Sharf, J. (Jan. 2003). The Industrial/Organizational Psychologist, 40(3), 31-40.

# The Third Circuit Again Rejects Relative Merit

James C. Sharf

Sharf and Associates, Employment Risk Advisors, Inc.[1]

Last October the Third Circuit[2] handed down their second look at South Eastern Pennsylvania Transit Authority's employment tests in Lanning v. decision interprets the Civil Rights Act of 1991 to require employers in the to "measure the **minimum qualifications** necessary for successful performance of the job in question in order to survive a disparate impact challenge." What follows is the background and facts of this case, relevant language from the Third Circuit's decision, and invited comments from knowledgeable employment attorneys.

## Background

The Civil Rights Act of 1991 defined the employer's rebuttal to a "disparate impact" discrimination claim involving objective assessment as "job related for the position in question and consistent with business necessity." As noted previously in a TIP[3] review of the Department of Justice's brief and the Third Circuit's June 1999 Lanning decision, because the legislative history of the Civil Rights Act of 1991 had conflicting interpretations of what was meant by "business necessity," the adversarial process of interpreting the meaning of this burden has now begun and will likely ultimately require Supreme Court clarification.

The Clinton Administration's Civil Rights Division in the U.S. Department of Justice had joined plaintiffs in February 1977 and had been successful in advocating the following:[4] "the 'business necessity' burden is separate and addition to the 'job relatedness' burden;" that even with criterion-related validity, the employer "must still demonstrate the 'practical significance' of any correlation between those criteria and the test;" and "even where a test itself is demonstrably job-related, an employer must still show that the chosen cutoff score predicts successful job performance and distinguished applicants who will be successful performers on the job from those who will be unsuccessful." One of the early moves of the Bush Administration was to withdraw the Civil Rights Division's participation in this case.

In my opinion, however, because the court took notice of DOJ's 'validate the cutoff' argument, the damage to selecting on the basis of relative merit had been done – at least to employers in the Third Circuit. So now that I have had a hand in designing the airport security screener employment tests, what am I ethically obligated to tell TSA (much less the traveling public in and out of airports in PA, NJ & DE)?[5] Passing scores are to be lowered so that public safety will be served by "minimally qualified" screeners who "are likely to be able to do the job"?

## Facts of the case

In upgrading the Philadelphia transit police (SEPTA), SEPTA's consultant, an exercise physiologist, determined that running, jogging and walking were important tasks for patrol officers and initially decided that an aerobic capacity of 50/mL/kg/min was necessary to perform the job. After determining that such a high standard would have a

draconian effect on women applicants, he decided that the goals of SEPTA could be satisfied by using a 42.5 mL/kg/min standard. Using this standard, the pass rates for women and men during the time period under challenge were 6.7% and 55.6% respectively - a 5.56 standard deviation disparity. After litigation commenced, SEPTA hired expert statisticians who demonstrated a statistically significant correlation between aerobic capacity and arrests, arrest rates and commendations.

The District Court held a bench trial[6] in January 1998, and rendered an opinion in favor of SEPTA in June 1998. Based on the consultant's reports, the District Court held that SEPTA had established that its aerobic capacity requirement was "job related and consistent with business necessity." This decision was appealed by the Department of Justice to the Third Circuit which in June 1999 remanding the case back to the District Court to develop the record as to what is meant by "business necessity" – an invitation for outside parties to submit *amicus* briefs to the court. Because the meaning of the Civil Rights Act of 1991's "business necessity" rebuttal burden has now been defined in one circuit court's "first impression," it is likely that future issues of TIP will follow other precedents as circuit courts opine in different fact situations.

