

#18

Memorandum

May 24, 1990



TO: Vic Martuza, Chair

FROM: Jan H. Blits *JH*

Your May 8th memo to Frank Murray illustrates exactly what concerns me. You show no interest in avoiding a recurrence of the improprieties that occurred this year in Linda Gottfredson's promotion process. As you present it, there wasn't even the appearance of unfairness this year. The only reason that I think otherwise is that I have "conspiracy theory." If I should encounter prejudice or unfairness next year, I will have only myself to blame. My "inappropriate behavior" in voicing my concerns may "result in creating the very conditions [I] claim [I] want to avoid."

You make seven specific claims. Only the first has any substance, though not much.

1. You are right in pointing out that each of the departments in the College nominates a full professor (as well as an associate and an assistant professor); the Dean selects four of the six, with each rank represented by at least one. However, by your own admission to me, you were sure that the Dean would not select one of your nominees (David Kaplan), because he is a member of this year's Department Committee. That would guarantee the selection of your two other nominees, including Roberta Golinkoff, and would also reduce the chances of the Dean's selecting the full professor nominated by the other department, since all ranks must be represented. My original statement of the point was somewhat inaccurate, but not misleading.

I note that your memo makes no effort to refute my objections to your other committee nominees for next year, including the reappointment of some of this year's committee members.

2. Little needs to be said about the department's "house rule." The facts speak for themselves. Even the most stringent interpretation of the constraints caused by the department's size and the requirements of the committee's composition would not begin to explain the department's repeated practice of appointing someone who has just been promoted. (There has never been a dearth of associate professors in the department.) The department's stated reason for the practice has always been that everyone can benefit from the knowledge the most recent candidate has gained from having gone through the promotion process.



3. What do you mean by saying that "the timing is not right" for Linda's appointment to a P&T committee? Why isn't it right? It is typical and extremely revealing, I think, that you speak of the controversy that has been "generated by her current grievance." The controversy was not generated by her grievance, but rather by what she is grieving. At every opportunity you blame her (for responding to unfair treatment) just as you blame me (for trying to avoid such treatment). As you consistently present it, there was nothing wrong this year until Linda complained. Nor do I understand what you mean by saying that it would make more sense to wait until the controversy "has faded"? Do you mean to suggest that you, as Department Chair, have no interest in making an active effort to find a reasonable and timely resolution of such problems? And if you are willing to be passive about resolving this year's difficulties, what assurance do I have that you will be active in avoiding difficulties next year?

4. You claim that two of your actions show "clear evidence" that you accept the legitimacy of Linda's promotion. The first is that you were one of the first people in the Department to congratulate her. The truth is that you did not go to her, but merely ran into her in the hallway, where some word of congratulations was hardly avoidable.

The second is that you met with her to find ways of resolving--not the conflicts contained in her grievance--but "the conflicts which have resulted from her grievance and the way in which she has conducted her affairs in regards to that grievance." While the former might well have been the sort of evidence you claim, the latter can hardly be called "clear evidence" that you accept the legitimacy of her promotion. Ironically, you accuse me of having "a propensity to use exaggeration and overstatement." Your specific example is interesting. You deny that you said that anyone recommended Linda's promotion because of "political pressure" or that you impugned the integrity of any of your superiors. At the same time, however, you admit that you said that "the threat of a grievance action and hints of civil litigation have the potential to influence such decisions." You don't seem to realize that what you deny you "said," you admit that you suggested. To say that "the threat of a grievance action and hints of civil litigation have the potential to influence such decisions" amounts to impugning the integrity of those who could be influenced.

5. If you never claimed or implied what Frank said that you did, then I can only imagine that he was trying to encourage me to believe that you were, after all, making an effort to address my concerns. I would have thought that, while you might not have agreed with the word "rebuke," you would nonetheless have gladly

taken credit for making a positive gesture. As it turns out, the only step that someone might give you credit for is congratulating Linda when you ran into her in the hallway.

6. I do, indeed, see as a warning the claim that there has been a "racheting up" of promotion criteria. It is true that Julius Meisel was turned down on a second try after a positive evaluation on the first. Julius' case, however, was very weak from the start, and the Department rejected him the second time because his record was not much better and the Department feared that he was certain to be turned down again at higher levels. That was an application, not an elevation, of existing standards.

Another case is more interesting. The University Handbook states that research done prior to joining the University "plays little or no role in the promotion [to associate professor] except to form a meaningful context against which later development and accomplishments can be judged" (III-K-10). Yet, in one of the three cases since Meisel's which involve promotion to associate professor, a member of the Department was promoted to associate professor--with a glowing recommendation from the Department--almost entirely on the strength of a book that was all but completed before joining the university (only the last chapter required some revisions). The same work that got him an appointment in the Department, got him his promotion. I say this not to criticize anyone, but only to refute your claim.

You speak of my "history of interpreting any advice that [I do] not wish to be given as either a threat or a warning" and cite our recent exchange of memos concerning my teaching assignment as "further evidence along these lines." You might note that in one of those memos (January 18) you explicitly threaten me with serious actions unless I either substantiate my charges (which I did) or write a written retraction and apology.

7. Frank keeps trying to reassure me that you could change your mind about my two articles with Linda. I wish I could believe him, but your remarks here about those articles demonstrate continued bias.

You say that you have not seen any information since writing your recommendation against Linda's promotion which has dispelled your concerns. I have, in fact, given you a copy of Frank Schmidt's paper, delivered recently at a SIOP Symposium, which strongly supports the argument that Linda and I made in our articles. Schmidt, who explicitly states his policy differences with us, goes well beyond the language of our articles (which you found so offensive), saying that "What the [NAS] committee did...is a serious offense against a cherished scientific and scholarly value: intellectual honesty." "What was intellectually dishonest was the [panel's] deceptive attempt to provide a bogus statistical, psychometric, and scientific justification for score

adjustments" (pp.4,5). Why would such corroboration not help to dispel your concerns? Why do you ignore even its existence?

That is not all. You also suggest that the reactions that appeared in Transaction/SOCIETY--including two written by members of the panel whose reports we criticize--support your concerns. This is even more disturbing than your ignoring favorable information. Apart from the fact that you ignore one reaction (from Mary Tenopyr) that corroborates all of our main points, you confuse controversy with disinterested evaluation. As Transaction/SOCIETY editor Irving L. Horowitz has emphasized, the articles that you refer to were chosen precisely for the sake of controversy. They were chosen because they could be expected to be strongly critical. Horowitz wanted, and got, controversy--"in the firm belief that the life of social science discourse is enhanced by genuine difference of opinion." When he complied with this year's College Committee's request to see the articles, Horowitz wrote, "To be frank, if these four papers are used as intellectual cannon-fodder to deny Professor Gottfredson a promotion..., I should be dismayed." Fortunately, Jack Pikulski, Chair of the College Committee, treated them appropriately, but you, by contrast, have already seized upon them to confirm your position.

In addition, I have given you a draft of a manuscript which I wrote replying to the opposing articles in Transaction/SOCIETY and have offered to discuss my replies with you. This new information, too, you seem to have ignored.

Is this what I can expect as "fair" treatment next year? You ignore even the existence of positive information and seize upon inappropriate negative information. Contrary to what Frank would have me believe, there is no reason in the world to believe that you will do anything other than try to defend this year's unfair negative evaluation at all costs.

Your memo to Frank, far from reassuring me, only heightens my concern that I will be treated unfairly when I come up for promotion next year. It demonstrates the bias it tries to deny. Just as you consistently blame this year's troubles on Linda's attempt to defend herself, so you already blame next year's likely troubles on my attempt to avoid them. It is hard to believe that you are willing to treat me fairly when everything that you say and do suggests that you would be perfectly content to see this year's treatment of Linda repeated next year in my case.

cc: Frank Murray
George Cicala
Gordon DiRenzo
Maxine Colm

Memorandum #19

May 29, 1990

TO: Frank Murray, Dean
FROM: Jan Blits *JB*
RE: My promotion

I find it intolerable, and I think any disinterested party would find it unfair and unreasonable, to face a Department Chair and a Department P&T Chair both of whom have not only demonstrated bias against some of my work, but who now, as a result of that bias, have a vested interest in reconfirming last year's evaluations of my joint work with Linda. You say in your May 3rd memo that the P&T process next year will provide, in part, the occasion you are looking for "with regard to a confirmation or rejection of this year's committee's evaluation." That is just the problem. Next year's process is precisely the opportunity for those who are currently being grieved for their actions this year to vindicate themselves and render the case against themselves moot. Unwilling or unable now to justify their actions this year in an appropriate fashion, they could escape accountability altogether by producing a majority vote on an ostensibly unbiased committee.

My concern is real. Vic's May 8th memo to you demonstrates that he is unwilling even to acknowledge the existence of evidence that corroborates our analysis of the NAS panel's reports. It also demonstrates that Vic will seize upon whatever he can find, even reactions written by those we criticize, to reconfirm his original evaluation.

I am also concerned about your comments concerning the need to avoid both false negative and false positive errors next year. By stating the problem in such terms, you deftly dispose of the question of prejudice. Bias against me becomes legitimated as mere error, while any support I might enjoy is lowered to the level of mere bias. Your way of reformulating the problem of unfairness obscures the crucial difference between professional judgment and unprofessional bias.

I am perfectly serious about my charges of bias and lack of fairness, and regret that we were not able to resolve the issues informally this spring. I intend to pursue my charges through a formal complaint. However, the intolerable situation of my being judged by those who could seemingly vindicate themselves by recommending against my promotion forces me to postpone my application. I do this with great reluctance, but Vic's recent



memo to you has made quite clear that no good purpose could be served by my coming up for promotion before the issues of bias and unfairness have been properly resolved.

cc: George Cicala
Gordon RiRenzo
Maxine Colm
Ralph Ferretti
Robert Hampel
Victor Martuza

Memorandum #20



June 21, 1990

TO: Faculty, Educational Studies
FROM: Frank Murray, Dean *Frank Murray*
SUBJECT: Linda Gottfredson's Grievance/Complaints

Quite apart from the formal grievance/complaint process, there are some aspects of the case that can, and should be clarified, and I hope settled. One of those is treated in the attached memoranda. As far as I am concerned, and I hope for all of you as well, my conclusion will stand as our position on the point raised in the memoranda.

FBM:jph
Attachment
cc: Maxine Colm
George Cicala
Gordon DiRenzo



Memorandum



June 7, 1990

TO: Vic Martuza, Chair
Department of Educational Studies

FROM: Frank B. Murray, Dean *frank*

You will remember, Vic, that some months ago I asked that the department reconsider what they meant by the term "misrepresent" in their evaluation of Linda Gottfredson's promotion dossier. As far as I know, the question was not revisited and formally taken up by the department. My own inquiry into the matter, in a discussion with the Departmental P&T Committee, indicated that they were not using the term "misrepresent" to connote a willful or intentional distortion of scholarly evidence. They, instead, meant only to say that in their opinion, she got it wrong and that her view of what the report was claiming and theirs differed. The committee members I met with also made it clear that some members of the committee in fact felt she did deliberately distort the record and that she intentionally misled her readers for her own purposes. However, this was not the view of the majority of the committee.

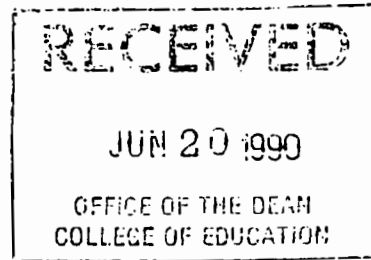
I think you should publicly clarify this point because it is harmful to us all if we have the idea that a member of our faculty lies about the evidence. My conclusion is that the committee was saying that Linda was trying to get the story right but she was in error.

FBM:jph

cc: J. Blits
R. Ferretti
L. Gottfredson
R. Hampel
D. Kaplan
R. Venezky



Memorandum



June 15, 1990

TO: Frank B. Murray, Dean
College of Education

FROM: Victor Martuza, Chair *VM*
Educational Studies

SUBJECT: Response to your memo of June 7, 1990 regarding the use
of the term "misrepresent" in the Department Promotion
and Tenure Committee's recommendation regarding Linda
Gottfredson's promotion application.

It is my understanding that the EDS P & T Committee's use of the term "misrepresent" carried the standard dictionary definition "to state incorrectly" and is quite consistent with the conclusion you state in your memo of June 7, 1990. It was not my understanding at the time the recommendation was issued nor is it now that the Committee meant to indicate that Dr. Gottfredson was guilty of "willful or intentional distortion of scholarly evidence".

I certainly have no objection to your making this memo public if you feel that would serve a useful purpose.

VM:blo

cc: Ralph Ferretti, EDS P & T Committee Chair
Linda Gottfredson
Jan Blits

Memorandum #21

DATE:

June 19, 1990

TO: Frank Murray, Dean
FROM: Linda Gottfredson *lg*
RE: Your Memo and the Chair's about the Term "Misrepresent"

In his June 15 memo to you, Vic says that he has no objection to your making his memo public if you feel that would serve a useful purpose. I do not believe that making it public will serve any useful purpose.

