

PRECEDENTIAL

Filed October 15, 2002

UNITED STATES COURT OF APPEALS
FOR THE THIRD CIRCUIT

No. 01-1040

CATHERINE NATSU LANNING; ALTOVISE LOVE;
BELINDA KELLY DODSON; DENISE DOUGHERTY;
LYNNE ZIRILLI

v.

SOUTHEASTERN PENNSYLVANIA TRANSPORTATION
AUTHORITY ("SEPTA"); RICHARD EVANS, Individually and
in his official capacity as Chief of SEPTA
Police Department

(D.C. Civil No. 97-cv-00593)

UNITED STATES OF AMERICA

v.

SOUTHEASTERN PENNSYLVANIA TRANSPORTATION
AUTHORITY

(D.C. Civil No. 97-cv-01161)

CATHERINE NATSU LANNING; ALTOVISE LOVE; BELINDA
KELLY DODSON; DENISE DOUGHERTY and
LYNNE ZIRILLI,
Appellants

Appeal from the United States District Court
for the Eastern District of Pennsylvania
(D.C. Civil Nos. 97-cv-00593 and 97-cv-01161)
District Judge: Honorable Clarence C. Newcomer

Argued: February 11, 2002*

Before: ROTH, McKEE and BARRY, *Circuit Judges*.

(Filed October 15, 2002)

Jules Epstein, Esquire
Kairys, Rudovsky, Epstein,
Messing & Rau, LLP
924 Cherry Street, Suite 500
Philadelphia, PA 19107

Michael Churchill, Esquire (Argued)
Public Interest Law Center
of Philadelphia
125 South 9th Street, Suite 700
Philadelphia, PA 19107

Counsel for Appellants

Saul H. Krenzel, Esquire (Argued)
Saul H. Krenzel & Associates
42 South 15th Street
The Robinson Building, Suite 800
Philadelphia, PA 19102

Counsel for Appellee

Terry L. Fromson
Women's Law Project
125 South 9th Street, Suite 300
Philadelphia, PA 19107

* This matter was originally heard on February 12, 2002 before Judges Mansmann, McKee, and Barry. Because Judge Mansmann died prior to an opinion being rendered, the Panel was reconstituted to include Judge Roth, and the appeal was submitted.

Sharyn A. Tejani
Feminist Majority Foundation
National Center for Women
& Policing
1600 Wilson Boulevard, Suite 801
Arlington, VA 22209

Counsel for Amici Curiae

OPINION OF THE COURT

BARRY, *Circuit Judge*.

In *Lanning v. Southeastern Pennsylvania Transp. Auth.*, 181 F.3d 478 (3d Cir. 1999) (hereafter "*Lanning I*"), we held that "under the Civil Rights Act of 1991,¹ 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." 181 F.3d at 481. We found that the District Court did not employ this standard, which was implicit in *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971), 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. *Id.* at 494. 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

1. 42 U.S.C. § 2000e-2(k).

disparate impact on female candidates by showing business necessity, we will affirm the judgment of the District Court in favor of SEPTA. We have jurisdiction pursuant to 28 U.S.C. § 1291.²

We clearly do not 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. *Lanning v. Southeastern Transp. Auth.*, 1998 WL 341605 (E.D. Pa. June 25, 1998). On appeal, we, too, rendered a lengthy opinion with a lengthy dissent. To be sure, the majority opinion spent 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. *Lanning v. SEPTA*, 2000 WL 1790125 (E.D. Pa. Dec. 7, 2000).

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. As Judge Weis explained in *Lanning I*:

Concerned about its inability to control crime on its property, SEPTA instituted a three-pronged attack on

2. This case was consolidated before the District Court with *United States v. Southeastern Pennsylvania Transportation Authority* (D.C. Civ. No. 97-cv1161). The United States filed a notice of appeal from the final judgment now under review, with the appeal docketed at No. 01-1297. On October 1, 2001, the United States moved to dismiss its appeal pursuant to Fed. R. App. P. 42(b). That motion was granted on October 12, 2001.

the problem. It added a substantial number of officers, implemented a zone method of patrol, and adopted standards to improve the generally poor physical condition of its officers. Unlike many metropolitan police departments, SEPTA officers are deployed alone and on foot, engaging in physical activities more frequently than other law enforcement agencies.

The patrol zones present significant variations in conditions that affect the physical exertion of officers in the performance of their duties. Zone One, for example, has a climb of 30 to 50 steps from street level. Zone Three, a mixture of above and below-ground locations, borders a large shopping mall, featuring retail theft and pursuits that lead into the SEPTA transit system. Zone Five, which includes sports complexes, is characterized by long distances between stations. Zone Six includes the Temple University area, a scene of frequent crimes against students.