## Third Circuit's October 2002 "Lanning II" Decision

"In Lanning v. Southeastern Pennsylvania Transportation Authority (3d Cir.1999), we held that 'under the Civil Rights Act of 1991, a discriminatory cutoff score on an entry level employment examination must be shown to measure the minimum qualifications necessary for successful performance of the job in question in order to survive a disparate impact challenge.' We found that the District Court did not employ this standard, which was implicit in Griggs v. Duke Power Co. and incorporated by the Act, and, thus, vacated the judgment of the District Court and remanded the appeal for the Court to determine whether the employer, the Southeastern Pennsylvania Transportation Authority ("SEPTA") had carried its burden of establishing that its 1.5 mile run within twelve minutes measures the minimum aerobic capacity necessary to perform successfully the job of a SEPTA transit police officer. We left it to the discretion of the District Court to allow the parties to expand the record in keeping with our newly-announced standard. Because we conclude that SEPTA produced more than sufficient competent evidence to support the finding that a pre-hire, pre-academy training aerobic capacity of 42.5 mL/kg/min measures the minimum qualifications necessary for successful performance as a SEPTA transit police officer and has, thus, justified the conceded disparate impact on female candidates by showing business necessity, we will affirm the judgment of the District Court in favor of SEPTA…"

"We clearly do no write on a clean slate. The District Court conducted a twelve-day bench trial in January of 1998 after which it rendered a 162-page opinion detailing 378 findings of fact and 107 conclusions of law… On appeal, we, too, rendered a lengthy opinion with a lengthy dissent. To be sure, the majority opinion spend much time explaining how the standard announced therein came to be, but that opinion, and the dissent, discussed much more, including why SEPTA's concern over public safety caused it to modify its hiring requirements, the history of this litigation, and key pieces of evidence. On remand, the District Court conducted a five-day hearing, after which it

rendered a 69-page decision detailing yet another 153 findings of fact and 34 conclusions of law…"

"So much has been written and so little remains for determination that we do not believe it necessary to repeat what has been said before or, as does the dissent here, poke a hole here or there in one or more of the District Court's extensive findings of fact and conclusions of law. There is, however, one undisputed fact which bears repetition because it sets the stage for what is to follow: it is undisputed that SEPTA management wanted to improve the crime fighting ability of SEPTA's force and the fitness of its officers…"

"We also reiterate what we explicitly stated in Lanning I: the business necessity standard takes public safety into consideration. We observed, in Lanning I, that Congress viewed the 'more liberal test for business necessity' adopted in Wards Cove Packing Co., v. Atonio as a significant departure from Griggs and intended, when it enacted the Civil Rights Act of 1991, to endorse the business necessity standard enunciated in Griggs and not the Wards Cove interpretation of that standard. Nonetheless, we noted: '(T)o the limited extent that the Supreme Court's pre-Wards Cove jurisprudence instructs that the public safety if a legitimate consideration, application of the business necessity standard to SEPTA is consistent with that jurisprudence because the standard itself takes public safety into consideration. If, for example, SEPTA can show on remand that the inability of a SEPTA transit office to meet a certain aerobic level would significantly jeopardize public safety, this showing would be relevant to determine if that level is necessary for the successful performance of the job. Clearly a SEPTA officer who poses a significant risk to public safety could not be considered to be performing his job successfully. We are accordingly confident that application of the business necessity standard to SEPTA is fully consistent with the Supreme Court's pre-Wards Cove jurisprudence as required by the Act.'"

"It is against this backdrop that we assess the sole issue we caused to be resolved on remand: whether or not SEPTA has proven that its 42.5 mL/kg/min aerobic capacity standard measures the minimum qualifications necessary for the successful performance of the job of SEPTA transit police officers. The District Court concluded that the answer was 'yes,' and that any lesser standard 'would result in officers… who were a danger to themselves, other officers, and the public at large, (and) unable to effectively fight and deter crime.' … (W)e conclude that the District Court's findings of fact were not clearly erroneous."