First, the P&T Committee's use of the term "misrepresent" in its recommendation against my promotion is but one small part in the larger set of problems which I have outlined in my AAUP grievances against the Committee and the Chair. To focus on that one term, taken out of context, is to ignore the web of the procedural irregularities and biases. Why did the Committee take the highly unusual step of purposely seeking a negative review? Why did the Committee place such exclusive weight on that one negative review (described by the College Committee as unprofessional), not only ignoring the eight overwhelmingly positive reviews but also the evidence that the one negative review was mistaken? Why did the Committee decide not to provide advance drafts of its recommendation to the department, as is required by our P&T guidelines, and then steadfastly refuse to explain the reasons for its negative recommendation, as is required by the Faculty Handbook? What is the evidence for its claims that my research has been declining in quality since 1983 and that I have a "tendency to misrepresent"? Softening the interpretation of the term "misrepresent" does nothing to address my concerns that the Committee and the Chair were unfair and biased.

Second, it is telling that the Committee does not retract or clarify any of its statements, but rather we are told what Vic thinks the Committee really meant. This is hardly a satisfactory substitute for the Committee providing reasons either for its harsh judgment of my research record or for its irregular and unfair treatment of my promotion application. Nor are Vic's beliefs about the Committee's beliefs a substitute for Vic explaining his own mishandling of my promotion application, or for retracting his own statements to the effect that I purposively mislead or misconstrue the positions of others.

Finally, you indicate in your June 7 memo ("Your

CONSERVE ENERGY SO ENERGY CAN SERVE YOU



Grievance/Complaint") that Ralph Ferretti feels that no useful purpose would be served by my meeting with his committee to find ways of resolving my grievance. This is unfortunate because, as I told both Vic and Ralph when I asked to work with them on the issue, it would have been more constructive for us to hammer out a mutually acceptable solution than for me to have to pursue the matter beyond the College level.

cc: Vic Martuza
Ralph Ferretti
George Cicala
Maxine Colm

Memorandum #22



DATE:

September 17, 1990

TO: Frank Murray, Dean

FROM: Jan Blits and Linda Gottfredson

JB

g

Your June 21 memo to the Department, concerning the P&T Committee's use of the word "misrepresent" in recommending against Linda's promotion, requires a reply.

The issue of the Committee's charge of misrepresentation cannot be properly dealt with apart from the larger web of bias and procedural irregularities of which it is a part. This web includes serious violations of the external peer review process. To ostensibly clarify and settle the issue of the Committee's use of the word "misrepresent" by considering it out of context distorts our grievance/complaint as a whole and gives the false impression that a key issue in the case has been settled when in fact it has only been swept aside.

Furthermore, even if the issue of the Committee's use of the word "misrepresent" could be dealt with in isolation, only the Committee itself can say what it really meant. In the memos that you circulated to the Department, however, the Committee does not retract or clarify any of its statements. Rather, Vic (who states only part of what he calls the "standard dictionary definition" of the word) says what he believes the Committee meant, and you accept his conclusion as your own.

The issues in our grievance/complaint are serious. They go to the very heart of professional and personal integrity in our Department. I'm sure you will recognize the need to treat them accordingly.

cc: Maxine Colm
Gordon DiRenzo
George Cicala
Vic Martuza
Ralph Ferretti
Bob Hampel
Dick Venezky
David Kaplan



SOCIAL SCIENCE AND BLACK/WHITE: II

EMPLOYMENT TESTING AND JOB PERFORMANCE

19

job relatedness, employers virtually ceased using general mental ability tests.

In 1981, however, as a result of new research on testing, the Labor Department's Employment Service began encouraging state employment services to use its General Aptitude Test Battery (GATB) for all job referrals and to rank candidates according to test scores. Intended to improve the state agencies' screening of nearly 20 million applicants a year, the Department's action was based on psychological research that undercut the EEOC's original criticism of general ability tests. The original criticism rested on the fact that a given test's ability to predict job performance—its "validity"—seemed to vary from job to job, and from location to location. Because a test that was valid in one location often appeared to be invalid in other locations for the same job, experts tended to believe that every test needed to be validated separately for each new setting.

The new research, pioneered largely by Frank Schmidt and John Hunter, challenged the need for separate validation. By applying new techniques, Schmidt, Hunter, and other researchers showed that well-developed mental ability tests can be generalized to new locations and even to new jobs. Most of the previously observed variation in a given test's ability to predict performance from one job or setting to another was shown to be merely a statistical artifact caused by small sample sizes.

This new research has disconcerted many opponents of testing, because, while showing that general ability tests are valid for most skilled and semiskilled jobs, it does nothing to reduce their adverse impact. In fact, it shows that the persistent adverse impact of these tests cannot be blamed on the inadequacies of the tests themselves. Several decades of other psychological research have also refuted the still popularly held belief that general ability tests are racially biased—that is, that they underpredict the job performance of blacks. Many separate studies show that, contrary to prevailing public opinion, lower test scores among blacks and Hispanics are accompanied by lower job performance, just as in the case of whites.

The new research has thus revived and intensified the policy storm over general mental ability tests; the question now is not whether but how such tests should be used. Specifically, can they be used without incurring adverse impact?

The Labor Department found evidence to support the following

Employment testing and job performance

JAN H. BLITS & LINDA S. GOTTFREDSON

FOR MORE than twenty years, a fierce public-policy storm has raged over how or whether general mental ability tests should be used in employment selection. Employers typically favor using such tests because they are generally the best single predictor of job performance. But the tests are controversial because they disproportionately screen out blacks, who as a group tend to score considerably lower than whites.

Even though the 1964 Civil Rights Act specifically sanctions general ability tests, the federal government sharply curtailed their use, ostensibly to comply with Title VII of the Act. In 1970 the Equal Employment Opportunity Commission (EEOC) issued guidelines that the Supreme Court soon adopted as law in the landmark case of *Griggs v. Duke Power Company* (1971); after *Griggs*, any employment test that had an adverse impact on blacks was considered discriminatory and therefore illegal unless it could be shown to be specifically job related. In the wake of subsequent court decisions and stringent federal regulations for demonstrating

not only the use of the GATB to improve productivity, but also the adoption of a race-conscious way of recomputing test scores to avoid adverse impact. The new procedure, called race-norming—or, euphemistically, “within-group scoring”—converts raw scores into group-based percentile ranks. Each candidate’s score is reported not in relation to those of all other candidates, but only in relation to the scores of applicants of the same group. Blacks are compared only to other blacks, Hispanics only to other Hispanics, and “others” to all but blacks and Hispanics. Top scorers from all groups are then referred on the basis of their group-based percentiles to employers, who are often not aware that the candidates’ scores have been so adjusted. In effect, race-norming gives bonus points to members of groups that tend to score lower than others, offsetting average differences in scores among groups. By 1986 state employment agencies in forty states had adopted the Labor Department’s new referral procedure.

Race-norming crucially affects employment referrals. Among candidates for semiskilled jobs, for example, blacks at the 50th percentile on the GATB have raw scores of 276; Hispanics, 295; and whites, 305. A score of 305 is at the 84th percentile among blacks. Thus the raw score that would place a candidate at the 50th percentile among whites would place him at the 84th percentile among blacks. Likewise, among candidates for many skilled jobs, a black applicant with a score of 283 would have the same ranking and referral status (70th percentile) as a white applicant with a score of 327.

Faced with a 1986 challenge from the Justice Department (which held that race-norming constitutes unconstitutional reverse discrimination), the Labor Department agreed not to extend its program until the National Research Council of the National Academy of Sciences (NAS) conducted a thorough scientific evaluation of it. The NAS panel recently issued its report, *Fairness in Employment Testing: Validity Generalization, Minority Issues, and the General Aptitude Test Battery*.¹ This report has inadvertently provoked other, even more troublesome questions concerning not the impact of tests on racial groups, but the impact of politics on science. What is the proper role of political views in formulating important “scientific” evaluations? Is science being used as a guise for purely personal political preferences?

The NAS report

In its report, the NAS panel, headed by Yale statistician John Hartigan, confirms that the GATB’s validity for predicting job performance is strong enough to enhance worker productivity; but it recommends that the tests not be used unless the results are race-normed. The panel was unable to justify race-norming on the grounds of racial bias—it confirms that tests like the GATB may even slightly overpredict the performance of blacks, particularly in higher level jobs; in this sense the tests are slightly biased in favor of blacks. Nor did it find that the tests were irrelevant—the panel found a 0.3 correlation between test scores and job performance, which means that employers can realize 30 percent of the gains in workforce productivity that a perfectly valid test would yield. Undeterred, the panel looked elsewhere and found—or at least claimed to find—justification for race-norming in the very nature of a selection system that favors people with high scores over those with low scores.

Any selection system that selects higher scorers over lower scorers necessarily favors those who are above a cut-off rather than those who are below it. However, some candidates who score below a cut-off may in fact be able to perform the job at least at a minimally acceptable level; such candidates are called “false negatives.” Others, “false positives,” score above the cutoff but are nonetheless unable to perform at a minimally acceptable level. The lower a test’s validity, the more such false predictions are likely to occur. Now, more blacks than whites fall in the category of “false negatives,” and more whites than blacks are “false positives.” The reason for these disproportions has nothing to do with race *per se*, but arises from the fact that more whites have scores above the cut-off and more blacks below it. “[T]hese effects,” the panel agrees, “are a function of high and low test scores, not racial or ethnic identity.” Yet even as it emphasizes that the effect is the same for all low-scoring individuals, regardless of their race, the panel nevertheless claims that “the disproportionate impact of selection error provides scientific grounds for the adjustment of minority scores.” It concludes, in other words, that the disproportions in false predictions provide a “scientific justification” for benefiting test takers whose scores—if anything—overpredict their job performance.

The central failing of the NAS panel is that it recommends a race-based solution to what it concedes is not a race-based problem.

¹John A. Hartigan and Alexandra K. Wigdor, eds. National Academy Press. 334 pp. \$20.95 (paper), \$39.95 (cloth).

thus it seeks a scientific justification for what is scientifically unjustifiable. The panel evidently set out to endorse race-norming; to do so, it had to make slanted analytic assumptions, to choose a questionable criterion for determining the tests' fairness, to disregard technical distinctions, and to frame policy considerations one-sidedly.

For example, the panel consistently interprets its data in ways that minimize the GATB's validity. Repeatedly choosing what it acknowledges to be the most "conservative analytic assumptions," it is able to limit the tests' estimated validity to 0.3. (Labor Department technical reports place it at 0.5 or higher.) Time and again the panel then belittles this correlation as "imperfect" or "only modest," despite admitting that it is strong enough to increase productivity. Since the tests' validity poses a stumbling block to any attempt to offer "scientific justifications" for race-norming, the panel is forced to disparage it. The better a test predicts job performance, the more "deceptive" (the panel's word) race-norming becomes.

Even as the panel chooses analytic assumptions that limit the tests' validity, it also sets a standard of perfect validity. The panel shows that the unadjusted GATB scores meet the only widely accepted criterion of fairness in testing—that a test predict equally well for all groups; it then drops that criterion, however, and adopts another—that the test's false predictions not adversely affect minorities—that the unadjusted scores fail to satisfy. In fact, though, any test that meets the commonly accepted criterion must necessarily fail to meet the panel's criterion. (Conversely, any test that meets the panel's criterion must necessarily fail to meet the commonly accepted one.) Only a perfectly valid test could meet both criteria; and perfectly valid tests do not exist. Most experts reject the panel's chosen criterion, viewing it as internally inconsistent; but the panel adopts it without reviewing the scientific literature critical of it or even acknowledging that serious criticism exists.

The panel also draws attention only to the crudest distinctions in job performance, ignoring differences between minimally competent and highly productive workers. For example, when asserting the unfairness of using unnormed test scores, the panel indiscriminately groups together as "able" all workers whose job performance is above a minimum level. This procedure has the effect of exaggerating the significance of false predictions, providing spe-

cious strength to the panel's "scientific justification" for race-norming. Similarly, the panel recommends reporting a second, unnormed score for all candidates already referred on the basis of race-norming; the second score would indicate the probability that a candidate will perform "above average." Although ostensibly meant to reveal intergroup distinctions that race-norming (admittedly) conceals, the unnormed score in fact conceals the full extent of differences in expected job performance because it makes no distinctions among higher levels of performance. Its sole standard is "above average" performance.

Policy considerations

The panel's discussion of policy considerations is similarly one-sided. The panel warns, for example, that the renewed popularity of mental ability tests could lead some groups to feel superior and others to feel inferior; but it says nothing about the feelings that race-norming might produce among those same groups. In discussing race-conscious selection procedures, it cites the opinion of the Lawyers' Committee for Civil Rights Under Law, but not that of any of the group's mainstream or conservative counterparts. It repeatedly warns that the adverse impact of general ability tests makes employers vulnerable to Title VII discrimination suits, but it never points out that its own findings demonstrating the validity of the tests may constitute an employer's defense. It frequently emphasizes the possible advantages of race-normed referrals, but it never clearly assesses the economic costs or legal risks to employers. Although it emphasizes that the problem of false predictions is a function of low scores and not of race, it nonetheless recommends adjusting scores to benefit only minority "false negatives," not white ones.