SEPTA officers must occasionally ask for assistance from their comrades in other zones. These calls are divided into two categories, "officer assists" and "officer backups." An "assist" requires officers to respond immediately. Often the only method available to get to the scene quickly is a run of five to eight city blocks. An officer responding to an "assist" must preserve enough energy to deal effectively with a situation once arriving on the scene. SEPTA averages about 380 running assists per year. "Backups" are not as critical as "assists," so officers generally use a "paced jog." SEPTA averages about 1,920 "backups" annually.

181 F.3d at 494-95 (Weis, J., dissenting).

We also reiterate what we explicitly stated in *Lanning I*: the business necessity standard takes public safety into consideration. 181 F.3d at 490 n.16. We observed, in *Lanning I*, that Congress viewed the "more liberal test for business necessity" adopted in *Wards Cove Packing Co., Inc. v. Atonio*, 490 U.S. 642 (1989), 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. We, thus, considered, in *Lanning I*, only *Griggs* and its pre-*Wards Cove* progeny. Nonetheless, we noted:

[T]o the limited extent that the Supreme Court's pre-*Wards Cove* jurisprudence instructs that the public safety is 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 officer 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.

181 F.3d at 491 n.16.

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." *Lanning* 2000 WL 1790125, at *28 (Conclusion # 32). In the course of our concluding that the Court's conclusion was correct, we have reviewed the relevant findings of fact under the clearly erroneous standard and the District Court's conclusions of law under the plenary review standard. *Newark Branch, NAACP v. City of Bayonne*, 134 F.3d 113, 120 (3d Cir. 1998); *Lanning I*, 181 F.3d at 484-85.

We start by noting that when this case was last before us, we expressed concerns as to certain aspects of the

statistical studies upon which the District Court relied, and we "encourage[d] the District Court to take a critical look at these studies, if necessary, on remand." 181 F.3d at 493 n.21. In its second memorandum opinion, the District Court noted that it had "indeed taken a second critical look at these studies as suggested and once again reaffirm[ed] their validity consistent with its prior extensive 160 plus page memorandum opinion." *Lanning*, 2000 WL 1790125, at *1 n.1. Significantly, after the five-day hearing on remand, the District Court made additional findings of fact that bolster its conclusion. As the following brief discussion demonstrates, it is evident that the District Court's findings regarding the studies in this case were not clearly erroneous.

In *Lanning I*, we were concerned that "the absolute number of arrests or 'arrest rates' [in certain studies did] not necessarily correlate with successful job performance." 181 F.3d at 492 n.21. While we continue to believe that a "SEPTA officer should generally attempt to control a situation without having to make an arrest," *id.*, the District Court found, based upon its review of the extensive body of evidence before it, that "lost arrests have a significant impact on the public safety" in the SEPTA transit system. *Lanning*, 2000 WL 1790125, at *26. This finding indicates that arrests do, in fact, correlate to successful performance as a SEPTA transit police officer.

We also questioned whether certain studies overemphasized the role of aerobic capacity in making arrests, and whether other studies placed too much emphasis on arrest rates for "serious crimes" given that, historically, the bulk of SEPTA arrests has been for lesser crimes. While not all SEPTA arrests are aerobic contests,³ nor are they always effectuated to apprehend "serious" criminals,⁴ the District Court found that "[a]n inability to

3. Although there is some dispute over the precise percentage of arrests that have required substantial aerobic exertion, the evidence shows that, at the very least, around one in ten arrests, or ten percent, have so required.

4. Arrests for "serious crimes" have similarly accounted for around one in ten arrests.

proficiently perform *any . . . task*] would compromise the effectiveness of the SEPTA transit police." 2000 WL 1790125, at *24 (emphasis added). In essence, the Court concluded what, to us, is now evident: 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.⁵

A final concern mentioned in *Lanning I* involved the statistical methods relied upon in some of SEPTA's studies. We noted, for example, that a study on arrest rates seemed to be based on "disproportionately large numbers of officers with an aerobic capacity over 42 mL/kg/min" and stated that, if the correlation coefficient of the study was low, the study would "be subject to close review." 181 F.3d at 492-93 n.21 (citing 29 C.F.R. § 1607.14(B)(6)). Upon further review, we are satisfied that the correlation coefficient was sufficiently high to be statistically significant.