"And so we move more directly to **the critical issue before us – the minimum qualifications necessary in terms of aerobic capacity to successfully perform as a SEPTA transit police officer** (emphasis added). Neither the District Court nor the parties have explicitly defined the key phrase 'minimum qualifications necessary,' but a definition is implicit in the parties' respective arguments and the District Court's acceptance of that of SEPTA. SEPTA argued that the run test measures the 'minimum qualifications necessary' because the relevant studies indicate that individuals who fail the test will be much less likely to successfully execute critical policing tasks. For example, the District Court credited a study that evaluated the correlation between a successful run time and performance on 12 job standards. The study found that individuals who passed the run test had a success rate on the job standards ranging from

70% to 90%. The success rate of the individuals who failed the run test ranged from 5% to 20%. The District Court found that such a low rate of success was unacceptable for employees who are regularly called upon to protect the public. In doing so, the District Court implicitly defined 'minimum qualifications necessary' as meaning 'likely to be able to do the job.'"

"The District Court cited numerous other studies that offer similar results. In one such study, 80% of those passing SEPTA's run test met minimum job standards, while only 33% of those failing did. Another study showed that 84% of those passing the test could carry out an 'emergency assist,' while only 14% of the failing group were able to do so. The consideration that the District Court gave to these studies lays to rest plaintiffs' claim that the cutoff time was merely the product of the judgment of SEPTA experts. **As we noted in Lanning I, a 'business necessity standard that wholly defers to an employer's judgment as to what is desirable in an employee… is completely inadequate'** (emphasis added). The factual record here, however, clearly demonstrates that SEPTA experts set the run time cutoff at 12 minutes for objective reasons, with the studies showing that the projected rate of success of job applicants dropped off markedly for those who ran 1.5 miles in over 12 minutes."

"Plaintiffs argued, however, that within the group that failed the run test, significant numbers of individuals would still be able to perform at least certain critical job tasks. They argued that as long as some of those failing the run test can do the job, the standard cannot be classified as a 'minimum.' In essence, plaintiffs proposed that the phrase 'minimum qualifications necessary' means 'some chance of being able to do the job.' Under this logic, even if those failing the test had a 1% chance of successfully completing critical job tasks, the test would be too stringent."

"We are not saying, as our distinguished brother in dissent suggests we are saying, that 'more is better.' While, of course, a higher aerobic capacity will translate into better field performance – at least as to many job tasks which entail physical capability – to set an unnecessarily high cutoff score would contravene Griggs. It would clearly be unreasonable to require SEPTA applicants to score so highly on the run test that their predicted rate of success be 100%. It is perfectly reasonable, however, to demand a chance of success that is better than 5% to 20%. In sum, SEPTA transit police officer and the public they serve should not be required to engage in high-stakes gambling when it comes to public safety and law enforcement. SEPTA has demonstrated that the cutoff score it established measures the minimum qualifications necessary for successful performance as a SEPTA officer…"

"The dissent concedes that SEPTA has the right to improve its workforce and does not suggest that that is not being done. Instead, the dissent concentrates its efforts on why, in its view, the 42.5mL/kg/min aerobic capacity cutoff score as an application requirement is wanting, concluding that '(a)fter all has been said and done, … one unassailable fact remains. The 42.5 mL/kg/min aerobic capacity (cutoff) is not required of transit officers before or after they begin policing. As for the 'before,' we reject without more argument that applicants – male and female – should not be tested until they have graduated from the police academy, perhaps two and one-half years after they first applied to SEPTA; indeed, the dissent recognized but relegates to a footnote the increase in SEPTA's costs and the uncertainty in planning and recruitment this would occasion. As for the 'after,'

all incumbents – male and female – are now required to take a physical fitness test every six months, another step toward improving the workforce. In this connection, it bears mention that SEPTA is unable to discipline incumbents who do not pass the test only because of the patrol officers' union's challenge, sustained by an arbitrator. With the union's blessing, however, SEPTA offers financial incentives to those officers who do pass. One final note. While it is undisputed that SEPTA's 1.5 mile run test has a disparate impact on women, it is also undisputed that, in addition to those women who could pass the test without training, nearly all the women who trained were able to pass after only a moderate amount of training. It is not, we think, unreasonable to expect that women – and men – who wish to become SEPTA transit officers, and are committed to dealing with issues of public safety on a day-to-day basis, would take this necessary step. Moreover, we do not consider it unreasonable for SEPTA to require applicants, who wish to train to meet the job requirements, to do so before applying in order to demonstrate their commitment to physical fitness. The poor physical condition of SEPTA officers prior to 1989 demonstrates that not every officer is willing to make that commitment once he or she is hired. In any event, the multi-agency training which SEPTA candidates receive does not provide sufficient physical fitness training to bring an unqualified candidate up to the physical standards requirement. Of course, yet another step in improving the performance of incumbents would be to require a physical fitness test not only upon application but also immediately prior to entry on duty."