Further, the panel misconstrues the obvious meaning of the phrase "adversely affect" in Title VII of the 1964 Civil Rights Act, so that it can claim that the title's definition of discrimination centers on unequal results for groups rather than on unequal opportunities for individuals. Contrary to the very words it cites, the panel concludes that "Title VII ... adopts a group-centered definition of discrimination, outlawing 'employment practices' that 'adversely affect' an individual's status as an employee because of that employee's race, color, religion, sex, or national origin." More generally, almost whenever the panel speaks of "equality of opportunity," it really means equality of results for groups.

Likewise, when the panel discusses equal-protection jurisprudence and affirmative action, it relies on the constitutional interpretation of only one scholar, Laurence Tribe. And when it presents the arguments for and against preferential treatment, it inverts the natural order, placing the argument for after the argument against, and then concludes with a section devoted entirely to discrediting the case against preferential treatment.

In that section (entitled "Beyond Philosophy"), the panel, evidently speaking for itself, provides two arguments "to cut through the intellectualization of the issue of preferential consideration for blacks or other disadvantaged minorities." It first contends that preferential treatment has always existed in America, with only the intended beneficiaries having changed. "Very powerful forces support preferential treatment for veterans, including hiring preference in the civil service and referral priority by the U.S. Employment Service." But preferential consideration on the basis of race differs from grateful repayment for service to the nation; more generally, granting people rights merely because of their race differs from grouping together individuals with the same preexisting rights and claims.

The panel's second argument asserts "a skeptical assessment of the liberal values of equality, color-blind law, merit, and fair competition," as seen by those barred from enjoying them until the passage of the 1964 Civil Rights Act. In their view, the panel claims, domination over blacks was sanctioned first by religion (in the colonial period) and then by Social Darwinism (after the Civil War); now it is buttressed by "the myth of equality," which "provides a veneer for further oppression." If the panel is to be believed, just as preferential treatment is nothing novel in America, so too equality of opportunity is only a twentieth-century mask to hide racial oppression.

Politicizing science

Early in its report, the panel observes that "[i]n hindsight it is clear that many of the advocates of early [ability] testing allowed their scientific judgment to be influenced by contemporary racial and ethnic biases and by unexamined assumptions about the social order". The same is no doubt true of the panel itself, whose "scientific judgment" is clearly influenced by policy preferences. Race-norming seems to have been the panel's intended solution from the start.

What the panel in fact shows is not that race-norming is scientifically justified, but that the decision to race-norm is not a scientific issue. The panel confirms that testing's adverse impact is not caused by defects in the tests themselves. Hence the adverse impact has no technical solution, let alone one that requires adjusting minority scores to compensate for "the inadequacies of the technology [of testing]." Nor is race-norming an issue on which scientists are in a special position to make policy recommendations. Instead, race-norming is a political question that ought to be answered politically. Whether the nation should adopt race-norming should be decided by public officials and ultimately by the citizens; to pretend that science can supply an answer is only to pervert science by politicizing it.

Equality or Lasting Inequality?

Jan H. Blits and Linda S. Gottfredson

In 1981, with little publicity and no public debate, USES (the United States Employment Service), a branch of the U.S. Labor Department, recommended that state Employment Service agencies no longer report job candidates' scores on its employment tests in relation to all other candidates, but only in relation to those of the same race. Under this plan, referred to as "race-norming," or (with euphemistic blandness) as "within-group scoring," black applicants are ranked relative only to other blacks, Hispanics only to other Hispanics, and "others" to all but blacks and Hispanics. Rather than being ranked on the basis of a common standard or measure, candidates taking USES's cognitive and psychomotor ability tests are assigned percentile scores according to their standing within a norm or a comparison group of their own race and are then referred to public and private employers on the basis of those percentiles. Employers are often unaware that candidates' scores have been so adjusted. This sleight-of-hand gives bonus points to members of groups that score on average lower than other groups. It offsets average differences in scores among groups and thus eliminates the adverse impact that normally occurs when employment testing is used, as was indeed its sole purpose. Forty states now use race-norming for at least a portion of their job referrals.

In 1986, the U.S. Justice Department challenged the referral plan on the ground that it constitutes illegal and unconstitutional racial discrimination against non-Hispanics and non-blacks, and therefore should be abandoned. Under the pressure of this challenge, the Labor Department agreed not to extend the referral system until a panel of the NAS (National Academy of Sciences) completed "a thorough, scientific evaluation" of it and made a recommendation to the Department. The NAS panel has now completed that study and recommends "the continued use of score adjustments for black and Hispanic applicants

in choosing which applicants one should refer to an employer."

If the Labor Department decides to implement the panel's recommendation, its decision will have widespread effects on employment policy and practice throughout the United States. Because ability tests are generally the best single predictor of job performance, as the panel itself suggests, race-norming the test results for millions of workers will greatly affect the economy of the nation, the composition of its work force, and the relative economic well-being of different groups of job seekers. But even more profound will be the effect on the way Americans think about the basic issues involved in racial equality.

Twenty years ago, Americans would have rejected the notion of race-norming without hesitation. Critics usually attacked employment tests as being unfair precisely because tests were thought to treat blacks and other minorities differently from whites. The assumption was that equality of opportunity for individuals, the traditional American touchstone of fairness, would produce equality of results for groups. Yet, large group differences remain even when tests are racially unbiased and people are treated alike regardless of race, as the NAS points out. Many people – including members of the panel – have been attempting to redefine fairness so that it emphasizes equal group results rather than equality of individual opportunities. As this new emphasis has slowly but surely gained ground, discussions of fairness in employment testing in particular and of racial equality in general have become increasingly muddled.

The NAS panel has added considerably to this confusion. In an Interim Report on race-norming, published in 1988, it did not take an explicit stand on the issue, but rather offered what it called an "impartial" framework within which race-norming should

be considered and judged. Yet, what it offered was not an impartial framework at all, but a quasi-Marxist framework. By ignoring, distorting, and obscuring the most important issues, this framework confused the founding liberal principles of our nation with their exact opposites and supported only one possible outcome in the debate – the one that its Final Report adopts. The Final Report, ostensibly more balanced, is largely concerned with technical issues of testing. Its examination of these issues is shaped by the Interim Report's distorted framework, particularly its conclusion – which becomes the Final Report's tacit premise – that the individual is subordinate to the group.

Our concern here is twofold. First, we are concerned that, despite several recent Supreme Court decisions on civil rights, the NAS's distortions of our nation's fundamental principles will receive the imprimatur of science and will appear to legitimate not only the practice of race-norming, but the questionable justification for it that groups rather than individuals possess rights. Second, we are concerned that, with tragic irony, race-norming will augment and make permanent the very inequality that it is intended to eliminate.

The NAS's Confusion

Like many others in today's jumbled debate, the Interim Report misuses key terms, thereby tacitly redefining them. It constantly speaks of "equality of opportunity" when it really means "equality of results" and of "equity" (with its heavy overtones of fairness and justice) when it means "equality," which may or may not be just. In this way, it takes advantage of the traditional American concern for equity (fairness) through equal opportunities for individuals, using it to promote the very different goal of equality of results across group lines.

The language in which the Interim Report presents alternative perspectives also subtly but effectively elevates some above others. When the Report speaks of classical economic theory, it invariably speaks of it in the past tense, and it always refers to its fundamental principles not as principles but merely as "premises," "assumptions," or "presumptions." By contrast, whenever it mentions the need for economic redistribution, which it often does, it uses the present tense and describes the need not only as a "conclusion," but one of which "many Americans" have become "convinced."

Having redefined equity in terms of group equality, the Interim Report then emphasizes the need

to find a workable balance, or appropriate trade-off, between the conflicting goals of economic efficiency and racial balance ("group equity") in the work-place. From an economic perspective, it argues, employment procedures should contribute to efficiency and productivity. Its section on economics is entitled "The Economic Perspective: Efficiency." Yet, when the Report discusses efficiency, rather than present the argument for it, it suggests that we redefine the term. Even as it states that "In the conceptualization described here and used in this report, economic efficiency or productivity is considered at the level of the individual employer," it criticizes that "conceptualization" precisely because it focuses on that level. Although understanding efficiency at the level of the individual employer "seems an appropriate first line of approach to the question, particularly since the subject of analysis involves employment testing,...time has shown the limitations of the classical economic approach of equating private gain and public good." It offers further criticism of the classical economic approach, on the grounds that:

There could be social costs attached to the very use of ability tests that are masked by the approach taken here. Selection based on test scores implies that those persons with the lowest scores will rarely be referred for jobs. Although the number of people who are unemployed might not change, widespread test-based selection would tend to place the same individuals perpetually at the bottom of the totem pole, creating a class of citizens with little chance of employment. An economic rationale for supporting policies that increase the employment of minorities and others outside the economic mainstream could be made on the grounds of better utilization of human resources or encouraging investment in human capital even in the face of some decrement of efficiency at the level of individual employers.

Since, according to the Interim Report, America's current economic context calls for redistribution rather than the more efficient production of economic goods, economic efficiency or productivity should be redefined in terms of redistribution. Far from posing them as counterweights, as it claims to do, the Report in the end suggests that economic redistribution is really a part of economic efficiency. The trade-off between them becomes no trade-off at all, as "group

equity" becomes an important component of efficiency.

Having already elevated group equality as a component of—indeed a principal standard for—efficiency, the Interim Report goes on to demote efficiency in its traditional sense by belittling the considerable effects of race-norming at the level of the individual employer, first characterizing losses as large as 13 percent in performance as "little," and then, in its Conclusions, as "very little." It also considers efficiency at the employer level in the narrowest possible way. It never considers whether the economic losses might extend beyond the short-term losses in worker productivity to be expected from hiring less qualified minorities—for example, whether obvious dual standards might lead to a progressive and pervasive degradation of morale and of hiring and performance standards among majority and minority employees alike.

Twenty years ago, Americans would have rejected the notion of race-norming without hesitation

But most insidious of all, the Interim Report's framework ignores or denigrates the principles of classical liberalism, grounding its so-called neutral framework on a philosophical theory that is hostile to individual rights. It takes as its starting point not the "self-evident truth" of the Declaration of Independence that all humans are born with certain inherent rights ("life, liberty, and the pursuit of happiness"), but rather "the important insight of early social science" that economic arrangements decisively determine the central values of a society. The Report then suggests that respect for individual freedom is nothing more than a relic of America's nineteenth-century free-market economy. Such respect may have suited America when it was undergoing great economic and geographical expansion. But it can have little or no relevance or validity in the context of contemporary America, the Report suggests, when the central economic issue has become the redistribution, rather than the production, of wealth. That contemporary American economic policy continues to be guided by respect for individual freedom shows not that the nation has retained timeless principles, but that "the nineteenth-

century formulation of capitalism has had—and retains—an important hold on the American mind."

The Interim Report tacitly denies that individual rights are inalienable or irreducible. In place of the liberal principle, upon which the nation was founded, that every person is born with certain inalienable rights which government is established to protect, it offers the Marxist conclusion that individual rights reflect the operations and conditions of the classical free-market system. Rights, according to this view, are derivative, not fundamental. They have no objective or independent existence. They are neither universal nor permanent, but merely an historical phenomenon. Contrary to what the Declaration states, they are not rooted in the nature of man, nor can their protection be the standard for just government. They are only the product of law, not its foundation or standard.

It is no wonder that the Interim Report seldom speaks of rights, and never mentions them in their own name. When it speaks for itself, it speaks of economic interests. When it speaks of rights, it does so in the name of law. ("Rather than encouraging a search for efficient trade-offs, as might have been done if the economic perspective held sway, the courts have adopted an analytical framework that pits an employer's right to make a profit against the rights of minority group members to more equitable treatment.") The Report never mentions a person's right to be hired on his or her own merit, not even when it discusses the law. As they are wholly derivative, individual rights are in no sense inviolable. While the Report cautions, if only ambiguously, that race-conscious employment practices should not "unduly trammel the interests of majority-group job seekers," it never says a word about violating their rights. One cannot accept the Interim Report's framing of the issues surrounding race-norming without at the same time rejecting both the importance of economic efficiency and the primacy of individual rights.

If the Interim Report undermines all serious objections to race-norming by distorting, ignoring, and discrediting them, the Final Report, picking up where the Interim Report leaves off, attempts to justify race-norming on the basis of a specious scientific argument. The NAS panel faced the difficulty that the tests in question are useful in selecting a more productive work-force and they are not biased against minority group members. As the Final Report acknowledges, the tests do not underpredict job performance for blacks; if anything, they slightly favor blacks by overpredicting their performance. As a

result, the panel had to find a scientific justification for race-norming other than the bias or irrelevance of tests. Undaunted, it claims to have found one in the very nature of a selection system that favors people with high scores over those with low scores.

Because such a selection system chooses higher scorers over lower scorers, it necessarily favors those who are above a cut-off rather than those who are below it. But because no test is perfect, some candidates who score somewhat below a cut-off may in fact be capable of performing the job at least at a minimally acceptable level, while others who score somewhat above it may prove incapable of such performance. There will be "false negatives" among the marginal scorers below the cut-off and "false positives" among those above it.