We have also examined the other relevant studies in this case for possible deficiencies and conclude that they were sufficiently reliable and in keeping with sound principles of statistical analysis. As scientists well know, it is not uncommon for competent statisticians to interpret data in different ways, sometimes reaching different results. While we concede that there may be some room for disagreement — as there often is in the realm of statistical analysis — the District Court carefully considered the relevant studies and credited them. Under the standard that governs our review, we 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

5. The District Court also credited studies which indicate that once SEPTA officers reach suspects, they often must engage in physical struggles which require substantial aerobic capacity. The Court further noted that SEPTA officers "must engage in at least one aerobic encounter during the course of his or her duties every month, either as an emergency assist or running backup." 2000 WL 1790125, at *24. These findings support the conclusion that aerobic capacity is an important factor in successfully making arrests.

aerobic capacity to successfully perform as a SEPTA transit police officer. 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%.⁶ 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 so doing, the District Court implicitly defined "minimum qualifications necessary" as meaning "likely to be able to do the job."

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

6. 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." *Id.* at 489 (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.

"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%.⁷ In sum, SEPTA transit police officers 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 42.5 mL/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." Dissent at 34. As for the "before," we reject without more the 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 recognizes 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

7. Even in other studies where those failing the test had higher predicted rates of success on job tasks, the rates were still only around 33%.

toward improving the workforce.⁸ 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 standard requirements.

The judgment of the District Court will be affirmed.⁹

8. 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.

9. We have considered the parties' other assertions of error and find them unavailing.

McKEE, *Circuit Judge* dissenting.

In *Lanning v. Southeastern Pennsylvania Transp. Auth.*, 181 F.3d 478 (3d Cir. 1999) (hereafter, "*Lanning I*"), 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." 181 F.3d at 481. The Southeastern Pennsylvania Transportation Authority ("SEPTA") requires all applicants for the position of Transit Police Officer to be able to run 1.5 miles in twelve minutes, which translates into an aerobic capacity of 42.5 mL/kg/min. Despite the numerous opinions that have been written in this suit, and an involved and lengthy bench trial that was followed by additional testimony on remand, I do not believe SEPTA is any closer than when we started to demonstrating that the *pre-training* 42.5 mL/kg/min cutoff — a standard that disqualifies 90% of female applicants from even beginning training — satisfies the business necessity standard. Accordingly, I respectfully dissent from the opinion of my colleagues.

I.

Before I explain why SEPTA has not met its burden and why we should therefore reverse the District Court's decision on remand, see *Lanning v. SEPTA*, 2000 WL 1790125 (E.D. Pa. Dec. 7, 2000) (hereafter "*Lanning II*"), I think it important to dispose of one concern at the outset.

My disagreement with my colleagues is not about public safety or the importance of safeguarding SEPTA'S ridership. No rational or informed person can seriously doubt the importance of qualified police officers to a large urban transit system. Yet, on remand, the District Court expressed the following concern after summarizing much of the evidence that had been offered in support of the 42.5 mL/kg/min aerobic capacity cutoff:

Significant gains in apprehensions and deterrence such as those demonstrated here are to be encouraged and supported by the federal courts. The Court simply will not condone dilution of readily obtainable physical

abilities standards that serve to protect the public safety in order to allow unfit candidates, whether they are male or female, to become SEPTA transit police officers.

Lanning II, 2000 WL at * 25 (E.D. Pa. Dec. 7, 2000). Similarly, the Majority states: "SEPTA transit police officers and the public they serve should not be required to engage in high-stakes gambling when it comes to public safety and law enforcement." Majority Op. at 10. I can not stress too strongly that the issue is not now, and never has been, whether SEPTA must jeopardize public safety in order to eliminate the disparate impact that SEPTA concedes the 42.5 mL/kg/min standard visits upon female applicants. Rather, the issue continues to be whether SEPTA can justify that cutoff under the business necessity test that is incorporated into Title VII by the 1991 Civil Rights Act.¹ Since "the business necessity standard takes public safety into consideration," Majority Op. at 5 (citing *Lanning I*, 181 F.3d at 478, 490) applying it here will not endanger the public.