"The judgment of the District Court will be affirmed."

## Invited Comments from Employment Attorneys

Walt Connolly, Foley & Lardner, Detroit[7]

"The Third Circuit has eschewed the Wards Cove standard in favor of the Griggs standard. Lanning is a pragmatic decision trading adverse impact against women applicants against public safety and SEPTA's business necessity burden. Frankly, I would be reluctant to translate this public safety precedent into a run of the mill entry-level employment decision where one would expect to find a correlation between the employment test and the performance on the job. Remember that courts in the past have given greater deference to universities and to public safety issues involving jobs such as truck drivers, airline pilots and bus drivers. Rank order testing has already been suspect and this is why we have recommended banding and the use of cut off scores where there is a clearly demonstrable relationship with job performance."

David Copus, Jones Day, Washington DC[8]

"Putting aside the RESULT and looking only at the RATIONALE -- it seems to me that somehow the court has lost sight of the fact that any test, physical or otherwise, is merely a rough predictor designed to increase the likelihood of success on the job. We just want to the test to be better than a toss of a coin. There will never be a test that can guarantee success. Even the most "valid" test -- i.e., one with a high correlation coefficient, is going to explain only a part (often a small part) of the variation in performance. Thus, to say that a test measures the "minimum qualifications necessary to perform the job" confuses me. We know that's not true -- otherwise, we'd give the test to all incumbents and

fire those who fail. Maybe I'm stupid, but I find all the appellate decisions in that case to be goofy.

Here's an example I've used to help our clients understand this concept. If you were selecting a team to climb Mt. Everest with you, would you require all applicants to have two legs? Would you require them all to have reasonably good vision? Sure you would, on both counts. You'd want the best, strongest, most able team you could find. But, a one-legged guy has climbed Everest; so has a blind guy. So, what ARE the minimum requirements to climb that mountain?

It does not make sense to say that any given set of qualifications is the MINIMUM NECESSARY to do a job. All we can ever say is that we are not going to take a chance on anyone below that level of qualifications."

Robert J. Malionek, Latham & Watkins, Los Angeles[9]

"Much has been written of Lanning v. SEPTA (Lanning I), including the inherent dichotomy between the legal standard announced by the Third Circuit for employers to justify the use of discriminatory cutoff scores on employment tests – *i.e.*, they must measure "the minimum qualifications necessary for successful job performance" – and the professional standards of the field of industrial and organizational psychology (indeed, the court in Lanning I even shunned the SIOP Principles as "not instructive"). Now that we have Lanning II, the latest decision on yet another appeal in the same case, we can expect that the writings on this subject will not stop any time soon. Because the panel of judges in Lanning II was legally bound to follow the standard that another Third Circuit panel announced in Lanning I, the standard remains the law of that Circuit. Nonetheless, it is the clarification and application of that standard to the facts in Lanning II that can be recognized as bringing about a subtle shift away from what could have been a draconian standard in practice.

What we knew from Lanning I about this standard and how employers could meet it was very little. In that case, the Third Circuit cited the choice by the employer's primary test validation expert to ignore the estimate of the subject matter experts (incumbents) regarding the minimum 1.5-mile run time (translated into aerobic capacity) necessary to perform the job of SEPTA transit officer. SEPTA's evidence of a simple analysis demonstrating the correlation between aerobic capacity and certain criteria of performance, such as absolute number of arrests and arrest rates – *i.e.*, "more is better" evidence – without regard to any link between the cutoff score on the test and minimum necessary job ability, the court stated, was insufficient to overcome a conclusion that the employer's screen was an "arbitrary barrier to employment opportunities." The court then remanded to the trial court to apply this new standard. The trial court held that the employer met it, and the plaintiffs – female applicants who failed the test and the U.S. Department of Justice – appealed again to the Third Circuit. (The DOJ originally supported the plaintiffs, but abandoned its participation from this appeal over a year before the Third Circuit reached its conclusion in Lanning II.)

