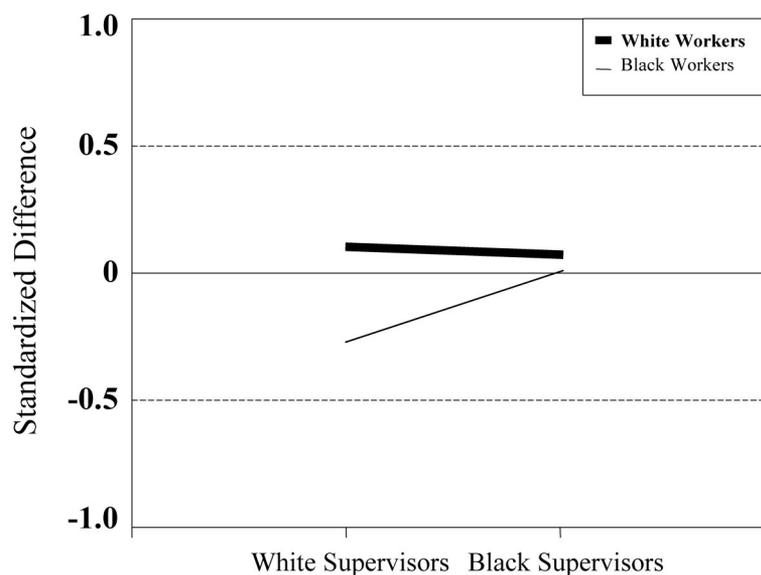


Figure 1. Standardized differences for civilian data.

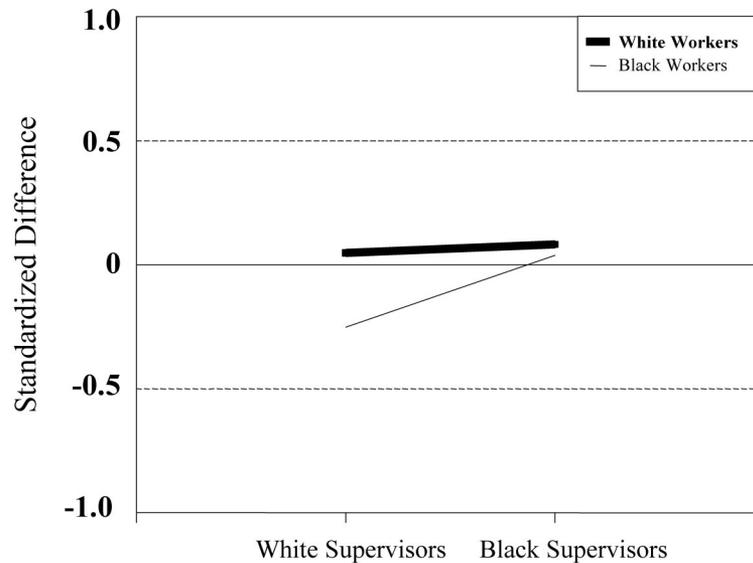


Figure 2. Standardized differences for military data.

or support such a conclusion. Different data (e.g., racial attitudes) must be collected if this area is to be addressed in the future. We suggest that it may be fecund to look in other directions. For example, the work of Chattopadhyay and colleagues (e.g., Chattopadhyay, 1999; Finn & Chattopadhyay, 2000; George & Chattopadhyay, 2002) may provide a stepping off point for us to explore why Black and White supervisors rate the performance of Blacks differently. Chattopadhyay and colleagues have used different theoretical frameworks in an attempt to understand how demographic dissimilarity influences people's attitudes and behaviors within organizations. For example, with regard to race, the similarity-attraction paradigm (Byrne, 1971) suggests that supervisors are attracted to members of their own race. The ratings process may be influenced by that attraction. Social identity theory (Tajfel & Turner, 1986) suggests that supervisors use demographic variables such as race to form in-groups and out-groups (theirs being the in-group) and behave in ways that promote the superiority of the in-groups over the out-groups, a phenomenon known as social identity enhancement. The findings of Chattopadhyay and other researchers working in this area may be consistent with the finding that both Black and White supervisors rate Whites higher and disagree only on their ratings of Black workers. That is, the demographic dissimilarity literature has demonstrated that there are many subtleties beyond the main effects predicted by initial theories, and the field is currently concerned with sorting out those contingencies. For example, George and Chattopadhyay (2002) reported that members of higher status groups view the lower status groups more unfavorably when the lower status groups compose the bulk of the workforce. Lower status work groups do not always exhibit this tendency when they are in the minority. These findings may provide an explanation of the findings with respect to the studies of the relationship between race and supervisor ratings.

