

THE FLIGHT FROM g IN EMPLOYMENT TESTING

Linda S. Gottfredson
College of Human Resources, Education, and Public Policy
University of Delaware
Newark, Delaware 19716
(302) 831-1650
gottfred@udel.edu

Many of you are aware of the controversy over the police entrance exam developed several years ago for Nassau County, NY. You can find the details of that controversy on the IPMAAC webpage (www.ipmaac.org/nassau). I won't be discussing them today, but will instead use the Nassau project to illustrate a broader social phenomenon in the United States that might be called the flight from g, or general intelligence.

The flight from g is a continually evolving, disparate-impact driven effort to minimize or avoid using cognitive ability in selection. I will also describe how the flight from g is pervasive throughout national life, is transforming that life, and is reflexively denied by organizations who engage in it. This flight represents an effort to solve some vexing racial problems, but it will do more harm than good. As I shall explain, the source of most disparate impact may not be racial at all, nor may its amelioration.

Both the Nassau consultants and their critics include professionals whom many of us admire, which has prompted some soul-searching in the profession about the extent to which it has blurred the distinction between measurement and social values (e.g., Mattson, 1997). The critics' chief complaint is that, under pressure from the U.S. Justice Department, the Nassau test battery was stripped of all meaningful cognitive demands in order to nearly eliminate disparate impact against blacks. As Russell (1996) has written, the project's technical report paints a picture of the test developers bending over backwards to eliminate the remnants of cognitive ability tests from their 25-test experimental battery, which the project had previously administered to over 25,000 candidates. The final battery recommended a year later for actually ranking the candidates consisted of eight personality scales and a reading test regraded pass-fail with the passing score set at the first percentile of incumbents. The Justice Department, which in an unprecedented move had jointly developed the test, immediately began pressuring police departments around the country to adopt the test or one like it. The Justice Department's chief litigator in the case said that the test was, in his words, as close to perfect vis-a-vis disparate impact as he had seen in his many years of practice.

Cognitive ability, however, is a strong predictor of learning police work, it is a useful predictor of supervisor ratings of police performance, and it shows up as critical in job analyses of police work, including the project's own job analysis in Nassau County. Schmidt (1996a, 1996b) has shown that the project vastly overestimated the validity of its cognitively-denuded final battery, and he warns that the test will be a public safety disaster wherever it is used (see also Gottfredson, in press).

Impact Driven

To restate the phenomenon more generally, the “flight from g” is the effort by organizations to reduce or eliminate racial differences in outcomes by minimizing or eliminating consideration of individuals’ cognitive skills and abilities in making decisions about them. What distinguishes it from merely sloppy or haphazard selection is that it is an active effort, even if sometimes unwitting, to avoid selecting for g or general mental ability.

This effort to minimize the role of cognitive ability in selection has been precipitated by the dilemma with which you are all familiar. Cognitive tests are the best single predictor of job performance overall, but they tend to have much disparate impact due to large average racial gaps in cognitive skills. As you also know, however, there is relentless legal and political pressure to eliminate disparate impact in all arenas of social life. Because mental standards have disparate impact, this means that there is enormous pressure to degrade or eliminate those standards, even where they matter most, say, in high-level or critical and sensitive jobs where mistakes can cost lives, property, and public trust.

To take Nassau County as an example, the Justice Department forced it during twenty years of litigation not only to lower its educational requirement for police candidates (a requirement which, by the way, it can use only if it does not have disparate impact) but also to abandon each valid entrance exam the county had produced under successive consent decrees. Justice was satisfied only when the county succeeded in virtually eliminating disparate impact—the disparate impact ratio was 77%--which, as I have described, entailed eliminating any meaningful mental standard.

Continually Evolving

Since 1971, when the Supreme court declared that disparate impact constitutes prima facie evidence of discrimination, the United States has become one vast experimental lab in how to reduce disparate impact. Employers’ first response to the Supreme Court’s Griggs decision was simply to drop mental tests. However, that exacted considerable cost in lost productivity. The reality of the importance of mental ability for job performance has a way of asserting itself when it is ignored.

Much validation research followed Griggs and showed that mental tests are not culturally biased and are valid for perhaps all jobs. Research in the field of intelligence simultaneously confirmed that all cognitive tests measure g more than they do any specific mental ability that a test might have been intended to measure, that average racial differences in test scores are larger on more g-loaded mental tests, and that racial differences in g are real, although their source is still unknown. By the 1980s, it had become clear that improving cognitive tests was not, as many had hoped, the solution to disparate impact. Indeed, higher reliability and validity often increased impact.