It is uncontested that plaintiffs have established a *prima facie* case of job discrimination because the 42.5 mL/kg/min cutoff has a disparate impact on females applying to be SEPTA police officers. *Dothard v. Rawlinson*, 433 U.S. 321, 329 (1977). However, that discrimination is permissible if the aerobic capacity represented by that cutoff is "job related for the position in question and

1. As the majority notes, in *Lanning I*, we held that the Civil Rights Act of 1991, codified the Supreme Court's holding in *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971). Under that holding, the employer had the burden of establishing that its "discriminatory cutoff score . . . measure[s] the minimum qualifications necessary to perform successfully the job in question." *Lanning I*, 181 F.3d at 489. We also specifically held that studies demonstrating "that the higher an officer's aerobic capacity, the better the officer is able to perform the job," i.e. that "more is better," have "no bearing on the appropriate cutoff to reflect the minimal qualifications necessary to perform successfully the job in question." *Id.* at 492. We noted that such a study "may suffice . . . in only the rarest of cases where the exam tests for qualities that fairly represent the *totality* of a job's responsibilities." *Id.* at 493 n.23 (emphasis added).

consistent with business necessity. . . ." 42 U.S.C. § 2000e-2(k). SEPTA may therefore justify its discriminatory cutoff by demonstrating that anything less jeopardizes public safety. We explained this in *Lanning I*. There, we stated:

If, for example, SEPTA can show on remand that the inability of a SEPTA transit officer 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.

Lanning I, 181 F.3d at 490. Accordingly, any suggestion that "federal courts" are advocating that diminished public safety is the social price that we must pay to eliminate gender discrimination is as unfortunate as it is inaccurate and misleading. Despite continuing protestations to the contrary, SEPTA has simply not demonstrated the business necessity of its cutoff though it has had several opportunities to do so. The implication that changing SEPTA's cutoff would endanger the public is therefore nothing more than the proverbial red herring. As I shall discuss in greater detail below, the effect of SEPTA's aerobic capacity standard on public safety is purely conjectural.

Moreover, as I shall also explain, the objective evidence on this record should weigh far more heavily in the balance than the conjecture that SEPTA's arguments rest upon. That objective evidence consists of the job performance of a female police applicant who was mistakenly hired although she did not meet the aerobic cutoff when she applied, SEPTA's failure to require officers to meet the standard when they first go on duty as police officers, and the fact that incumbent police officers have regularly failed to meet the 42.5 ml/kg/min cutoff with no demonstrable impact on public safety. Given this record, SEPTA can not shield its unnecessarily exclusionary job requirements from judicial review by relying on a threat to public safety where none has been demonstrated.

Moreover, even if SEPTA had demonstrated the necessary correlation between public safety and the 42.5 mL/kg/min

cutoff, "the plaintiffs may still prevail if they can show that an alternative employment practice has a less disparate impact and would also serve the employer's legitimate business interest." *Id.* at 485 (quoting *Albemarle Paper Co.* 422 U.S. 405, 425 (1975)). In *Lanning I*, we suggested several alternatives that could address SEPTA's justifiable concern about the fitness of its police force while mitigating or eliminating the discriminatory impact of the 42.5 mL/kg/min cutoff.

SEPTA could [for example]: 1) abandon the test as a hiring requirement but maintain an incentive program to encourage an increase in the officers' aerobic capacities; 2) validate a cutoff score for aerobic capacity that measures the minimum capacity necessary to successfully perform the job and maintain incentive programs to achieve even higher aerobic levels; or 3) institute a non-discriminatory test for excessive levels of aerobic capacity . . . that would exclude 80% of men as well as 80% of women through separate aerobic capacity cutoffs for the different sexes.

Lanning I, 181 F.3d at 490.

SEPTA currently uses the aerobic cutoff as an incentive rather than a disqualifier for incumbent transit police as is suggested by the first alternative. There is no evidence that public safety has been compromised by doing that. Therefore, I fail to see why (assuming that everything SEPTA says about the necessity of the standard is true) SEPTA's objectives can't be achieved by offering incentives for applicants just as SEPTA now does for incumbent police officers.

There is an even simpler and more effective option that would also further SEPTA's objective more than the current method of using the aerobic cutoff. SEPTA could impose the cutoff as a condition of graduating from the police academy rather than as an application qualifier. The period between application and actually becoming a transit police officer is sometimes as long as two and one-half years. If this performance standard is as important to the work of a transit police officer as SEPTA claims, the quality of SEPTA's police force would only be enhanced, not

weakened, by requiring officers to meet it when they graduate from the academy.

That was one of our suggestions and it is one of the plaintiffs' proposals. See *Lanning I*, 181 F.3d at 504. That proposal is eminently reasonable, it has several benefits over SEPTA's current application process, and it is a practice that other police forces use.² In fact, the more vital the 42.5 mL/kg/min cutoff is, the more logical it is to enforce it upon graduation from the police academy rather than when one applies a year or two earlier. Doing so would have a less discriminatory impact because it would allow females to complete the application process if they could not run 1.5 miles in the required time, but it would require them to train to the extent necessary to meet the 42.5 mL/kg/min cutoff before starting on duty. It would also be more consistent with insuring that *all* officers satisfy the standard that SEPTA claims to be the minimum necessary for the public's safety.³

The fact that SEPTA does not retest before a new recruit begins patrolling should cause us to be very skeptical about SEPTA's claim that the 42.5 mL/kg/min cutoff is necessary to public safety. That skepticism is reinforced by several undisputed facts on this record.