Because the question of criterion bias is inextricably linked to that of predictor bias, it is essential to recognize the impact this finding has on our investigations of test bias. Much of the research

on criterion bias that we have cited was prompted by the concern of the NRC (Hartigan & Wigdor, 1989) and others (e.g., Burke, 1984; Linn & Werts, 1971; Petersen & Novick, 1976) that criterion bias could influence conclusions regarding test bias. Although the finding that Black and White supervisors disagree substantially on the mean performance of Blacks does not mean that, for example, ability tests are biased against Blacks, it does mean that the issue deserves more scrutiny. Traditional methods of investigating bias rely on the assumption that the criterion is free from bias. What is needed are more studies similar to Rotundo and Sackett's (1999) study that try to look at this issue in novel and creative ways—ways that do not rely on the assumption of an unbiased criterion or studies that use more disciplined criteria that are found to be free from bias. In sum, we feel that a renewed call for more research into the cause of this bias and its impact on myriad outcomes, including test bias, is necessary. Martocchio and Whitener (1992) have encouraged "others to evaluate fairness in their validation work and submit their findings to public review in scholarly journals" (p. 502). We certainly agree with them and encourage this scrutiny.

References

- Burke, M. J. (1984). Validity generalization: A review and critique of the correlation model. *Personnel Psychology, 37*, 93–115.
- Byrne, D. (1971). *The attraction paradigm*. New York: Academic Press.
- Campbell, J. P. (1990). An overview of the Army Selection and Classification Project (Project A). *Personnel Psychology, 43*, 231–239.
- Chattopadhyay, P. (1999). Beyond direct and symmetrical effects: The influence of demographic dissimilarity on organizational citizenship behavior. *Academy of Management Review, 86*, 781–788.
- Finn, C., & Chattopadhyay, P. (2000). Managing emotions in diverse work teams: An affective events perspective. *Academy of Management Best Paper Proceedings*, d1–d6.
- George, E., & Chattopadhyay, P. (2002). Do differences matter? Understanding demography-related effects in organizations. *Australian Journal of Management, 27*, 47–55.

- Hartigan, J. A., & Wigdor, A. K. (1989). *Fairness in employment testing: Validity generalization, minority issues, and the General Aptitude Test Battery*. Washington, DC: National Academy Press.
- Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.
- Kraiger, K., & Ford, J. K. (1985). A meta-analysis of ratee race effects in performance ratings. *Journal of Applied Psychology, 70*, 56–65.
- Landy, F. J., Shankster, L. J., & Kohler, S. S. (1994). Personnel selection and placement. *Annual Review of Psychology, 45*, 261–296.
- Latham, G. P., & Wexley, K. N. (1994). *Increasing productivity through performance appraisal*. Reading, MA: Addison Wesley.
- Linn, R. L., & Werts, C. E. (1971). Considerations for studies of test bias. *Journal of Educational Measurement, 8*, 1–4.
- Martocchio, J. J., & Whitener, E. M. (1992). Fairness in personnel selection: A meta-analysis. *Human Relations, 45*, 489–506.
- Petersen, N. S., & Novick, M. R. (1976). An evaluation of some models for culture fair selection. *Journal of Educational Measurement, 13*, 3–29.
- Pulakos, E. D., White, L. A., Oppler, S. H., & Borman, W. C. (1989). Examination of race and sex effects on performance ratings. *Journal of Applied Psychology, 74*, 770–780.
- Rotundo, M., & Sackett, P. S. (1999). Effect of rater race on conclusions regarding differential prediction in cognitive ability tests. *Journal of Applied Psychology, 84*, 815–822.
- Sackett, P. R., & DuBois, C. L. Z. (1991). Rater–ratee race effects on performance evaluation: Challenging meta-analytic conclusions. *Journal of Applied Psychology, 76*, 873–877.
- Schmidt, F. L. (1988). The problem of group differences in ability test scores in employment selection. *Journal of Vocational Behavior, 33*, 272–292.
- Tajfel, H., & Turner, J. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–24). Chicago: Nelson Hall.
- Waldman, D. A., & Avolio, B. J. (1991). Race effects in performance evaluations: Controlling for ability, education, and experience. *Journal of Applied Psychology, 76*, 897–901.

Received January 20, 2003

Revision received December 16, 2003

Accepted January 7, 2004 ■

Low Publication Prices for APA Members and Affiliates

Keeping you up-to-date. All APA Fellows, Members, Associates, and Student Affiliates receive—as part of their annual dues—subscriptions to the *American Psychologist* and *APA Monitor*. High School Teacher and International Affiliates receive subscriptions to the *APA Monitor*, and they may subscribe to the *American Psychologist* at a significantly reduced rate. In addition, all Members and Student Affiliates are eligible for savings of up to 60% (plus a journal credit) on all other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the American Association for Counseling and Development, Academic Press, and Human Sciences Press).

Essential resources. APA members and affiliates receive special rates for purchases of APA books, including the *Publication Manual of the American Psychological Association*, and on dozens of new topical books each year.

Other benefits of membership. Membership in APA also provides eligibility for competitive insurance plans, continuing education programs, reduced APA convention fees, and specialty divisions.

More information. Write to American Psychological Association, Membership Services, 750 First Street, NE, Washington, DC 20002-4242.