What we know from Lanning II is that the standard is not one of absolutes. Recognizing that statistics tell different stories when told by different experts, the court was satisfied with 'sufficient competent evidence' by the employer that its cutoff score meets the

"minimum qualifications" standard. Parsing through the opinion, one can see the court attempting to bridge the relevant professional standards to its fledgling legal standard, and in particular one can identify several factors which seem paramount considerations for any employer or testing expert studying an allegedly discriminatory cutoff score on an employment test in the Third Circuit:

First is the business justification of the employer. While the court in Lanning I expressed skepticism about accepting an employer's justification for using a discriminatory employment practice at face value, the Lanning II court took stock in the undisputed fact that SEPTA's stated goal was to improve the physical fitness of its officers. In fact, crediting the employer's iteration of its business justification for the screen served as the backdrop for the court's entire opinion.

Second is the relationship of the employer's business justification – and particularly, in the case of a criterion-related validity study, the criteria studied – to the requirements of the job itself. Here, the court placed the *criticality* of the job tasks and abilities that formed the criteria for the study above the frequency at which they are performed; in other words, those tasks and abilities need not dominate the requirements of the job. Not all arrests are 'aerobic contests, nor are they always effectuated to apprehend 'serious' criminals,' and SEPTA officers generally were required to engage in at least one aerobic encounter with a suspect every month, either as an emergency assist or running backup of another officer, the court noted. But the evidence at trial supported the conclusion that the inability of an officer to perform *any important task* proficiently would compromise the effectiveness of SEPTA, and that demanding anything less would pose a danger to the officer, other officers and the public at large. In short, the court recognized that, 'a SEPTA transit police officer must be ready and able to apprehend not just the numerous sedentary, petty criminals, but also the fleet-footed few who, from time to time, wreak serious harm on the people of Philadelphia.'

Third is the relationship of the predictor to those criteria. Easily satisfied with a showing of a statistically significant relationship between aerobic capacity and one of the criteria (*e.g.*, arrest rates), the court could turn its attention to an analysis of the cutoff score. This is where the court demonstrated that the 'minimum qualifications' standard is not a hard and fast line that must be met with precision (and probably could never be, given what we know about the properties of tests generally). In applying that standard, the court easily rejected the plaintiffs' argument that as long as some test failers were able to perform the job (*e.g.*, some incumbents in the validation study), the cutoff score cannot be considered the 'minimum.' The court considered as sufficient evidence to justify the cutoff score a validation study which demonstrated that test passers had a success rate on the various 'job standards,' or criteria, of 70-90%, while the success rate of test failers was only 5-20%. The court was more concerned with the somewhat ambiguous conclusion that the success rate of failers 'dropped off markedly' from that of passers than it was with the plaintiffs' false negatives argument (an argument which is all the rage within the Department of Justice in recent litigations).

What the court in Lanning II made clear is that the "minimum qualifications" standard is not meant to demand a perfect cutoff score which separates out *all those who can* perform the job from *all those who cannot*. Rather, the court held, it means the cutoff score should differentiate between those who are "likely to be able to do the job" and

those who are not. It is 'perfectly reasonable' for an employer to expect that the applicants it hires possess – at the time of hire and not after training or some experience on the job – the abilities that are necessary to enable him or her to perform any and all important tasks of the job. An employer need not "engage in high-stakes gambling" by hiring applicants with unacceptable probabilities of performing those tasks (even if there is *some* probability that they could perform them). And with that, employers can thus rest assured that the standards of industrial and organizational psychology are alive and well in the Third Circuit. But we'll all need to wait and see how those standards will fit in to the next set of facts to reach that court."