This recognition fueled the move by many employers to use different hiring standards for different races in order to eliminate disparate impact. The most sophisticated form of double standards was to race-norm test scores. Some personnel psychologists advocated this form of race-conscious selection in order to reduce impact because it exacts fewer short-term costs in lost productivity than does lowering mental standards across the board for all races. However, the procedure’s naked race-consciousness led Congress to outlaw it in 1991.

This ban in turn reinvigorated the search for useful non-cognitive predictors of job performance, because they have little or no disparate impact. Researchers have indeed been able to develop and validate less cognitive selection devices, such as biodata and personality tests. They have also shown that when used to supplement cognitive tests, they can often raise a selection battery's validity while lowering its disparate impact.

If that is the good news—and good news it is—, the bad news has been that adding less-cognitive tests to batteries containing cognitive tests does not lower their impact enough to pass the federal government's 80% rule for determining disparate impact. As Schmitt and his colleagues (in press), among others, have recently shown, passing that rule requires virtually eliminating any cognitive content. Not surprisingly, then, this is where many psychologists have been investing their creative energies in recent years—developing ways to minimize the use or impact of mental standards while not eliminating them altogether. Techniques include, for example, setting low minimum cut-offs on cognitive tests, banding the scores, or reducing the tests' weight in the overall selection process. However laudable such g-minimizing efforts may seem from some perspectives, they should be recognized for what they are—racially-motivated efforts to minimize the use of a known valid predictor.

Racially gerrymandering the content of tests, as was done in the Nassau County, is a more sophisticated and harder to detect method of minimizing g in selection. It involves finding technical pretexts for picking and choosing content according to the racial results it produces. Such racially-motivated gerrymandering of test content was also a predictable development in the evolving flight from g; indeed, it was the logical next step when race-norming was banned. Although Sackett (1997) has recently defended the Nassau project, he and Wilks (1994) have discussed how employers, now denied the use of race-norming, might begin reducing the efficiency of their selection systems in order to reduce disparate impact. They might "'bury' predictors with adverse impact by using them in such a way that they have a negligible effect on selection decisions (such as setting a cutoff so low that virtually all applicants pass)" (p. 951). But Sackett and Wilks also discussed content-selection procedures which they thought might be prohibited by the 1991 ban on race-norming. They said specifically that they thought that after-the-fact removal of test items based on group differences on those items could be "construed as group-based score adjustment" (p. 940).

In his defense of the Nassau study at the SIOP meetings this spring, Sackett (1997) also touched on what I suspect will be the next major development in the flight from g—changing the criteria against which batteries are validated. Army Project A (Campbell, 1990) showed that cognitive tests best predict the "can-do" dimensions of job performance while non-cognitive tests generally best predict the "will do" component. I predict that we will see more emphasis on the latter dimensions of criterion performance in future validation studies, specifically for the purpose of reducing disparate impact. But I would caution those who are tempted to follow this path. You will get what you select for and lose what you don't. Once again, the reality of the importance of g for core technical job performance, if ignored, will eventually assert itself in practical ways the organization will not appreciate. Too many incompetents, no matter how conscientious and cooperative they may be, can quickly demoralize coworkers, gum up the works in any project, and even do great damage to an organization or its reputation when they make critical mistakes.

Pervasive

Attempts to reduce the role of g are widespread in personnel selection. I suspect it has even become a preoccupation for many in the field. However, we can expect the same phenomenon in all social institutions in which greater mental competence provides individuals greater access to resources, services, rewards, or official certification of skills, knowledge, and abilities. This means just about everywhere. It may be most obvious, however, in education, a very g-demanding arena to which all citizens are subjected for most their early life, regardless of how able they are.

Like in the employment setting, any educational practice that is g-related routinely comes under heavy fire as racially discriminatory, and is often either abolished or watered down. For example, many schools have eliminated ability grouping and broadened the entrance criteria for gifted classes to include non-cognitive talents in order to prevent racial disparities in class assignments. Disproportionate assignment to special education classes, minimum competency tests for high school graduation, and teacher certification tests are also perennial objects of criticism, if not litigation, due primarily to their disparate impact.

Many colleges and universities use racial preferences in order to admit more minority students. However, most have also watered down their intellectual standards for admission. For example, many have added social characteristics to their admissions criteria and greatly reduced the proportion of admissions based strictly on academic merit. These additional criteria tend to be weakly or even negatively correlated with g in order to dampen the advantages of academic credentials that minorities disproportionately lack. The most striking example of avoiding g in admissions, and of that avoidance's disastrous consequences for educational quality, are the open admissions policies adopted by colleges such as the once-great City University of New York.