Race-norming will make permanent the very inequality that it is intended to eliminate

More blacks than whites fall in the category of "false negatives" and more whites are "false positives." These disproportions have nothing to do with race per se, but arise from the fact that more whites have scores above the cut-off and more blacks have scores below it. As the panel itself points out, "These effects are a function of high and low test scores, not racial or ethnic identity." Yet, even though it concedes that the tests are unbiased and that the disproportion of blacks among "false negatives" arises from their lower scores and not from racial identity, the panel claims that "the disproportionate impact of selection error provides scientific grounds for the adjustment of minority scores." It concludes that the disproportions in false predictions provide a scientific justification for benefitting test takers whose scores, if anything, already overpredict their job performance. The panel, cloaking itself in the mantle of science, argues for a group-centered "remedy" to an outcome that the panel itself admits is not rooted in group membership, but instead affects all marginal scorers alike, regardless of race. Early in the Final Report, the panel writes that "In hindsight it is clear that many of the advocates of early [ability] testing allowed their scientific judgment to be influenced by contemporary racial and ethnic biases and by unexamined assumptions about the proper social order." The same may be said of the panel itself.

Given the NAS's need to strain for both a theoretical and a scientific justification for race-norming, one wonders why USES recommended race-norming in the first place. The answer lies in the combined effects of years of Executive branch orders and guidelines and subsequent Supreme Court decisions, which, by applying an adverse impact, or a group-centered definition of discrimination, have virtually forced employers to adopt the race-conscious use of tests in order to retain testing itself.

The Pressure to Race-Norm

Beginning in 1971 with the Supreme Court's *Griggs v. Duke Power Co.* decision, adverse impact (the hiring or promotion of proportionately fewer minorities) became prima facie evidence of discrimination. *Griggs* was the first racial case to reach the Court under Title VII of the 1964 Civil Rights Act. Title VII (Equal Employment Opportunity) forbids discriminatory practices in employment, but authorizes employers to use "professionally developed ability tests" in selecting workers, provided the tests are not designed, intended, or used to discriminate because of race, color, religion, sex or national origin. While the testing provision prevented the Supreme Court from declaring that adverse impact alone renders an employment test illegal, the Court interpreted the scope of the provision as restrictively as possible. Paying "great deference" to the EEOC's (Equal Employment Opportunity Commission) 1970 Guidelines for demonstrating the validity of a test, it ruled that the testing provision did not authorize general ability tests, but only tests that measure the applicant's ability to perform a particular job. Any test which adversely affects blacks is illegal under Title VII, the Court announced, unless the employer can prove that it has a manifest relationship to the performance of the particular job in question.

Under *Griggs*, no discriminatory treatment or intent need be shown, since adverse impact alone was deemed sufficient to create a prima facie presumption of illegality. An employment practice can be completely race-neutral or unbiased on its face – as are standardized aptitude tests and educational credentials – and applied to all applicants alike regardless of race, but still held to be discriminatory under *Griggs* if it has adverse impact and is not demonstrated to be specifically job-related.

By adopting the adverse impact standard for discrimination and interpreting Title VII's testing provision so narrowly, *Griggs* turned the judicial focus of discrimination claims in employment from

issues of fairness or justice to technical questions such as the identification of the relevant pool of applicants, the statistical significance of disparate results, and, above all, the validation (the demonstration of job-relatedness) of particular tests.

The *Griggs* legacy has posed two formidable problems for employers. First, the Guidelines are almost impossible to satisfy inasmuch as they represent not a reasonable minimum legal standard for judging test validity, but a scientific ideal to which it was once thought that professionals should strive in developing tests. The Guidelines adopted the APA's (American Psychological Association) 1966 Standards for test development and administration, even though the APA itself warned that those Standards were never intended to be, and were not well-suited to become, rules of law. The Guidelines go beyond the 1985 APA Standards in requiring employers to conduct a separate validation study for each and every job and work-setting in which a particular test is used, as well as for each race or other population subgroup of job applicants. Validation studies of this kind are not only difficult and expensive (frequently costing \$100,000 or more per job), but they are often impossible to conduct in practice. In recent years, some courts have recognized the nearly impossible burden of the now-outmoded Guidelines and have discarded some of their requirements. Yet, while this is an improvement in one sense, this development has at the same time increased the uncertainty of employers concerning the standards to which they might be held because judicial deference to the Guidelines has varied from court to court and from case to case.

An invitation to dinner promises not equality of opportunity but satisfying results

The second formidable problem posed by *Griggs* is that its group-centered definition of discrimination virtually assures that test use will lead to prima facie presumptions of discrimination, since even unbiased and job-related tests tend to have adverse impact. Their use typically makes employers vulnerable to adverse-impact litigation, forcing upon them the exceedingly difficult and expensive burden of proving the job-relatedness of their tests. (Whether the Supreme Court's recent *Wards Cove v. Antonio*

decision will ease this burden, as many expect, remains to be seen.)

Not surprisingly, then, following *Griggs* and its progeny, almost all employment litigation has been decided in favor of the plaintiffs. This judicial climate has caused a dramatic change in the use of employment tests. Some employers abandoned tests altogether rather than go to the expense of conducting test validation studies which, in any case, were no guarantee for successfully countering charges of discrimination. Others, wishing to retain some of the benefits of testing, sought refuge in the so-called "bottom line" defense, which combined tests and equality of end results. This approach typically involved a multi-step procedure in which tests were used in an initial stage to screen out applicants who fell below some minimum cut-off score. The final selection then compensated for the adverse impact of the first stage by giving preferential treatment to the minorities who remained after the initial cut. Overall racial balance – equality at the bottom line – would, it was thought, bar an adverse impact claim and thus preclude the necessity of validating the tests.

In 1982, the Supreme Court rejected the bottom line defense, holding that adverse impact scrutiny should be applied not only at the final selection but at every step of the selection process. That case, *Teal v. Connecticut*, involved a two-step process to select employees for permanent promotion to supervisory positions in the state welfare department. On the first step, 54 percent of the black candidates (26 of 48) and 77 percent of the white candidates (205 of 259) passed a written examination, making the passing rate for blacks 68 percent of that for whites. Yet, on the second step, 11 blacks and 35 whites were promoted, giving blacks an overall passing rate 170 percent of that for whites, as 23 percent of the original 48 black candidates were promoted compared with only 13.5 percent of the original 259 whites. The Court ruled that this was discrimination against blacks. The Court's rejection of the bottom-line defense left employers with little choice but to abandon tests altogether or else to use them in a race-conscious way.

Despite the enormous expense and legal liability now associated with employment testing, many employers are reluctant to abandon testing because of its proven effectiveness in selecting a significantly more productive work-force, especially at the entry-level. By adopting race-norming, employers believe they can realize some of the benefits of testing and at the same time avoid the onerous requirements of the

federal Guidelines. Among the various race-conscious uses of tests, race-norming is particularly attractive to employers because it enables them to select the best workers from each racial group and to avoid the difficult, risky, and expensive burden of proving the validity of their tests. So long as private employers avoid adverse impact at every step of the selection process, which race-norming does, they are free to use whatever selection device (rational or otherwise) they choose.

Lasting Inequality

By eliminating competition across racial lines, race-norming repudiates the language and the legislative history of Title VII and the 1964 Civil Rights Act, which, even the NAS panel concedes, call for color-blind employment practices. More than that, it repudiates the fundamental principles of the American regime and of liberalism itself. With the advent of race-norming, liberalism comes full circle. It returns to the anonymity and the arbitrary inequality of medieval feudalism, which it was intended to replace.

Liberalism, with its fundamental principles of individual freedom and equality, arose in opposition to medieval feudalism's group-centered concept of justice. Consisting in an elaborate structure of distinct and fixed ranks or orders, feudal society was based on the primacy of the group, not of the individual. Which group someone belonged to — serfs, free vassals, clergy, burghers, nobles, or any of the numerous subdivisions within each of these groups — made all the difference in every aspect of life. In contrast to liberalism's emphasis on a common standard of justice for all, under feudalism there were no common rights or common justice. Because the group rather than the individual was the primary unit, a person's rights and disabilities were determined by one's membership in a group. Even courts of law differed from group to group. Serfs were tried under one code of law and before one judicial tribunal; lords another; burghers a third; and clergy a fourth. Criminal punishments not only differed from group to group, but could be imposed on entire groups as such. Whole towns and villages — the guilty and innocent alike — were sometimes collectively punished.

Group membership, which was permanent, was determined solely by birth. Regardless of abilities or talents, individuals were born into the rank they would occupy for the rest of their lives. A man was a serf because he was born a serf; a lord because he was born a lord. Lineage was destiny. Even membership

in a particular guild (such as shoemaking or weaving) often depended on a person's birth.

In feudal society, people enjoyed neither freedom nor equality. The arbitrary conditions of birth limited opportunities, and no one could be the equal of anyone except another member of one's own group. Between members of different ranks or orders there was no common measure (except in the eyes of God), and therefore no possibility of equality.

Liberalism sought to replace the arbitrary conditions of birth with the common nature with which all people are born. It replaces the arbitrary and the fixed with the natural and the free. Because the purpose of government is to protect the rights of individuals, just government rests on the principle that no one should be limited in life by the condition of his or her birth. Just government requires equality of opportunity for individuals. Because everyone is born with the same basic rights, everyone is entitled to be judged on the basis of individual merits, limited only by one's natural abilities, talents, and motivation.

Not content with equal opportunity for individuals, proponents of race-norming demand equal results for groups. Yet what they would actually achieve would be just the opposite of what they seek. Under race-norming, blacks as a group might enjoy parity with whites. By design, race-norming promises racial balance in the work force. Yet, in spite of this, under race-norming, blacks and whites as individuals could never be equal. Race-norming would perpetuate the very inequalities it seeks to remove. Instead of narrowing inequality, it would make equality impossible, for blacks, never competing with whites, could never be compared with them. Because blacks would compete only with other blacks, they could be the equals only of other blacks. Instead of being comparable, members of different races would become different in kind.

Race-norming represents a giant step backwards in the struggle for human dignity and equality. It would restore feudalism's group justice ("group equity," the NAS calls it). Lineage, once again, would become destiny, as a person's rights and standing in society would be determined by the arbitrary conditions of birth.

The New Segregation

The adverse impact approach to racial discrimination, which the Supreme Court first adopted in *Griggs* and which the NAS panel takes for granted, assumes that rights belong to groups as groups rather than to their individual members. Whether a person has been

discriminated against is determined not by reference to the individual, but by reference to his or her group. Membership in the group determines an individual's injuries and claims. On the one hand, individuals in similar circumstances have similar claims only if they belong to the same racial group, and, on the other, no individual can claim to be a victim of discrimination unless he or she can show that others in the group were also victims. It is the group, not the individual, which is thought to bear rights.

The panel suggests that class action suits provide a precedent for the adverse impact approach. Yet the two are the exact opposites of each other. In a class action, individuals become class members by having similar claims. In a disparate impact claim, members are not grouped together because they have similar claims. Rather, they have similar claims because they are grouped together. The group is not an aggregation of individuals, but a fundamental entity in itself.

Race-norming returns us to the anonymity and the arbitrary inequality of medieval feudalism

Adverse-impact justice is depersonalized justice. Just as the harm suffered must be corporate harm, so the responsibility borne is corporate responsibility. According to the traditional notion of justice, an individual should be held responsible only for what the person himself did. Guilt implies causality. No one should be punished for what someone else (or no one) did. Yet, in adverse impact claims, employers can be held responsible for effects which they did not cause and over which they may have had no control. Just as group membership determines rights and injuries, so group statistics determine responsibility and guilt.

With the adoption of the adverse impact approach, the group as group necessarily eclipses the importance of its individual members. And, as discrimination becomes defined in terms of statistical results, the judicial focus shifts from equality of opportunity for individuals to equality of results for groups. This is evident in the way *Griggs* turned Title VII on its head. Whereas Title VII ("Equal Employment Opportunity") is plainly concerned with opportunities, the *Griggs* Court declared that "Congress directed the thrust of the Act to the consequences of employment practices" (the Court's emphasis). And whereas Title

VII is concerned with individuals, not groups ("It shall be...unlawful...to discriminate against any individual...because of such individual's race..."), the Court never mentioned the word "individual" (except in two footnotes when it could not possibly avoid it) and spoke instead only of groups. What Title VII proscribes, it said, is "precisely and only" discriminatory preference "for any group."

If the *Griggs* Court never mentioned individual rights, neither did it support its adoption of the adverse impact approach with any specific references to the language or the legislative history of Title VII, or to any legal precedent or even a legal theory. Instead, it relied entirely on a child's fable. "Congress has now provided," it announced, "that tests or criteria for employment or promotion may not provide equality of opportunity merely in the sense of the fabled offer of milk to the stork and the fox." In the fable, a miserly fox invites his neighbor the stork to dinner, but serves the meal in a shallow plate, preventing the stork, whose long beak is an obstruction, from enjoying it. The fable is cleverly chosen, for it completely obscured the very distinction in question — that between opportunity and results. An invitation to dinner promises not equality of opportunity, but satisfying results. No guest, no matter what his ability, should leave dissatisfied with his portion. The standard for hospitality is not the fairness of the host, but the pleasure of his guest. "Resort[ing] again to the fable," the *Griggs* Court concluded that Congress had "provided that the vessel in which the milk is proffered be one all seekers can use." For, it explained, by enacting Title VII, "Congress has now required that the posture and condition of the job-seeker be taken into account." According to the Court, Title VII, in short, has mandated race-conscious employment practices because it requires equal results. Lest equality of opportunity be "merely" equality of opportunity (as the Court infelicitously put it), employment testing must guarantee equality of results.