Yet, on remand, the District Court never even mentioned these options and it appears that SEPTA never attempted to explain why these less discriminatory (and more effective) alternatives would not adequately address any legitimate concerns.

2. Cf. Brief of Amici Curiae at 13 (noting that in recognition of training interval, some agencies, such as New York State Troopers, have different fitness requirements for entrance into, and graduation from, the training academy).

3. I realize that this would increase SEPTA's costs and introduce a new level of uncertainty in planning and recruiting because SEPTA would be less certain of how many of its recruits will ultimately survive the academy and become SEPTA officers. However, common sense suggests that such a "washout" factor already exists because it is a fair assumption that not everyone who begins training ultimately graduates and accepts a position with SEPTA. These are considerations that the District Court never factored into its analysis of business necessity.

II.

The District Court's analysis never focused on the fact that the 42.5 mL/kg/min cutoff is enforced when someone applies for the position of a SEPTA police officer, not when an offer of employment is extended or when an applicant graduates from the police academy and actually goes on the job. All of the studies the District Court relied upon purport to correlate success in job-related tasks to fitness level at the time of the task, not at some time prior to hiring and training. However, an offer to hire may be extended as much as two and one-half years after the aerobic running test is administered and the 42.5 mL/kg/min cutoff applied, and there is absolutely no retest before beginning as a SEPTA officer. See *Lanning v. Southeastern Pennsylvania Transp. Auth.*, 1998 WL 341605 (E.D. Pa. June 25, 1998), and Findings of Fact and Conclusions of Law therein (hereafter "Initial FOF"), ¶¶ 150-151. Accordingly, there is no way for SEPTA to know if even a male incumbent police officer has the aerobic capacity deemed so necessary to the job when he actually begins patrolling. Common sense establishes that no matter how quickly and enthusiastically an applicant may be able to scamper one and one-half miles when initially tested, "ole man time"⁴ may have reduced the applicant's aerobic output beneath the 42.5 mL/kg/min cutoff at the very point that SEPTA claims it is necessary for public safety.

There has been no showing - and no finding by the District Court - that fitness level at the time of application is a reliable proxy for fitness level on the job over the ensuing years. On the contrary, as the majority implicitly acknowledges, being able to run 1.5 miles in 12 minutes prior to training is neither a necessary nor a sufficient condition for being able to run that fast (and consequently, according to SEPTA's theory, being able to perform various police tasks) thereafter.

On the one hand, the majority concedes that "nearly all the women who trained were able to pass after only a moderate amount of training," Majority at 11, so the ability

4. Not to mention the possible intervening intake of pizzas, burgers, and that extra helping of desert every now and then.

to pass without training is not a necessary prerequisite. On the other hand, pre-hire fitness does not guarantee continued fitness on the job. The failure to establish the required nexus between fitness at the time of application and fitness on the job is a major gap in SEPTA's proof. My colleagues in the Majority gloss over this gap noting that it is "not unreasonable" to require women to train in advance, "in order to demonstrate their commitment to physical fitness." Majority at 11. However, it is undisputed that SEPTA does not require applicants or incumbents to train "in order to demonstrate their commitment to physical fitness." There are two additional problems with the Majority's reliance on the relevance of applicants demonstrating a commitment to fitness.

First, the District Court made no finding that applicants who can run fast are more "committed" to physical fitness (or even more likely to remain fit) than someone who may exhibit greater endurance than required for the aerobic cutoff, but not have the speed necessary to satisfy it. Moreover, since running is only one of the physical demands made of an applicant, focusing only on it may ignore the commitment of an applicant who can do a greater number of push-ups or sit-ups, or bench press more weight than someone who can run 1.5 miles in 12 minutes.⁵

Second, since men are far more likely than women to be able to pass the running test without training, this newly-minted commitment rationale imposes an additional discriminatory criterion — viz, that women must possess a demonstrated level of commitment not required of men. Given the substantial physical advantage enjoyed by men, a woman who runs 1.5 miles in just over 12 minutes may thereby demonstrate more commitment to training and fitness than a man who runs the course in just under 12 minutes or 12 minutes "flat." Yet, SEPTA enforces its cutoff in a manner that would exclude the female in favor of the male, despite the fact that the relative achievement level

5. The number of push-ups, and sit-ups an applicant can do is also tested, as is the amount of weight that the applicant can bench press.

may demonstrate a greater commitment to fitness on the part of the female.⁶

III.