Institutions that have recently been barred from using racial preferences, like the Universities of Texas and California, are all now contemplating other means of increasing minority admissions. Most turn out to involve some flight from g. To illustrate, the University of Texas has decided to offer admittance to the top 10% of all graduating seniors in each of the state's high schools. Given the enormous variability among high schools—the student bodies of some perform much above the national average and others much below it—this practice would admit less able students on the average and create a much more ability-heterogenous student body overall. To the extent that high schools are racially segregated, however, such a policy constitutes a crude form of race-norming and thus yields greater racial parity in admissions.

I should also mention in passing that just as g-related forms of competence are deemphasized in preferring candidates when they have disparate impact, so too are g-related forms of incompetence, inadequacy, and misbehavior for rejecting candidates. Employers, including police departments, are not allowed to turn away job applicants convicted or charged with serious crimes because criminal records have disparate impact. And as Zelnick (1996) describes in his book *Back Fire*, the Justice Department has also forced insurers and mortgage lenders to ignore major actuarial risks in providing and pricing home insurance and mortgages, once again to reduce disparate impact.

In short, the dictates of the disparate impact standard are creating a topsy-turvy world that punishes institutions for rewarding competence or for sanctioning social pathology when

competence and pathology have disparate impact, as they usually do. This practice stigmatizes merit and promotes multiracial incompetence.

Transforming American Life

The incessant, impact-driven pressure for lowering common standards of performance and behavior is transforming organizations and American life to a degree we have yet to appreciate. Let's take a small example. Police hiring standards in New York City collapsed some years ago in the effort to meet racial goals, which in turn led to numerous highly-publicized problems on the force. The point I wish to make here, however, is that the new lows in performance led to changes in the job itself, specifically, to reducing the discretion allowed police officers. I am told, for example, that for some years now, York City police officers have not been allowed to make arrests without first obtaining approval from their supervisors, who often actually come to the arrest scene.

If there are not too many incompetents on a work team or in a workforce, then more able workers can supervise or finish the work that the less competent cannot, albeit not without resentment for the lack of reciprocity involved and not without hurting their own efficiency. But once the number of poor performers reaches a critical mass, the job itself has to be narrowed to prevent costly mistakes, meaning that work is shifted to other classes of workers such as supervisors, as was the case with New York City police.

As I noted earlier, some personnel professionals argue that organizations should use racial preferences to reduce disparate impact rather than lower standards for all races. Let us set aside for a moment the implication that minority skill levels cannot be raised and the disturbing ramifications of racial preferences. The point I would make is that racial preferences eventually lead to degraded common standards anyway, but just generate more racial animosity in the process. Let me provide an example of how this happens, especially when the organization does not want to admit having lowered standards for anyone, which is almost invariably the case.

Bernard Davis describes how, during the 1970s, Harvard Medical School began reserving 20% of its medical school openings for minorities. His 1986 book, *Storm over Biology* (as reviewed by Gordon, 1988, pp. 85-87), recounts the silent erosion of standards that Davis witnessed there as a faculty member.

Because black students experienced their greatest difficulty in basic science courses, it was suggested that the "long tradition of building on these courses as a foundation for clinical training might have been wrong: perhaps one really did not need to be competent in science in order to be a good physician." Letter grading was replaced by the less informative pass-fail criterion, and incompletes were rendered invisible on student records once the missing coursework had been made-up. Such changes made it easier for the dean to claim that performance records of minority students were indistinguishable from those of other graduates. Departments were expressly to permit repeated re-examinations for failing students, "and inevitably these examinations became less demanding." As a by-product, the standards for passing crept downwards for all students.

Before long, the dean's office discontinued yearly reporting of the school's students in the National Board Examinations, until then a ritual. Eventually, the faculty came to rely on passing the National Board Examinations as evidence that its standards had not declined

too far, although Harvard would have considered such a criterion excessively permissive for its students in the past. But the National Board Examinations are renormed each year, Davis informs us in another essay, "and so the absolute norm for passing is necessarily lowered by any nationwide increase in admission of students with substandard academic qualifications."...(At Harvard) a failing student could retake the National Board Examinations five times, but eventually that anemic standard was itself waived and a diploma awarded in the case that at last caused Davis to publish a 1976 guest-editorial in the New England Journal of Medicine in which he sounded the alarm....