One may hope that the new majority on the Supreme Court would prohibit race-norming, for it is no exaggeration to warn that race-norming threatens a return to the infamous "separate but equal" doctrine of *Plessy v. Ferguson* (1896). In that case, the Supreme Court found it no infringement of equality to group individuals by race: "[A] statute which implies merely a legal distinction between the white and colored races...has no tendency to destroy the equality of the two races," the Court declared. Likewise, under race-norming, while blacks and

whites would work side-by-side, they would be judged separately, though guaranteed equal group results. No Court would ever openly return to the phrase "separate but equal," for everyone recognizes that "separate" always meant "unequal." A new euphemism would be needed. But if history is any guide, there is no reason to doubt that one will be forthcoming. "Newspeak" is not new to the rhetoric of racial discrimination. An accurate phrase for what the panel seemingly intends would be "mingled but unequal." Though physically mingled, black and white workers would be judged on unequal bases. For blacks as a group, the result would be "separate but equal," for blacks as individuals, "mingled but unequal." Racial balance among groups of workers might be achieved, but at the price of permanent social inequality among the workers themselves.

READING SUGGESTED BY THE AUTHORS

Blits, Jan H. and Linda S. Gottfredson. "The Political Science of Employment Testing," *The Public Interest*, Winter 1990.

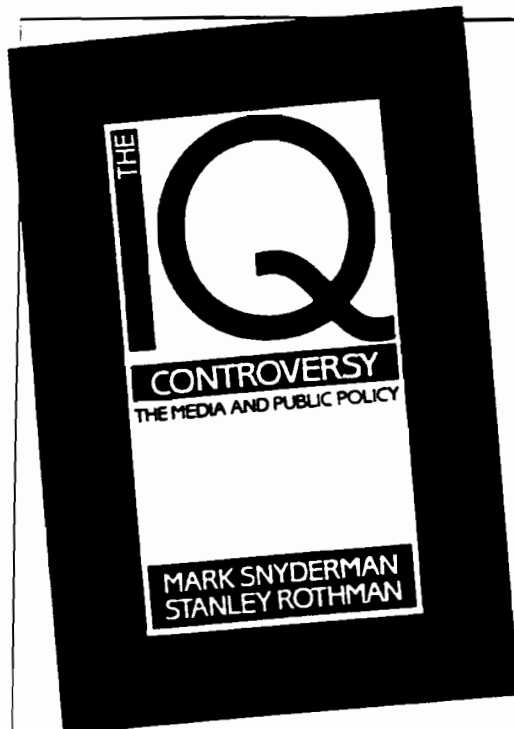
Gottfredson, Linda S. and James C. Sharf (eds.). "Fairness in Employment Testing," *Journal of Vocational Behavior* 33:3 (1988).

Hartigan, John A. and Alexandra K. Wigdor (eds.). *Fairness in Employment Testing: Validity Generalization, Minority Issues, and the General Aptitude Test Battery*. Washington, DC: National Academy Press, 1989.

Hartigan, John A. and Alexandra K. Wigdor (eds.). *Interim Report: Within-Group Scoring of the General Aptitude Test Battery*. Washington, DC: National Academy Press, 1988.

Jan H. Blits teaches in the University Honors Program at the University of Delaware and was previously Secretary of the Navy Distinguished Fellow at the U.S. Naval Academy. He has published two books and numerous articles on educational and political philosophy.

Linda S. Gottfredson is associate professor of educational studies at the University of Delaware and co-director of the Project for the Study of Intelligence and Society. She has published in personnel psychology, counseling psychology, and sociology on issues of testing, inequality, and fairness in education and employment.



"A major purpose of *The IQ Controversy* is to talk back to the media about intelligence testing..."

"An interesting study of important problems: differing judgments regarding the development of intelligence, and the conflict between expert and public views on the significance of intelligence testing.... Well-written and well-researched, this first-rate title is an illuminating piece of scholarship."

—K.F. Widaman, *Choice*

"Shed(s) important new light on the role of the media in shaping public opinion on the subject of intelligence testing."

—Terry Teachout, *National Review*

"The book's survey of the experts is a major event."

—Daniel Seligman, *Commentary*

ISBN: 0-88738-151-0 (cloth) 192 pp \$24.95
ISBN: 0-88738-839-6 (paper) \$18.95

At your bookstore or direct from the publisher.
Major credit cards accepted. Call (201) 932-2280



transaction publishers

Department IQ
Rutgers-The State University
New Brunswick, N.J. 08903

transaction

The Case for Fairness

Alexandra K. Wigdor and John A. Hartigan

There is little to be gained from a point-by-point rebuttal of the inaccuracies and distortions that Blits and Gottfredson have visited upon the work of the National Research Council's Committee on the General Aptitude Test Battery. They refer so little to the Committee's Final Report that our time is better devoted to describing our study and the policy recommendations that grew out of it.

There is a difference in perspective, or better said, world view, between the Blits/Gottfredson piece and the work of the Committee that deserves mention at the outset. The Blits/Gottfredson article posits a world of artless simplicity: individualism is all; groups are illegitimate categories; social policy that confers benefits on groups is not only bad, but Marxist (farmers, home mortgage holders, and veterans beware!). Its authors do not admit that the issue of allocating employment opportunity equitably is genuinely perplexing – just as they see no reason to worry about discrimination against groups or about the effects of past discrimination.

Our report recognizes the very real complexities inherent in realizing equality of opportunity, particularly in a society that until 1964 – a mere 25 years ago – operated under a legally sanctioned and pervasive system of discrimination against blacks, and on a regional basis, certain other groups of citizens. We present in some detail both sides in the public debate on affirmative action and preferential treatment. We make it clear that there are several possible definitions of fairness and that they conflict. In discussing the economic and social implications of using test scores to make employment decisions, we are careful to demonstrate explicitly that the statistics showing the job prospects and likelihood of job success of black and white workers tell many different stories to those whose minds are open.

The most important difference between the work of Blits and Gottfredson and that of the Committee is that we recognized that social groups of many sorts are

relevant categories, both for understanding how America has operated in the past and for creating public policy in the present. Blits and Gottfredson seem unaware of (or unwilling to admit) the fact that in the United States people have been discriminated against because of group membership. Blits and Gottfredson have many complaints about the Supreme Court's decision in *Griggs v. Duke Power* (1971), but demonstrate no appreciation of the fact that until July 2, 1965 (the effective date of the Civil Rights Act of 1964), the Duke Power Company (among many others) openly discriminated on the basis of race.

At the company's Dan River Steam Station, blacks were permitted to work in only one of the five operating departments (the Labor Department), which meant exclusively in menial jobs performed out of doors. The highest salary paid to workers in this department was lower than the lowest salary in any of the other departments – Coal Handling, Operations, Maintenance, and Laboratory and Test. With a rather extraordinary sense of timing, Duke Power instituted a high school diploma requirement for transfer from Labor to any other department on July 3, 1965. The company also introduced an intelligence test requirement for new hires in all but the Labor Department. The effect of these requirements was to continue the segregationist policy that had been in place until July 2, an outcome that the Supreme Court found untenable.

A generation before in *Brown v. Board of Education* (1954), the first landmark civil rights case, the Supreme Court recognized that racial discrimination works not only against individuals, but against a people. The remedy ordered in that case was not just that a particular black child named Brown be allowed to attend a particular white school. The Court ruled that a state-run segregated school system is per se a denial of the equal protection of the laws to all black pupils and ordered that the dual system of schools be dismantled wherever it existed.

In the *Griggs* case, the Court was faced not with the overt, state-imposed discrimination of *Brown*, but with a no less severe de facto segregation of black workers in the least desirable jobs. To its credit, the Court recognized that this case was not a simple problem involving an individual employer with possibly invidious intentions, but a systemic problem – the fruits of segregated and inferior schooling and the operation of barriers erected in the past to favor white employees over black. Rather than get caught up in the quagmire of motivation or intent, the *Griggs* Court focused on the operation of an employer's procedures. Unless they can be shown to be related to job performance, the Court ruled, "practices, procedures, or tests neutral on their face, and even neutral in intent, cannot be maintained if they operate to 'freeze' the status quo of prior discriminatory employment practices." The judgment about discrimination would henceforth take place at the bottom line – in the composition of the workforce.

In its first interpretation of the equal employment opportunity clauses of the Civil Rights Act, the Supreme Court recognized what *Blits* and *Gottfredson* apparently do not yet discern: that a whole edifice of law and custom existed up to 1964 to prefer whites as a group over blacks in American society.

If *Blits* and *Gottfredson* are not attuned to the deeply rooted, systemic discrimination against blacks with which the Court was trying to grapple in these cases, they seem similarly unaware that the Civil Rights Act of 1964 is not the first law to be passed that extends special protection to particular social groups. Their attempt to reduce the liberal tradition to an Ayn Randish brand of unfettered individualism simply does not jibe with the realities of American history. Beginning with the Brandeis Brief of 1908, which argued successfully for limiting to ten hours per day the work of women in the laundry industry (*Muller v. Oregon*), group analysis has been an important engine of social reform. The laws regulating the conditions of labor for women and prohibiting child labor are analogous attempts on the part of government to protect groups at risk by constraining the freedom of the employer to define the conditions of labor. Louis Brandeis understood, even if *Blits* and *Gottfredson* do not, that the precepts of liberalism are simply that – precepts, not the Ten Commandments.

Background of the FAIRNESS Report

The National Research Council study of the General Aptitude Test Battery was commissioned by the DOL (Department of Labor). DOL was seeking advice on the use of the test in the Public Employment Service,

a cooperative federal/state program begun during the Depression to help employers find workers and workers find jobs. Nowadays, more than 19 million people pass through the 1,800 local offices of the Public Employment Service annually. About 3.5 million of them are actually placed in jobs.

In order to better serve employers, the USES (United States Employment Services), the federal partner in the Employment Service system, began in 1980 to encourage the states to try out a new referral system that uses the GATB (General Aptitude Test Battery) to screen virtually all job seekers. If the plan to use a general aptitude test to refer job seekers to all kinds of jobs is found legally acceptable under the civil rights laws, it is likely to attract large numbers of employers to the Employment Service who seek the security of hiring on the basis of a federally sponsored employment test.

The fact of the matter is that employers are in a bind. Declining American competitiveness in world markets has renewed their interest in employment testing as a means of hiring more proficient workers. Yet Federal equal employment opportunity policy has made employers vulnerable to legal challenge if their selection procedures have an adverse impact on minorities and women or, with the advent in recent years of reverse discrimination suits, if they take affirmative measures to assure that minorities and women are not screened out. Caught between the twin threats of discrimination and reverse-discrimination suits, many employers have turned away from objective selection procedures to avoid costly litigation. What employers need, and what the U.S. Employment Service hoped to provide with its test-based referral system, are the tools to pursue both economic efficiency and fairness.

The Validity Generalization-GATB Referral System

The GATB is not a new test. It was developed in the 1940s and has been used for vocational counseling and for some job referrals since then. What is new is validity generalization, a theory and mode of analysis that is in the process of overturning the traditional view that the predictive validity of a test must be established for every new use. By showing test validities to be much more uniform across studies than was previously thought, validity generalization analysis supports the idea that if the validity of a test is known for a reasonable sample of jobs (the predictive power of the GATB has been studied for more than 500 different jobs), then its validity for other similar jobs will fall within a relatively small range.

In practical terms, validity generalization analysis gives researchers a way to estimate, on the basis of existing research studies, what the predictive validity of a test would be for jobs not studied.

On the basis of validity generalization analysis of the GATB data base, USES staff concluded that the GATB is a valid predictor of job performance for all 12,000 jobs found in the U.S. economy – everything from steam fitter to school teacher to accountant. In 1980, USES began promoting the GATB as the centerpiece of a new referral system that would allow more exact assessments of Employment Service clients than the existing system based on brief interviews. The intent was to fill job orders by referring clients who otherwise meet employers' skill or experience requirements in order of test score. DOL officials felt that this test-based referral system would offer employers the maximum productivity gains from improved employee selection.

As one of the major implementing authorities of the Civil Rights Act of 1964, the Department of Labor was also sensitive to its obligation to avoid a referral system that would have an exclusionary effect on minority clients. Strict top-down referral would mean that most black and Hispanic job seekers would not be referred, since these groups score on average substantially below the majority. To avoid that outcome, U.S. Employment Service staff proposed that each registrant's score be computed as a percentile score within his or her own group (defined as black, Hispanic, and other). Black, Hispanic, and other clients at the 70th percentile would all be referred at once, even though their raw scores were very different.

As the Committee's Interim Report demonstrated, this solution kept much of the benefit of top-down selection for the employer, while at the same time avoiding the extreme exclusionary effects that testing would otherwise cause. Many employers, particularly large employers of skilled and semi-skilled labor who can afford to be selective, were enthusiastic about the validity generalization referral system.

The DOJ (Department of Justice) was not. In 1986, Wm. Bradford Reynolds, Assistant Attorney General for Civil Rights, challenged the validity generalization-GATB system, saying that the within-group scoring strategy illegally and unconstitutionally advances the interests of some groups (black and Hispanics) at the expense of other groups. An Assistant Attorney General for Civil Rights faced with a validity generalization-GATB system that did not make score adjustments would doubtless have challenged it as well

because of the degree of adverse impact it would have on black and Hispanic job seekers.