Although the requirement of running 1.5 miles in 12 minutes and the corresponding 42.5 mL/kg/min cutoff may not appear that daunting a requirement for someone who exercises regularly and is in fairly good condition, it is nevertheless more than certain branches of the United States military demand of *incumbents*.⁷ I can understand and appreciate SEPTA's argument that its zone system, and its reliance on foot patrols explains why the 42.5 mL/kg/min cutoff is more demanding than the standard set for New York City Police. Differences between what is required of SEPTA police and New York City transit police perhaps rebut the testimony of Former New York City Transit Police Chief, Michael O'Connor, as to the need for SEPTA's aerobic standard. He led a significant reduction of crime in New York City's transit system. When he testified before the District Court as an expert witness in this suit he emphatically rejected the concept of comparing police officers' running ability with that of fleeing felons as a job requirement. SEPTA had argued that its officers must have an aerobic capacity at least equal to that of the "perpetrator" population SEPTA officers may have to chase. Chief O'Connor disagreed. He testified: "[h]ow fast you run is not a measure of how good a cop you are. It takes a lot more than just running fast to be a good cop."⁸

6. The use of differing cutoffs, corresponding to the differing average aerobic capacities of men and women, was one of the less discriminatory alternatives we suggested in *Lanning I*. This alternative was ignored by the District Court.

7. The United States Army requires males ages 22-26 to run 2 miles in 16.36 minutes; the Navy requires males ages 20-29 to run 1.5 miles in 13.45 minutes and; the Marines require males ages 17-26 to run 3 miles in 28 minutes. See Gordon Strong, *Descriptive Comparisons of United States Military Physical Fitness Programs*, 2 Sport J. (1999), <http://www.thesportjournal.org/VOL2NO2/STRONG.HTM>.

8. See J.A., Vol. III at A-725.

SEPTA's own experience confirms this. Crime has been reduced on SEPTA facilities even though SEPTA does not require incumbent officers to achieve the 42.5 ml/kg/min cutoff. This undermines SEPTA's assertion about the relationship between the aerobic cutoff and effective policing, and it corroborates Chief O'Connor's view of the cutoff.

Although differences between New York City and Philadelphia may cast doubt on the Chief's attempt to assess the value of applying this cutoff to the Philadelphia transit system, it does not explain why SEPTA insists upon a more demanding standard than is required of young people training for the Army, Navy or Marines. Yet, SEPTA argues that its standard is the *minimum* necessary for public safety, and that it is necessary to screen out applicants who fail to meet that standard without regard to whether the standard will be attained after they are hired and without regard to whether incumbent police officers can meet it.

IV.

SEPTA's insistence that the 42.5 ml/kg/min cutoff satisfies the business necessity test is undermined by the fact that it does not require its incumbent police officers to meet that standard. Yet, there is nothing to suggest that public safety has been jeopardized. SEPTA first administered a running test in 1991. "SEPTA's own internal memoranda document that incumbent transit officers of all ranks have failed SEPTA's physical fitness test . . . , " Initial FOF ¶ 242, since SEPTA began testing.

One of plaintiff's trial exhibits purports to show that 62.20% of incumbent officers have failed the aerobic capacity test since it was first administered in 1991. Initial FOF ¶ 245 (citing Plaintiff's Exhibit 106). The District Court discounted the probative value of that evidence however, because an individual officer would have been counted "a number of times if this officer failed the test a number of times." *Id.* The Court concluded, "[t]hus, this evidence is not entitled to much weight." *Id.*

Nevertheless, even assuming that the 62.20% failure rate is somewhat inflated because it fails to adjust for multiple failures of individual incumbents, and even assuming that the failure rate has decreased with time, nothing on this record suggests the dramatic turn around one would have a right to expect if meeting that standard is so necessary to public safety, and SEPTA has the burden of proof here, not plaintiffs.