[T]o contain the embarrassment...the dean sent out a letter to all medical schools denying that standards had been lowered at Harvard and issued a misleading press release castigating Davis, as Davis's colleagues abandoned him publicly, as blacks debated whether or not he was a racist, and as the Harvard Crimson and [others] rushed to depict him as indeed a racist who questioned the ability of all black medical students if not all blacks.

Now, one seldom gets such inside accounts, for obvious reasons. However, this one alerts us to what may be happening now in the Nassau County police training academy as it struggles with its first class of trainees hired with the new test. The new class is rumored to be failing academy exams at much higher rates.

Denial

Let me be clear that we cannot assume that all disparate impact results from racial disparities in g or other job-related skills and abilities. We cannot presume that discrimination has disappeared. However, we do know that still-stubborn racial disparities in g guarantee much disparate impact when unbiased, valid cognitive tests are used in the absence of discrimination to select individuals into curricula, colleges, jobs, and the like.

However, a scripted public rhetoric is enforced in American life to deny that racial differences in outcome might be due in part to racial disparities in key skills and abilities. Adherence to this script is maintained by massive self-censorship so that individuals withhold signalling any awareness that such skills gaps exist, by ritual denials by organizations that they have lowered standards even in the face of obvious evidence otherwise, and by instant vilification of those who have revealed such evidence. Recall what happened to Bernard Davis when he finally went public about how racial preferences had caused Harvard medical school's standards to fall like dominos.

As a society we are engaged in what social scientist Timur Kuran has called "living a lie," and which he describes in his 1995 book, *Private Truths, Public Lies*. Drawing on a comparison with Eastern Europeans under communism, most of whom felt compelled to feign agreement with the official dogma, Kuran notes that many Americans self-protectively report opinions about affirmative action and race that they do not believe. Many others tacitly support a public opinion sharply at variance with their private convictions by simply holding their tongues.

Such reticence to question the reigning dogma makes it easy for institutions to deny that they have substantially reduced disparate impact by lowering their standards. Few individuals will contest the denials. None may actually know the weight of private opinion, but all know that expressing ideologically incorrect beliefs will surely provoke unpleasant social sanctions.

Although some of the avid supporters of a public lie know it to be false, others will believe it because they have been misled or themselves resolutely avoided the truth. When information is suppressed, ignorance reigns. And it reigns piously when, as in universities today, liberal arts and education courses routinely teach falsehoods about intelligence and mental tests as if they were incontrovertible truths.

When such falsisms abound and remain uncontested by skeptics, is it any wonder that the Nassau County project could claim the equivalent of having found that water runs uphill—namely, that cognitive ability is not useful in hiring police officers? And that is precisely what the Nassau report concluded—that cognitive tests add no meaningful validity to non-cognitive tests in predicting who will be a good cop. That the project nonetheless decided to give candidates at least some credit for passing their watered-down reading test at the incumbent first percentile suggests that at least some of its members knew otherwise.

While the project's conclusion lacked credibility in view of the extant literature on g and its predictive validity, so did the project's starting premise that the consultants could develop a test battery with significant cognitive content that would not have much disparate impact in Nassau County. The general structure of mental abilities is well known and, at least for our purposes, has been described in excruciating detail (Carroll, 1993). All cognitive abilities are moderately to highly correlated with g, which is their common and dominating element.

Nor is there anything mysterious about the practical, everyday meaning of g. As described in the just-published special issue of *Intelligence*, "Intelligence and Social Policy" (Gottfredson, 1997a), g can be conceptualized at the molar level as the relative rate of learning reasonably complex material (Carroll, 1997). This conceptualization can be translated into two others that are also useful: g as the ability to deal with complexity (Gottfredson, 1997b) and g as the probability of not making cognitive errors (Gordon, 1997). The manifest mental behaviors that are prototypical of g include problem-solving, abstract thinking, and reasoning. We often designate these in short-hand terminology as information-processing or critical-thinking skills. The g loadedness of a test depends on the complexity of test items, for example, how abstract the content is or how many bits of information must be apprehended and processed. A test's g loading does not depend on its manifest content (say, verbal versus numerical) or its mode of administration (say, oral vs. written).

The magnitude of racial differences on different cognitive tests is highly correlated with these tests' g loadings, that is, their cognitive complexity. It seems fanciful to suggest, as did the Nassau report, that one might materially reduce the disparate impact of cognitive tests without changing the mental skills they are thought to measure by, for example, going to a video format requiring no reading or writing.