Because of the extreme complexity of the issues, the two agencies agreed to maintain the status quo – DOJ would not take legal action and DOL would not expand the validity generalization-GATB system beyond its 1986 boundaries – until a study was conducted by a group of outside experts. The National Academy of Sciences agreed to take on the task. Through its working arm, the National Research Council, a committee of experts conducted a two-year study of the General Aptitude Test Battery.

Technical Findings

The Committee's report presents a thorough critique of the technical quality of the GATB, and on the basis of these findings, offers a number of policy recommendations for the fair and reasonable use of the GATB in Employment Service job referrals.

The Committee found that the GATB is a general ability test of reasonable quality, roughly comparable in technical quality to other big testing programs such as the Armed Services Vocational Aptitude Battery, although it is not supported by as strong a research and development program as the latter instrument. In terms of predictive power, GATB scores have a meaningful, though only modest, relationship to performance on the job, again roughly comparable to other general employment tests.

An important point, both theoretically and in terms of testing practices, is the Committee's support of validity generalization and its applicability to the GATB data base. Based on its reanalysis of some 755 validity studies involving more than 70,000 workers, the Committee estimated that the "true validity" of the GATB for the kinds of jobs typically handled by the Employment Service ranges from .2 to .4 with an average value of .3. If prediction from GATB scores were perfectly accurate (validity = 1.0), then 100 percent of the people who ranked in the top half on the test would also rank in the top half in terms of their job performance. With a validity of .3, only 60 percent of the people who ranked in the top half of the test would be expected to rank in the top half on job performance. Conversely, 40 percent of those in the bottom half of the distribution of test scores would rank in the top half of job performance.

In other words, there is a relationship between how well job candidates do on the test and how well they do on the job – but there is also a lot of error in the prediction of job performance from GATB scores.

The critical question for the Department of Labor to decide is whether the GATB is a good enough test to play an important role in the way the Employment Service refers people to jobs. In the Committee's judgment, the answer is a qualified yes. Even modest validities such as these translate into performance gains for the individual employer, whereas interviews, biographical inventories, and other alternative procedures provide less accurate prediction on the whole.

Yet the GATB is not good enough to be the sole means of referring candidates. The report recommends that multiple criteria (e.g., educational requirements or years of experience as stipulated by the employer) be used along with GATB scores. We have also urged that every local office provide alternate paths to job opportunities through the Employment Service. Some people have special skills that are not well measured by the GATB. Some people simply do not test well. Some groups and communities are very resistant to testing. Older workers might justifiably feel that their employment record is a far better indicator of their qualifications. The meaning of GATB scores for people with certain kinds of handicapping conditions is not known.

The economic gains from using an instrument of .3 validity are not sufficient, in the Committee's judgment, to justify its widespread use without some form of score adjustments, given the enormous adverse impact it would have on black and Hispanic job seekers.

The Distinction Between Referral and Hiring

Before describing the Committee's position on score adjustments, a word is necessary about the way the Employment Service's test-based referral system works. A substantial portion of the Blits/Gottfredson argument is based on the presumption that GATB scores serve as the sole basis for hiring decisions. This is simply not the case. The validity generalization-GATB system is a referral system. It functions as an initial screen, determining which clients will have the opportunity to compete for positions with a given employer. From among those registrants who meet the requirements imposed by the employer, the Employment Service local office will build up a referral pool in order of test scores, sending, say, 10 or 15 clients to an employer who has five openings.

The employer is free to consider a variety of applicant characteristics other than the GATB scores: past work history, regularity of attendance, ability to work as part of a team, possession of specialized skills and knowledge, and other attributes that may be relevant to a particular job, yet are untapped by the

GATB. Studies examined by this Committee suggest that employers do not rely heavily on GATB scores in deciding which referred applicants to hire.

The mundane facts of the way the referral system really operates bear little relationship to the impassioned argument made by Blits and Gottfredson that under "race-norming" (which, by the way, is not what the Committee recommends), "blacks and whites as individuals could never be equal....for blacks, never competing with whites, could never be compared with them. Because blacks would compete only with other blacks, they could be the equals only of other blacks. Without a common measure – without a common basis for comparison, they would be literally incommensurable with everyone else." In operational terms, there is direct competition among all of those, black and white, in the pool of applicants to be referred.

Equality Under the Law and the Fair Use of Tests

We have, as the next section explains, proposed a "performance-fair" system of score adjustments for use in building the referral pool that provides what Blits and Gottfredson call for: a common measure, a common basis for comparison – one that is directly linked to job performance. In such a system, the test scores would be so adjusted that the distributions of scores *at a given level of job performance* is the same for blacks, Hispanics, and whites. (And in the interest of full information, we have also recommended that both an adjusted and an unadjusted score be reported to employer and applicant alike.)

Until now, policy debate on the issue of employment testing has focused on test scores and on fairness as a matter of the equation used to predict job performance from test scores. That was Mr. Reynold's orientation when he found the within-group score adjustments a denial of the constitutional right to equal protection of the laws. It has also been the point of departure for most psychologists and other testing professionals.

The Committee's report on the GATB urges Federal officials to shift the grounds of discussion from test scores to the actual object of interest – realized job performance. At the heart of a performance-fair system is the following proposition: *Employers should place equal value on black and white workers who do equally well on the job.*

The earlier discussion has established that the GATB, like other employment tests, is a very imperfect indicator of how well people will do on the job. One consequence of imperfect prediction is that some

people who score low on the test would do well on the job. When applicants are referred in order of test score, such people will be subject to "false rejections," – they will not get a chance at the job even though they would have done well, in fact, better, than many of those who were referred. The converse is also true: some people with high scores will be "falsely accepted," that is, they will turn out to be unsatisfactory workers.

All test takers are at risk. In every group there will be some people who could perform successfully on the job but whose low scores will keep them from being referred. There will also be some people whose high scores lead them to be referred and perhaps hired, but who nevertheless perform poorly on the job.

But – and this is the crucial point – the interplay of prediction error and group differences in average test score means that far larger proportions of blacks and Hispanics than of majority candidates will be subject to false rejections. And far larger proportions of majority candidates will likewise get the benefit of false acceptances. In practical terms, this means that if unadjusted GATB scores are used to refer job seekers, black and Hispanic Employment Service clients will be disproportionately subject to false rejections. In the terminology of the *Griggs* case, use of the test scores without adjustments would erect an "artificial barrier" to their employment chances. The Committee felt strongly that the inadequacies of testing technology ought not to fall disproportionately on those already burdened with the effects of past and present discrimination.

The Committee has recommended that GATB scores be adjusted – not to create proportionality as Blits and Gottfredson have implied by erroneously describing our system as "race-norming," but to equalize the chance of referral for whites, blacks, and Hispanics who would do equally well on the job. The size of the appropriate adjustment can be derived from the degree of validity of the test and the size of the gap in average score between the groups. If test validities rise, the size of the adjustments would be reduced; if group differences in average score disappear, so would the adjustment.

It is true that this sort of performance-fair referral system is likely to result in some economic loss. The Committee estimates that between 10 and 20 percent of the gain in worker efficiency due to using the GATB could be lost, because the average job performance of the black workers referred will be lower than the average performance of the white workers referred. (To carry on the earlier illustration, the proportion of false positives will be equalized along with the propor-

tion of false negatives.) This is the cost of giving able black workers the same chance of referral as able white workers.

The Committee's response to the problem of allocating employment opportunity fairly is contingent, not absolute. If there are dramatic advances in the predictive accuracy of employment tests, the use of employment tests scored by performance-fair methods could result in a referral system with legally or politically unacceptable levels of adverse impact. But in the present state of the art, performance-fair score adjustments make sense.

If no score adjustments are made, we know in advance that black and Hispanic workers who would be rated by their supervisor as the equal of their white colleagues, will not in fact be referred to employers in a test-based referral system. By adjusting the scores of blacks and Hispanics upward so that people from all three groups who would be rated equally good workers have the same distribution of test scores, the Public Employment Service can assure that employers are in a position to place equal value on black, white, and Hispanic workers who do equally well on the job. Put another way, given the appropriate score adjustments, individuals from all three groups will be equally at risk of false rejections due to the technical weaknesses of a flawed, but economically attractive assessment device. This seems a definition of equality and an approach to fairness that adheres to the spirit of Federal equal employment opportunity law.

READING SUGGESTED BY THE AUTHORS

Hartigan, J.A. and A.K. Wigdor, (eds.). *Fairness in Employment Testing: Validity Generalization, Minority Issues, and the General Aptitude Test Battery*. Washington, DC: National Academy Press, 1989.

Alexandra K. Wigdor is study director of the Committee on the General Aptitude Test Battery, and also serves as study director of the Committee on the Performance of Military Personnel. She has co-edited (with Wendell R. Garner) *Ability Testing: Uses, Consequences, and Controversies*.

John A. Hartigan is chairman of the Committee on the General Aptitude Test Battery, as well as the Eugene Higgins professor of statistics and director of the Statistical Computing Laboratory at Yale University. His teaching and research interests center on the foundations of probability and statistics, Bayes theory, classification, statistical computing, and graphical methods.

Fairness in Employment Testing

Mary L. Tenopyr

The timing of *Fairness in Employment Testing* is regrettable. Appearing just before the *Wards Cove v. Atonio* decision (1989) and the other major civil rights decisions of this U.S. Supreme Court term, it will not have the practical impact it otherwise might have had. Nevertheless, the report may play a part in the formulation of legislation designed to mitigate the effects of the recent Supreme Court decisions.

The *Wards Cove* decision clarified the burdens of proof in an employment case so that when a facially neutral selection procedure such as a test is used, the first step is that the plaintiff must show an adverse effect on his or her group. The burden of producing evidence, not persuasion, then shifts to the defendant, who must, as a second step, either refute the plaintiff's statistics or articulate a legitimate business reason for the employment decision. As a third step, the plaintiff bears the burden of disproving the legitimate business reason given by the defendant or proving the availability of an alternative with less adverse impact and equally serving of the defendant's business needs.

As the employer has only to produce a legitimate business reason for a contested employment decision, validation in the strictest sense is probably no longer required in all Title VII cases. But there are contrary indications. The record-keeping requirements of the *Uniform Guidelines on Employee Selection Procedures* (1985) are endorsed in the *Wards Cove* decision. These same record-keeping requirements are also being used now in Office of Federal Contract Compliance and Programs reviews of employers.

The legal status of validity research with respect to testing is in question. Having done thorough validity research is the surest defense. The fact that the *Uniform Guidelines* are incorporated, by reference, in the *Wards Cove* decision, may bring forth plaintiffs' efforts to enforce the whole of the *Guidelines*, not just the record-keeping procedures. The inevitable new civil rights legislation that will be introduced in the U.S. Congress can be expected to emphasize valida-

tion. The validity of the GATB (General Aptitude Test Battery) may or may not be a legal issue in the future. From ethical and economic viewpoints, the validity of every selection device should continue to be a matter of concern.

There is a major problem in examining race-based norming, because it was introduced simultaneously with validity generalization and the use of race-conscious norms have been hopelessly confused in many people's minds. The two are separate and should be treated as distinct in any discussion of the report.

My general impression of the report is that it is thorough, but leaves major questions in a number of areas. The fact that *Fairness in Employment Testing* deals with such sensitive and controversial political, legal, cultural, and scientific issues made it a difficult report to prepare. The authors have had an arduous task. Also, since the research base for the GATB needs improvement, there is more speculation than one would like in the report. At times, figures and tables are not so clear as one would wish.

The perspective of the report ranges from even-handed balance on some topics to what appears to be some slanting on other topics. There seems to have been some selective citing of sources relative to the legal issues involved. The reader would have benefited from more inclusion of the perspectives of more conservative attorneys.

Inconsistent Definitions

There are also some contradictions and omissions that should be recognized relative to some of the conclusions. An example is the failure to review thoroughly the literature criticizing the definition of discrimination which is the basis of the GATB race-norming. The definition has been identified with Cole and criticized severely by Petersen and Novick, Hunter and Schmidt, and others, and has been found to be internally inconsistent. This fact again was recently pointed out by Humphreys. The definition is

based only on two types of workers: "good" and "poor," and assumes all workers in the "good" category have the same level of job performance. The emphasis is on attempting to reduce racial differences in the number of "good" workers selected. The definition does not adequately consider the consequences of selecting "poor" workers, pointed out by Humphreys. It has also been noted that apparent discrimination will be found with this definition any time prediction is not perfect. Even with perfect validity, apparent discrimination will be found when the predictor and the criterion are not dichotomized (broken in half) at the same relative point.

Employers do not know how job applicants are being referred to them

The Cole definition was not among those supported by the National Council on Education, the American Educational Research Association, and the American Educational Association in *Standards for Educational and Psychological Testing*. The only method for determining testing fairness endorsed by these organizations uses the regression model or variants thereof. This position is also reflected in *Principles for the Validation and Use of Personnel Selection Procedures* prepared by the Society for Industrial and Organizational Psychology. What is perplexing to me is that the National Research Council Committee carefully determined the degree of fairness of the GATB by the regression model and then, later in the report, recommended that the test be used at least partially on the basis of the Cole model.