It is uncontested that, despite SEPTA's claim that officers who can not meet the 42.5 mL/kg/min standard endanger public safety, SEPTA has promoted officers who failed the running test, and given commendations to others who failed a component of that test. See Initial FOF 256-259. It is also uncontested that SEPTA has never suspended, reassigned, disciplined or demoted any officer "for failing to perform the physical requirements of the job." Initial FOF 259.⁹ It has also promoted officers who could not meet the standard. SEPTA's policy of promoting candidates who fail to demonstrate a 42.5 mL/kg/min aerobic capacity is perplexing given the purported importance of that performance criteria. The District Court noted:

SEPTA has promoted incumbent officers who have failed some or all of the components of the physical fitness test at any time. Since July 1994, the Chief of SEPTA Transit Police Department has had the authority to remove candidates from promotional lists for failing to achieve their interim fitness goals. Despite

9. Although the District Court is certainly correct that an employer has the right to improve its workforce, and is not bound by the standard of its incumbents, a claim that a highly discriminatory standard that is not required of incumbent police officers is the *minimum* necessary to satisfactory job performance should be met with skepticism.

Cf. *Scott v. City of Anniston*, 597 F.2d 897 (5th Cir. 1979) ("Validation, in general, requires a demonstration that 'the qualifying tests are appropriate for the selection of qualified applicants for the job in question.'" (quoting *Washington v. Davis*, 426 U.S. 229 (1976))). In *Scott*, the Fifth Circuit observed that the "evidence casts doubt on the validity of the tests" where some individuals "who had satisfactorily performed jobs were demoted after [failing] the written tests and others "who had failed written tests performed satisfactorily when they were promoted after passing a performance test." *Id.* at 902.

the authority to remove officers from the promotional lists, no SEPTA officer has ever been removed from a promotional list for failure to pass physical fitness testing for incumbents.

Initial FOF 256. The court attempted to rationalize this apparent contradiction and credit SEPTA with making some effort to consider aerobic capacity when making promotions by noting: "Nevertheless, only ten officers who have failed their physical fitness tests have ever been promoted." *Id.* However, as I have already noted, SEPTA has the burden of proving "necessity." Accordingly, the fact that a relatively small number of officers have been promoted while failing to satisfy the cutoff is not nearly as probative as the fact that 10 officers who failed were promoted with no resulting impact on public safety. After all, SEPTA obviously believed that those 10 officers could do their jobs effectively, and adequately protect the public or it would not have promoted any of them.

The District Court explained that SEPTA's failure to discipline transit police who could not meet the 42.5 mL/kg/min standard was due to a union challenge to management's unilateral imposition of that requirement on incumbent officers. The union's position was sustained by an arbitrator. See *Lanning II*, 2000 WL 1790125 at *6. Management responded by offering monetary rewards whenever an incumbent officer "passed" the running test, and by offering to pay for gym memberships to help incumbent officers train to the level necessary to meet the 42.5 mL/kg/min cutoff. *Id.* Although the union's objections and the resulting arbitration explain why no incumbent was fired for not meeting that cutoff, they don't explain why SEPTA has commended and promoted officers who are purportedly so physically unfit that they are unable to perform the minimum job tasks that are now supposed to be so necessary to public safety. It also does not explain why SEPTA could not offer female applicants who do not meet that cutoff the same support it offers male incumbent transit police officers.

The District Court concluded that:

The experiences SEPTA had with its incumbent officers serves to further illustrate the importance of requiring

incoming officers to meet certain minimum fitness standards, as SEPTA has much less ability to influence its force once they become members of the collective bargaining unit.

Lanning II, 2000 WL at * 6 n.5. Of course, this does not explain why SEPTA does not retest applicants before offering them a position, and it does not even begin to show why the 42.5 mL/kg/min threshold is the *minimum necessary* for public safety.

We are reminded that SEPTA initiated the policy of retesting incumbent police officers every six months in 1991, but the manner in which the "policy" is administered undermines SEPTA's claim of public safety more than it supports it. The District Court concluded that "[d]espite this policy, there was evidence introduced at trial that incumbents are not always retested every six months." *Lanning II*, 2000 WL at * 6 (E.D. Pa. Dec. 7, 2000). Since one's job status is not affected by the outcome of the retests, those examinations are merely an adjunct to the system of rewards and encouragement employed for incumbent officers. As I noted above, SEPTA has not explained why it can not create a similar mechanism of reward and encouragement for applicants who are otherwise qualified to become police officers but can not initially demonstrate an aerobic capacity of 42.5 mL/kg/min.