Although the Nassau team continues to defend virtually all the project's odd technical decisions which allowed it to justify its improbable conclusion, even supporters of the study don't seem to take the defense seriously. When asked to discuss the Nassau study at the SIOP meetings this April, Murphy (1997) and Sackett (1997) both avoided discussing that defense or any specifics of the study, and spoke instead only in generalities about the complexity of test validation. While refusing to criticize or even discuss particulars of the project, both discussants nonetheless spoke as if the study team had in fact done precisely what it has stoutly denied doing, namely, restricted the cognitive content and subsequent validity of its test battery for purely racial reasons. Specifically, the two discussants gave reasons for why

and when they thought one should disregard valid cognitive tests to meet racial goals. Sackett argued that if that is what an organization wants, then that is what the consultant should do. Taking a different tack, Murphy suggested that it might show a pro-segregation attitude if one chooses not to reduce the cognitive component of a test battery. In other words, both defended the study by suggesting that the team was justified in doing what it had denied ever doing.

It is worth looking at Murphy's argument in more detail because it also illustrates how social pressure is used to maintain public lies in the face of contrary knowledge. He stated quite emphatically that cognitive ability "is special." To quote him, it "is relevant to every job we know about" and it "is the best single predictor" of job performance. He continued, however, that if we believe that we must therefore include cognitive tests in selection batteries, we must announce that we also believe that it is "better to have racial segregation than other outcomes." We are, according to him, "definitely also saying that we should not have as many black police officers."

What he is saying here quite explicitly is that if you favor maximizing validity you are declaring yourself not just a racist but a segregationist. Being pro-validity is being pro-segregation. What once was taught as good professional practice—enhancing validity—now apparently transforms us into Bull Connors. To avoid becoming moral pariahs we must limit validity when that allows us to reduce impact. Murphy's is but a new variation on the old discussion stopper that is used so regularly to squelch honest debate, namely, threatening to label people racist if they don't fall into ideological line. The tactic is not meant to silence the speaker so much as it is to intimidate third parties who might have been tempted to express agreement with the speaker. It is in this manner that the public dogma about race is enforced in the United States.

I would remind you at this point that disparate impact is lawful if you can demonstrate test validity and that the 1991 Civil Rights Act expressly forbids hiring practices that are racially motivated. The implication of Murphy's argument is that adhering to civil rights law by seeking the most job-related tests makes you a closet racist whereas violating the law by racially gerrymandering your tests to reduce impact protects you from such insinuations. As I mentioned earlier, the flight from g has created a topsy-turvy world.

Iatrogenic

However, might not it be better for social harmony to live a lie? Might not it be a useful fiction to maintain that all races come to the labor market, the mortgage lender, and the insurance broker equally qualified? Wouldn't it be harmful to puncture that fiction? Currently fashionable ideology would have you believe so, but let me raise a few troubling questions about its unexamined assumption that this lie is beneficial.

Let us take the matter of productivity first. What does it mean for a national economy when it spends so much effort deliberately limiting or rolling back the valid selection of perhaps its most valuable asset—its intellectual capital—especially in this information age? I sometimes wonder just how many employees in large companies and public agencies are devoted full- or part-time to restricting or reducing the reliability and validity of selection systems. We do know that there is a very large and intimidating division of the U.S. Justice Department pressing them to do so. All this effort surely constitutes a rather large tax on our economy,

one not shared by our global competitors. If we are not careful, we may end up fighting over slices of an ever shrinking economic pie.

Second and perhaps more insidious is the impact on race relations of the public lie which requires and then camouflages the flight from *g*. If a substantial proportion of racial inequalities in life chances are due to racial gaps in skills but we cannot say so, how then is inequality to be understood by the public? The ideologically correct answer today, of course, is white racism. But note what that means. If we cannot address the real problem of cognitive skill deficits, those continuing deficits will assure us in perpetuity the appearance of ubiquitous and unyielding white racism, even where none exists, and in virtually all facets of social life.

For blacks, this apparently pervasive racism is all the more diabolical for seeming to be unseen, unconscious, institutional, and literally woven into every nook and cranny of American life. For their part, whites resent the constant imputation of racism, especially when the charges become self-evidently ludicrous and when personal experience begins to suggest contrary explanations. In short, the public lie about racial parity in basic skills not only guarantees a continued flight from *g*, but also that both races will continue to accumulate grievances against the other. Note that these racial grievances stem not from problems of race per se but from the race-neutral problems of cognitive skill deficits. This is extraordinarily important. We are creating divisive racial problems out of non-racial ones.