False Negatives

Another contradiction is the fact that the report clearly admits that the false negatives, people who could perform the job successfully yet who fail the test, are not a function of race. Yet the Committee goes on to recommend race-conscious decision making.

Another anomaly is that relief is sought only for the false negatives who are minorities. Simply because minorities are just that — "minorities," — there will usually be numerically more false negatives who belong to the majority group than who belong to a minority group. The exact ratio of majority to minority false negatives depends on the situation, but it can easily be shown that, in the test score and performance

ranges typically found in an employment situation, the majority of the false negatives will be members of the majority group. By what logic does one give relief to a low-scoring minority and not give relief to an equally low-scoring majority group member?

There is another anomaly in basing conclusions about race-conscious norming on a model based only on dichotomous data, when on a continuous basis, persons scoring at each successive higher score level would be predicted on the basis of a valid test to have successively higher average levels of performance. This fact is obscured. Also, with a model based on a dichotomy, there is a different apparent degree of discrimination every time the passing score is changed.

Rationale for Criteria

A further criticism of the study involves what I consider serious omissions. Repeatedly it is said in the report that the validities for the GATB are "modest." Yet there is no serious discussion of the criteria involved. It is clear that supervisory ratings are usually the criteria and that many of these ratings were based on a single form that has been used in common for many jobs. There should be a discussion of how this rating form was developed and how it was established that the criterion itself was valid for so many different jobs with different job duties. I recognize the infinite regress problems implied here, but a thorough discussion of the rationale for criteria involved would seem mandatory. Perhaps the "modest" validities reflect a criterion problem, not a test problem.

There are also questions about the effects of combining data from various employers in one sample. Some more studies involving the correlations within employer sample groups would be useful. Another omission is the discussion of utility. There is much coverage of the effects of various referral methods on the percentage of minorities selected, yet there is no refined discussion of the economic impacts of various referral methods. Granted that there will always be subjective judgments regarding usefulness of any selection procedure, the report should have contained more data to inform the reader what dollar usefulness might be implied by an increase in validity.

Another area of concern is the Committee's apparent failure to get out in the trenches to obtain which may have offered an explanation for some findings. The decline in validities over the years may have an explanation found only at the lower levels of the Public Employment Service. I was told that in the

1960s, there was competition among various states to "find" validity. One can only surmise how much inflation of validities, through design or unconscious action, may have resulted from this competition. A mystery to me was why the Employment Services found such high validities for the manual dexterity tests in the GATB when no one else seemed to find dexterity tests very useful.

The apparent speededness of the tests may be simply a result of using tests designed for a full-service employment agency which once served many well-educated people (for whom the tests would not have been so markedly speeded) when in fact the applicant population had changed to one largely at the lower end of the economic spectrum. These lower ability applicants could be expected to answer fewer questions within the time limits.

Lack of Full Disclosure

Reports that local employment agencies are offering employers groups of referrals "representative of the local available labor pool" highlight the need for information from the lower levels of the Public Employment System. No mention is apparently made of race-conscious referral, and employers do not know how job applicants are being referred to them. Anyone who has tried to explain even a simple expectancy table to employment personnel working for a small employer realizes that these people are not going to understand readily the within-group referral system, let alone the compromise plan alluded to by the Committee. I believe in full, honest disclosure when it comes to informing the employer, and I do not think that it is being done now in the field.

Race-conscious referrals should be made only with the employer's informed consent

I am also concerned about the employer who applies further screening of Public Employment Service referrals. Assuming the employer's screening methods, be they tests, interviews, or other methods, have at least some validity and are to an extent correlated with the GATB, the employer is going to encounter much greater adverse effect on minorities from use of his/her internal procedures than otherwise would be the case. In other words, the employer's own selection proce-

dures will screen out a disproportionate percentage of the minorities referred under the lower race-conscious GATB standards.

Under these circumstances, is it in any way appropriate for the Public Employment Service to shift the burden of adverse effect to the employer who may face costly litigation? The race-conscious referrals should be made only with the employer's informed consent. Certainly the employer may well be placed in the untenable position of either adopting what may be an illegal quota or screening beyond the GATB and encountering great adverse impact.

Recommendations

There are 19 central recommendations in the National Research Council report. I agree with many of these and will mention them and also the ones with which I have problems briefly.

- Recommendation 1: *That there be research and development to correct problems with test security and speededness.* Test security should be addressed. The data the Committee presented about speededness (the extent to which examinees can finish the test within the time limits) are cause for some concern. The report shows effective lengths of power portions of the tests — for one test 90 percent of examinees answer 18 questions. Yet for this same test, the average score is 20, two more than the so-called effective length of the test, which is 18 questions. There is a positively skewed distribution, and the data presented by the Committee do not present so conclusive a picture as one might hope. The report should not have been so quick to decry speededness. For many jobs, speed in certain activities may be criterion-related. More research is in order.

- Recommendation 2: *That no job seeker should be obliged to take the GATB, that alternate paths to referral should be used.* This is fine if the alternate methods are valid and the employer is informed that he or she is getting employees who have not been uniformly tested.

- Recommendation 3: *That multiple criteria worked out with the employer be used for referral.* This, again, is appropriate if valid alternatives are available.

- Recommendation 4: *That race-conscious referral be continued.* Efforts to aid minority job applicants are laudable, but employers have an obligation to increase competitiveness and productivity. This is not a scientific controversy; it is a question of goals and values, and not in the least sense, the law. The key here is full, honest disclosure

and informed consent of any employer who obtains referrals through a race-conscious system, or any other referral method. The employer must be informed that the imprimatur of "science" is not appropriate for any particular way of using test scores.

- Recommendation 5: *That better norms should be obtained if within-group referral is continued.* This, I do not dispute.

- Recommendations 6 and 7: *That feasibility of reporting two types of information to the employer be considered.* Providing performance expectancy data as well as within-group centiles may seem like a Solomon-like recommendation, but it may only confuse employers.

- Recommendation 8: *That claims for the validity of the GATB be tempered.* I am somewhat uncomfortable about the use of the term "modest" with reference to GATB validities. Although extreme corrections for restriction in range may not be warranted, the report should accentuate some indication that validities would have been higher had even a modest correction been applied. The decision whether to correct for restriction in range should be noted as a subjective one. The question of relevance of the criteria used in the various studies needs further explication before the GATB validities are deemed only "modest."

- Recommendation 9: *That the Employment Service refrain from cost estimates.* Given the peculiar circumstances of the GATB, I agree with this conclusion.

- Recommendation 10: *That it should be made clear that employers are ultimately responsible for selection.* I agree, however, the principle of full, honest disclosure applies here.

- Recommendations 11, 12, and 13: *That there be more research on the GATB referral system.* This must be carefully planned (not cookbook) research with generic rating forms and the like. Massive, but poor, research is not better than none.

- Recommendations 14 through 19 refer to special consideration for veterans and the handicapped. I find little to disagree with in any of them.

There are many portions of this report with which I agree, but I have considerable discomfort with some of the major recommendations. This report and some others I have read recently are apparently attempting to extend the purview of science and psychometric theory far beyond the bounds of generally understood usage of these terms. One might debate for years about the role of values in science. One might ask whether any science can be meaningful without some sense of values. Also one might ask at what point value-laden considerations distort the true role of science in systematic investigation and honest discourse thereon.

The greatest worth of the report *Fairness in Employment Testing* will be that it will generate a much needed renewed debate about the roles of values in science and the responsibilities of scientific organizations in defining those roles and promoting a proper place for science within the social, political, economic, and ethical systems within which it must exist.

READINGS SUGGESTED BY THE AUTHOR

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington DC: American Psychological Association, 1985.

Equal Employment Opportunity Commission, Department of Labor, Department of Justice, Office of Personnel Management. *Uniform Guidelines on Employee Selection Procedures*. CFR Part 1607 (1985).

Humphreys, L. Letter to the Editor, *Science*, 245 (1989).
Hunter, J.E. and F.L. Schmidt. *Psychological Bulletin* 83 (1976).

Petersen, N.S. and M.R. Novick. *Journal of Educational Measurement* 13 (1976).

Society for Industrial and Organizational Psychology, Inc. *Principles for the Validation and Use of Personnel Selection Procedures*. Third Edition. College Park, MD: author, 1987.

Mary L. Tenopir has directed AT&T's selection and testing program for the past 17 years. She has been employed in various academic, governmental, and business settings. She was a member of the National Research Council's Committee on Ability Testing, and has published widely in measurement and industrial psychology.


Diet and Cancer Risk

**YOU CAN CHANGE
THE ODDS**

Dietary Fat

Too much fat in your diet?
That's a health risk . . . that's a cancer risk.
Learn more. Write:

American Institute for Cancer Research
Dept. FC2
Washington, D.C. 20069



American
Institute for
Cancer
Research

The Problem of False Negatives

Mark Kelman

It is sorely tempting for someone unsympathetic to the substantive positions that Blits and Gottfredson take on the race-norming issue to focus on the piece's worst flaws, regardless of their ultimate relevance to the underlying debate. It is more than a bit frightening to think that university professors so uneducated about basic legal and political theory feel comfortable writing about it. It is scarier still to think that they are sharing their wild misperceptions with a whole generation of students.

The authors' notion that Marx, whatever the manifold errors in his analysis, was some sort of corporatist, is a notion that no person who had even the most cursory familiarity with Marxist theory could maintain. Not only was Marx not wedded to the idea of group rights, he was a virulent anti-nationalist who romanticized the proletariat precisely because he believed their historical role was to destroy the last group cleavage in civil society.

Unfamiliarity with Marx, though, is so commonplace as to be forgivable. It is not forgivable to believe that opposition to a natural rights viewpoint is "a Marxist conclusion." On the contrary, positivism/consequentialism has dominated mainstream Anglo-American jurisprudence from Austin to Bentham to H.L.A. Hart to Chief Justice Rehnquist. One hardly needs to be a Marxist to believe the strong (Austinian) positivist message that rights are simply a reflection of sovereign decisions. One need not be the least bit radical to accept the weaker, less morally relativistic positivist position which argues that collectivities protect particular rights for reasons. There are, in this view, no natural, preexisting determinate rights; each claim of right must be justified in terms of the social consequences of recognizing the validity of the claim. When a conservative Justice Rehnquist rules in the *PruneYard* case that a shopping center owner need not be compensated when the state of California mandates that orderly pickets be permitted on the owner's property, he is simply recognizing that the "right to

exclude" is not naturally preordained, but, where recognized, simply a creature of state law. While it cannot be arbitrarily granted some owners and denied those we believe are similarly situated, the question of whether we should recognize it as a general right must be decided upon by a legitimate policymaker, presumably for reasons.

Similarly, it is almost embarrassing to read Blits and Gottfredson speak of a "person's right to be hired on his merit." Certainly no such right exists in traditional libertarian theory (the theory I suspect embodies their vaguely formulated "natural rights" perspective). In a world ordered by libertarian principles, employers have the right to hire whomever they please, and the state should never compel contracts, even to meet a presumptively worthy goal (like insuring that merit is rewarded). No such right has ever existed in the United States as a matter of positive law. Unless an employer discriminates for a prohibited reason (such as race or national origin), she is surely entitled to hire for entirely whimsical reasons, or to refuse to hire the most qualified applicant simply because it suits her fancy.

Private and Social Costs

Blits and Gottfredson are as woefully unaware of conventional neoclassical economics as they are of political theory. They mysteriously treat the unexceptionable assumption that profit-maximizing individual employers need not inevitably take steps that result in social product maximization as if it were some wild-eyed radical theory. Private and social costs will in fact diverge wherever transaction costs preclude all those who would like to manifest a preference for a certain outcome from doing so. (A standard instance of this phenomenon is that a profit-maximizing manufacturer will over-pollute the air if those harmed by air pollution have no property right to block pollution and cannot readily organize bids to bribe the polluter to reduce pollution levels, owing to

free rider effects, difficulties of assembling affected parties, etc.) It remains an open question in particular cases whether the divergence between private and social costs should result in a corrective intervention, but the idea that noting the possibility of such divergence "confuses" ideas of efficiency is itself yet one more elementary confusion.

I will curb (though not wholly eliminate) the temptation to hope others simply dismiss the piece because the authors are irresponsible exponents of their own viewpoint. Instead, I want to translate their arguments a bit, and focus on their most substantive aspects: Is the Report's concern with the disproportionate distribution of false negatives and positives in screening through imperfectly predictive tests a spurious concern, as the authors suggest, or is it legitimate? Does the existence of false negatives and positives breach the demands of a meritocratic system? Is it racist? Is the authors' suggestion that the focus of the Report is on "equality of results" rather than "equal opportunity" correct? Could one follow the system suggested in the Report without getting "equal group results"?

The Issue of False Negatives

The issue of false negatives should pose problems for completely individualistic meritocrats even if false negatives were not concentrated on a particular demographic group. It is reasonably simple to deal with a possible assertion that the applicant has been given an equal opportunity by virtue of having had the chance to compete on a test that is equally predictive for himself and others. Tests which bore no relationship to performance would be equally valid for everyone; they would have a universal validity of zero. Yet it would make little sense to say that people had an equal opportunity to get particular jobs just because they had been given a test that did not systematically underpredict their capacity along more relevant dimensions. The point, though extreme, is worth making. Because tests are "objective" and measurably differentiate those who take them, there is a strong tendency to believe, wrongly, that doing better on tests itself gives rise to a claim of merit.