Keith Pyburn, Fisher & Philips, New Orleans[10]

"The most recent Lanning v. SEPTA decision continues the debate over the dilemma created by the 3rd Circuit's initial conclusion that the Griggs "business necessity" standard includes a requirement that a "cut off" score be shown to be "valid." The original *en banc* decision held that if any selection procedure had a disparate impact, then its use could only be justified if the "cut score" did not exclude any "qualified candidates."

The majority in the latest decision, while reiterating the language of the original holding, proceeds to allow a "reasonable" cut score based on the factual record of this case, perhaps most importantly the "public safety" nature of the jobs in question. The dissent logically points out there are no suggested "standards" for determining when such a "discriminatory" cut score is "reasonable."

The future of the "no false negatives" rule remains in doubt. A second *en banc* review is certainly possible."

Ted Schroeder, Littler Mendelson, Pittsburgh[11]

"Revisiting a case that has for several years troubled testing experts, the United States Court of Appeals for the Third Circuit has issued its second opinion in *Lanning v. SEPTA* ("*Lanning II*"). *Lanning II* clarifies the Third Circuit's "minimum qualifications" standard for satisfying Title VII's business necessity defense in a manner that supports the use of properly validated testing as a legally defensible method for hiring and promotion decisions

Plaintiffs in *Lanning* were unsuccessful female applicants for police jobs with the Southeastern Pennsylvania Transit Authority ("SEPTA"). They challenged SEPTA's requirement that applicants have an aerobic capacity of 42.5mL/kg/min, as demonstrated by completing a 1.5 mile run in 12 minutes. Plaintiffs claimed this requirement is illegal sex discrimination under Title VII, because it has a disparate impact on women.

SEPTA conceded that its aerobic capacity test has a disparate impact but argued that the test was a legitimate selection device for police officers. Therefore, the critical issue was whether SEPTA could establish that the test was "job related and consistent with business necessity." After a trial in January 1998, the trial court held that SEPTA had satisfied this standard. In *Lanning I*, The Third Circuit reversed, holding that the district court had applied the wrong standard for business necessity, and that SEPTA must establish that its test "measure[s] the minimum qualifications necessary for the successful performance of

the job in question." *Lanning I* was troublesome because the case suggested that a test must not screen out any minimally qualified applicant and that employers could not demand a higher level of performance expectation.

*Lanning II* suggests that these fears are unfounded. The Third Circuit affirmed the trial court's judgment, after additional hearings, upholding the use of the test. In doing so, the court clarified the meaning of its "minimum qualifications" test, adopting the trial court's interpretation that the test means "likely to be able to do the job." The court held that SEPTA had satisfied this standard where the studies performed by its experts showed that the success rate on the job was 70-90% for those who passed the test, but only 5-20% for those who failed. Notably, the court held that it "would clearly be unreasonable to require" SEPTA's test to predict success at a rate of 100%. The court also noted that the minimum qualifications standard does not deny companies the ability to improve their workforce."

---

[1] jim@jimsharf.com

[2] PA, NJ, DE & U.S. Virgin Islands

[3] Sharf, J. (Oct. 1999). Third Circuit's Lanning v. SEPTA Decision: 'Business necessity' requires setting minimum standards. The Industrial/Organizational Psychologist, 37(2), 138-149.

[4] See Sharf (Oct. 1999) (fn 3 above) for U.S. Department of Justice's brief to the Third Circuit.

[5] American Psychological Association (1992). Ethical Principles of Psychologists and Code of Conduct. Principle 1.02, Relationship of Ethics and Law: "If psychologists' ethical responsibilities conflict with law, psychologists make known their commitment to the Ethics Code and take steps to resolve the conflict in a responsible manner."

[6] Argued before a judge alone and not before a jury which would have heard arguments had plaintiffs alleged "disparate treatment."

[7] wconnolly@foleylaw.com

[8] dacopus@jonesday.com

[9] robert.malionek@lw.com

[10] kpyburn@laborlawyers.com

[11] tschroeder@littler.com