What we have is a common non-racial problem—cognitive deficits in substantial portions of all races—that affirmative action and anti-racism programs do nothing at all to ameliorate in any race. Both blacks and whites have many skill deficient members who are finding it increasingly difficult to succeed in our increasingly complex post-industrial world. And both races are becoming increasingly bifurcated along economic lines, with both the black and white underclasses continuing to grow as the economy's use of low-skill workers shrinks.

Finally, policies grounded in falsehoods provide only illusory solutions. More resources are being poured today into denying than into constructively dealing with the fact that a disproportionate number of blacks are seriously skill deficient. Life gets markedly more difficult the lower one's general ability to deal with complexity, solve problems, and learn efficiently. That is a cross-racial problem that hits the black community harder, but it cannot be adequately addressed in any group without directly tackling the cognitive disadvantages involved. We should not overpromise but be realistic about our currently limited ability to ameliorate *g*-related social inequalities. We can continue to flail away at such realism as hard-hearted while leaving the skills problem untouched, or we can work slowly but directly and doggedly toward reducing it. The flight from *g* only hobbles our ability to make a difference.

Personnel psychologists have tried hard to ease the disparate impact problem. Perhaps no profession has tried harder. But we must remember that its primary mission is to husband human capital. It certainly should not be obscuring the import of one of its chief components. At some point the profession must clarify what it cannot do and what it should not be asked to attempt. Professionalism always requires specifying its limits. Perhaps we are at this point now, judging from the soul-searching I mentioned earlier. President Clinton has said that we need more candor in our discussions of race. We should accept that invitation. We can urge standing fast on standards while trying to raise skills. It may be slow, hard work, but both national productivity and racial progress depend on it.

References

- Campbell, J. P. (1990). An overview of The Army Selection and Classification Project (Project A). *Personnel Psychology, 43*, 231-239.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Carroll, J. B. (1997). Psychometrics, intelligence, and public perception. *Intelligence, 24*, 25-52.
- Gordon, R. A. (1988). Thunder from the left. *Academic Questions, 1*(3), 74-92.
- Gordon, R. A. (1997). Everyday life as an intelligence test: Effects of intelligence and intelligence context. *Intelligence, 24*, 203-320.
- Gottfredson, L. S. (Ed.) (1997a). Intelligence and social policy [special issue]. *Intelligence, 24*(1).
- Gottfredson, L. S. (1997b). Why g matters: The complexity of everyday life. *Intelligence, 24*, 79-132.
- Gottfredson, L. S. (in press). Racially gerrymandering the content of police tests to satisfy U.S. Justice Department: A case study. *Psychology, Public Policy, and Law*.
- Kuran, T. (1995). *Private truths, public lies: The social consequences of preference falsification*. Cambridge, MA: Harvard University Press.
- Mattson, J. (1997, July). President's message. *PTC Newsletter*. Arlington, VA: Personnel Testing Council of Metropolitan Washington.
- Murphy, K. (1997). *Discussion*. A presentation in the symposium "Police Selection in Nassau County: Validity and Demographic Diversity" at the annual meeting of the Society for Industrial and Organizational Psychology, St. Louis, April 12.
- Russell, C. J. (1996, July). *The Nassau County police case: Impressions*. Unpublished manuscript, University of Oklahoma. Available www.ipmaac.org/nassau/.
- Sackett, P. R., & Wilks, S. L. (1994). Within-group norming and other forms of score adjustment in preemployment testing. *American Psychologist, 49*, 929-954.
- Sackett, P. R. (1997). *Discussion*. A presentation in the symposium "Police Selection in Nassau County: Validity and Demographic Diversity" at the annual meeting of the Society for Industrial and Organizational Psychology, St. Louis, April 12.
- Schmidt, F. L. (1996a, December 10). New police test will be a disaster [Letter to the editor]. *Wall Street Journal*, p. A23.
- Schmidt, F. L. (1996b, July). *Some comments on the Nassau County police validity case*. Unpublished manuscript, University of Iowa. Available www.ipmaac.org/nassau/.

Schmitt, N., Rogers, W., Chan, D., Sheppard, L., & Jennings, D. (in press). Adverse impact and predictive efficiency using various predictor combinations. *Journal of Applied Psychology*.

Zelnick, R. (1996). *Back fire: A reporter's look at affirmative action*. Washington, DC: Regnery Publishing.