The far trickier point is to justify a claim that a person whose *ex ante* probability of success is in fact lower than another person's is nonetheless treated unfairly if not hired. The applicant with a low test score's claim must be as follows: I am, in fact, *ex post*, if allowed to perform a job, someone who is either a good worker or a bad worker. If I am in fact a good worker, indistinguishable in the relevant sort of merit

(ability to do the job, not ability to take tests) from high-test-score good workers, my chances of having been hired are relatively low, compared to the group of high-test-score good workers. If I am in fact a poor worker, indistinguishable in actual merit from a high-test-score poor worker, my chances of being hired are also dramatically poorer. At any given level of merit, I will be treated worse. It is not clear how a system can be considered meritocratic when people of equal factual *ex post* merit systematically end up in different places.

Assume a medical school admissions officer who has narrowed the applicant pool for a class of 500 to 1000 applicants who seem to him equally capable, 500 women and 500 men. He believes, quite reasonably, that it is bad to waste school resources on those who will not practice medicine, and knows that 96 percent of admitted women become doctors while 98 percent of men do. He also knows there are no more particularized predictors that work better than gender. Admitting only men will indeed meet a valid goal, maximizing the production of doctors, but it should drive meritocrats crazy; each woman is in fact someone who will either become a doctor or not, but the chances that those women who would in fact practice (who are, as it turns out, 96 percent of the population) get admitted is zero, while the chances that those men who will not in fact practice get admitted is 100 percent. From the purely individualistic meritocrat's perspective, nothing at all turns on the fact that the discriminating device in this case is group membership itself. A psychological test that measured "propensity to finish school" which was able to distinguish those who fail the test (and in fact become doctors 96 percent of the time) and those who pass (who become practitioners 98 percent of the time) would be similarly troublesome if we focused our attention on whether merit was inexorably rewarded.

The fact that false negatives are concentrated on a particular group (whether women in the med school hypothetical or blacks in the real world of GATB screening) is independently, additionally bothersome, whether one believes groups have "rights" or not. Normative individualists can still recognize the social reality of groups. Individuals within socially recognized groups may, for instance, have atypically interdependent welfare functions (each is more prone to be injured by losses to others in the group and made happy by gains to those in the group). Similarly, longer-term gains in motivation may depend on the visible success of same-group role models. Group

membership may be a decent, administrable proxy for the presence of current victimization that cannot be cheaply remedied on a case-by-case basis. Blacks are still victimized by racism whether they have been victims of discrete acts of discrimination by particular individuals whom it is worthwhile to sue or not.

The argument that Title VII might impose even costly duties on employers to insure that blacks wind up in the same positions as factually equal whites is far more complex and ambiguous than I can detail in this sort of brief reply. But Blits and Gottfredson's casual dismissal of the consistency of race-norming with meritocratic ideals is just one more instance of thoughtless naiveté. It is especially troublesome in dealing with tests like the GATB that are only minimally predictive of performance, explaining roughly 5 percent of the variance in worker performance as measured by supervisors.

Compatibility with Result-Oriented Quotas?

Race-norming tests insures that the same proportion of each normed group will be selected at the initial screening stage. It is completely silent, though, on the question of whether the employer may take subsequent steps (such as probationary employment periods) which would result in disproportionate dismissals and a "racially unbalanced" work force. There may or may not be independently satisfactory reasons why one would demand racial balance in the work force. Such reasons might well be wholly separate from those proffered to justify the demand for racial balance at the screening stage, given the conceded concentration of false negatives on blacks at this screening stage.

It may well be true that people who advocate the abolition of test-based screening with racially disparate impact independently favor racial quotas for some or all jobs, but there is no necessary or logical connection between the two positions. The *Connecticut v. Teal* court rejects the use of a certain sort of unjustified screening device (an unvalidated one) even when group quotas are met, precisely because the use of such a test breaches the right each minority group member has as an individual to meritocratic treatment.

The issue of work force quotas, a complex one in its own right, must be far more sharply differentiated from the question of the propriety of using screening devices with racially disparate impact in screening workers who are in fact of equal quality. Once one recognizes that GATB screening poses meritocratic problems, it remains a tricky issue whether quotas at the screening stage, rather than, say, test abolition, is the appropriate response to this particular sort of disparate impact.

Mark Kelman is Professor of Law at Stanford University, and is the author of A Guide to Critical Legal Studies, as well as a number of articles about law and economics, taxation, criminal law, and legal theory.

transaction

THE SUPREME COURT'S CONSTITUTION

Bernard H. Siegan

This critical examination of the Constitution and Supreme Court cases argues that the Court has frequently ruled without regard to Constitutional meaning and purpose. It provides instructive distinction between activist and literalist traditions in the high court.

ISBN: 0-88738-127-8 (cloth) \$29.95

ISBN: 0-88738-671-7 (paper) 240 pp. \$14.95

THE ART OF JUDGING

James E. Bond

Bond characterizes the controversy about the role of the Supreme Court and its interpretation of the Constitution as a debate between advocates of two different styles of judging, craftsmen (who look backward for guidance), and statesmen (who look forward to moral and political ideals). Only the former, he argues, can preserve both law and constitutionalism in democracy.

ISBN: 0-9120-51-13-2 (cloth) \$19.95

ISBN: 0-9120-51-14-0 (paper) 110 pp. \$12.95

JURISCULTURE

GREECE AND ROME

Gray L. Dorsey

In this first of a definitive seven-volume work, Dorsey locates the roots of Western jurisprudence. Here he explores the formation and regulation of early Greece and classical Rome in relation to prevailing beliefs about reality, knowing, and desiring. Dorsey reveals interrelations between ideas of antiquity and legal institutions and principles.

ISBN: 0-88738-237-1 (cloth) 92 pp. \$24.95

At your bookstore or direct from the publisher.
Major credit cards accepted.
Call (201) 932-2280



Transaction Publishers
Department LS
Rutgers-The State University
New Brunswick, N.J. 08903

Court Innocence

O. Peter Sherwood

Alexandra Wigdor and John Hartigan have ably and with commendable restraint responded to Blits and Gottfredson by focusing the reader on what the Final Report of the National Research Council's Committee on the GATB (General Aptitude Test Battery) contains. The following addresses directly an aspect of the Blits/Gottfredson essay in an effort to identify the sources of the legal policies they challenge.

When Sue Harris moved into her new apartment five years ago, she received a flowering plant from her sister, Joan. Sue placed it on the wide ledge outside a window three stories above Main Street. There it thrived until today when it fell and struck Jim Cole as he walked along the sidewalk. These facts are sufficient to raise a presumption in favor of Mr. Cole that Ms. Harris was guilty of negligence and is responsible in damages for injuries he suffered. This result flows from the application of one of the oldest doctrines in the common law, called *res ipsa loquitur* ("the thing speaks for itself"), see, for example, *Jesionowski v. Boston & Maine R.*, 329 U.S. 452 (1947); and *Scott v. Lincoln & St. K. Docks Co.*, 3 Hurlst & C. 595, 159 English Reports Reprint (1865). The simple fact that the flower pot was under Ms. Harris' control (it was on her window ledge), it fell and it caused injury, should result in a determination that she is responsible for the injury unless she can prove that the pot fell for reasons other than her negligence.

Turning a Congressional Statute on its Head

Other examples of "burden shifting," some ancient, some recent, abound in the law. Yet, two social scientists now attack a line of United States Supreme Court decisions, beginning with a unanimous 1971 decision entitled *Griggs v. Duke Power Co.*, holding that under Title VII of the Civil Rights Act of 1964, 42 U.S.C. Section 2000e, if the plaintiff proves that a facially neutral employment practice has an adverse effect against a protected class, the burden of proof shifts to the employer to demonstrate that the practice has a manifest relationship to the job in question. If that

relationship is established, the employer may continue the practice without alteration even if it disproportionately affects protected class workers. The courts, for example, have repeatedly upheld the high school diploma eligibility requirement which most jurisdictions require of applicants for employment as police officers even though the requirement adversely affects black and Hispanic applicants. Blits and Gottfredson suggest that this "effects" test is a judge-made rule which turns the statute Congress wrote on its head. In support, they quote an excerpt from Section 703(a)(1) of Title VII that "it shall be...unlawful...to discriminate against an individual...because of such individual's race." They do not pause to acknowledge the very next sentence, section 703(a)(2), which declares as unlawful, any employment practice that would "adversely affect [a worker's] status as an employee because of such individual's race." Congress examined *Griggs* when, in 1972, it amended Title VII to extend its reach to cover virtually all government employers and greeted it, not as an example of a ruling by an errant court, but as a proper interpretation of the law.

Blits and Gottfredson also criticize the Court in *Griggs* for never mentioning the word "individual." *Griggs* addressed only the first part of plaintiff's case *i.e.*, whether the employment practice complained of had the effect of discriminating on the basis of race. Apparently no one directed these researchers to *Franks v. Bowman Transportation Co.*, 424 U.S. 747 (1976) and its progeny, the Supreme Court decisions that address the second part of plaintiff's claim, *i.e.*, whether the individual plaintiff or plaintiff class members were affected by the practice complained of. Once a pattern and practice violation is proved, individuals must establish that they were affected by the unlawful practice, see *James v. Stockham Values & Fittings Co.*, 559 F.2d 310 (1977). Individuals who were not affected by the practice have no right to a remedy, see, for example, *Int'l Bro. Teamsters v. United States*, 431 U.S. 324 (1977).

Attacking Description

Apart from failure to recognize that the *Griggs* "effects" test applies the plain language of a Congressional enactment and that the notion of burden shifting is solidly rooted in ancient legal principles, the sustained attack on *Griggs* in an article ostensibly devoted to criticism of the Interim and Final Reports of the National Academy of Sciences Committee on the General Aptitude Test Battery seems odd. The Committee's description of the law and competing interpretations of it offered by lawyers who appeared before it, was simply that—a description. The Committee's reports merely reported on existing law. They made no effort to resolve the divergent views that were expressed.

Most of the Committee's Final Report is devoted to describing and analyzing a wealth of data it reviewed on the much studied General Aptitude Test Battery. Some of that data is reported for the first time and they call into question aspects of some often cited studies, including studies on which these two researchers have relied. One suspects that these disclosures and their implications could be grist for lively discussion and debate within the community of social scientists and elsewhere. Yet, in a mere phrase, the Blits/Gottfredson piece dismisses the entire Final Report [it "is largely concerned with technical issues of testing"] and then launches into criticisms of the social policy choice

Congress made over a quarter of a century ago when it enacted Title VII and has since reaffirmed.

In this Nation, anyone may express disagreement with social policies even if they are written into law. Whether Congress may impair the "right" of citizens guaranteed by the Second Amendment to bear arms by barring the sale or possession of automatic weapons; whether government should subject legitimately acquired automobiles, planes, boats, and houses used in connection with illegal drug transactions to forfeiture; or whether cigarette smoking should be banned on commercial airlines are all examples of important social policies that have generated debate in our society. While disagreement with a given social policy may well serve as a catalyst to spur individuals to social activism, it is hardly a basis for a reasoned attack on a simple description of that policy. Neither law nor social science is advanced by sacrificing fair treatment of the facts on the altar of ideology.

O. Peter Sherwood is Solicitor General of New York State. A litigator, he has tried cases and argued appeals in the state and federal courts, including several cases involving challenges to employment testing practices under Title VII of the Civil Rights Act of 1964. He lectures frequently and has taught constitutional law and fair employment practices law.

NEW FROM MASSACHUSETTS

FRENCH PHILOSOPHY OF THE SIXTIES

An Essay on Antihumanism

Luc Ferry and Alain Renaut

Translated by Mary H. S. Cattani

"The most powerful analysis of Foucault, Derrida, Bourdieu, and Lacan ever to come out of France. . . .

Ferry and Renaut's fair-minded and subtle book serves as an excellent and often entertaining guide to the hidden message of contemporary French philosophy."

—Edmund Leites

Paper \$13.95, cloth \$35

REMEDIAL LAW

When Courts Become Administrators

Edited by Robert Wood

Essays by distinguished practitioners and scholars examine judicial intervention in the administration of various public institutions, including schools, prisons, public housing, and mental health facilities.

\$25

AT BOOKSTORES OR FROM:

The University of Massachusetts Press
Box 429 Amherst MA 01004

THE EXEMPLARY PRESIDENCY

Franklin D. Roosevelt and the American Political Tradition

Philip Abbott

"Professor Abbott stands in the forefront of American scholars on the presidency, especially those who relate the presidency to public philosophy. His *The Exemplary Presidency* is a work of the highest moral and political imagination that promises to endure as a seminal contribution to political thought."—Kenneth W. Thompson

Paper \$13.95, cloth \$35

MANAGING LIVES

Corporate Women and Social Change

Sue J. M. Freeman

"I think the book will be read by psychologists, people interested in management and business administration, women's studies scholars, and a more general public consisting of the kinds of women Freeman has interviewed for the book. It's clearly argued, well organized and nicely written."—Joan W. Scott

May, paper \$13.95, cloth \$35

Phone orders: 413-545-2219

Fax orders: 413-545-1226