We are told of how crime in and around SEPTA property has been significantly reduced despite the apparent inability of many transit officers to meet the cutoff, and SEPTA's inability or unwillingness to enforce it. Accordingly, I can only conclude that factors other than the 42.5 mL/kg/min cutoff are responsible for reducing the levels of crime. Crediting the 42.5 mL/kg/min cutoff with reducing crime while not requiring incumbents to meet it totally ignores the other steps SEPTA took that explain the reduction in crime. For example:

SEPTA initiated a complete overhaul of the police department to make the subways on the SEPTA system the "safest place in the city." This overhaul included the announcement that transit police were to

be primarily dedicated to the subway and were not to serve as guards to protect personal or physical property at depots. [And] SEPTA increased the number of officers from 96 to nearly 200 and introduced a "zone concept" for the area they patrolled.

Lanning II, 2000 WL 1790125 at 1.¹⁰

Given the tenuous relationship between aerobic capacity and public safety on this record, it is perhaps not surprising that incumbent officers with an aerobic capacity too low to apply for the job had an arrest rate that was virtually identical to another group of officers who satisfied the 42.5 mL/kg/min cutoff. In *Lanning I*, we noted that "officers at less than 37 mL/kg/min had average arrests of 13.6 compared to officers with at least a 48 mL/kg/min level who had average arrests of 13.9." 181 F.3d at 493 n.21.

V.

Ironically, serendipity has furnished us with a model that we should not so readily dismiss in assessing the business necessity of the 42.5 mL/kg/min cutoff. At the outset I alluded to an instance where SEPTA mistakenly hired a female applicant who did not meet the aerobic cutoff. That officer was hired because of a clerical error. SEPTA concedes that she did not meet the 42.5 mL/kg/min cutoff when she applied and was therefore not fit to even be considered for the position of SEPTA Police Officer under SEPTA's discriminatory application requirement. That officer's performance on a job that she was not "qualified" to even apply for under the 42.5 mL/kg/min aerobic cutoff is significant indeed. Her presence patrolling SEPTA

10. See FOF ¶26 ("In response to these problems, SEPTA initiated a complete overhaul of the police department This overhaul included the announcement that transit police were . . . not to serve as guards to protect personal or physical property at depots. SEPTA increased the number of officers from 96 to nearly 200 and introduced a 'zone concept' for the area they patrolled."). Cf. Brief of Amici Curiae at 14-15 (citing research demonstrating that decreasing crime on transit systems is a function of controlling environmental factors and police presence, not aerobic capacity).

property has apparently not endangered the citizens of Philadelphia at all. On the contrary, as the District Court originally found:

Officer Thomas was hired in 1991 despite the fact that she did not complete the 1.5 mile run in 12 minutes and failed the bench press, sit-up and push-up components of SEPTA's physical fitness test for applicants. Officer Thomas has gone on to become a decorated officer who has repeatedly been nominated for awards such as Officer of the Year and Officer of the Quarter. In fact, SEPTA has commended Officer Thomas for her outstanding performance as a police officer. Moreover, Officer Thomas serves as one of SEPTA's two defensive tactics instructors.¹¹

Lanning v. Southeastern Pennsylvania Transportation Authority, 1998 WL 341605 at * 24, FOR 261 (E.D. Pa. June 25, 1998).

I realize, of course, that the job experience of a single officer is not statistically significant in evaluating the relevance of a standard that is supposed to apply to an entire applicant pool and (at least in theory) an entire police force. However, I do not offer this officer's job performance to counter the statistical record. Rather, I submit that the officer's performance is highly relevant to measuring the necessity of the aerobic cutoff in the context of this entire record. Her performance is consistent with other evidence that undermines SEPTA's justification for gender discrimination.

One of SEPTA's experts, Dr. Bernard Siskin, admitted that correlating arrests and aerobic capacity was not a good

11. Officer Thomas' prowess as a defensive tactics instructor despite her subpar aerobic capacity confirms what common sense would only suggest. The most effective police officers may be those who are skilled enough to arrest someone without engaging in the physical struggle that SEPTA argues partly necessitates an aerobic cutoff at 42.5 mL/kg/min. Indeed, one can not help but be concerned about the idea of a police officer armed with incapacitating chemical spray and a police baton, allowing him/herself to be drawn into a protracted struggle with an arrestee who could grab the officer's gun and use it against the officer or anyone else.

way to predict the quality of an officer's policing. In fact, a study found that individuals with an aerobic capacity of 33 mL/kg/min made more arrests for serious crimes than individuals with an aerobic capacity of 44 and 46 mL/kg/min. J.A., Vol XII at 1837. Accordingly, I question the District Court's conclusion that "aerobic capacity is an important factor in successfully making arrests." Maj. Op. at 8 n. 5. If it is, some of the same officers that SEPTA has promoted and awarded and given commendations to should not be patrolling Philadelphia's transit system.

VI.

The District Court's approval of the 42.5 mL/kg/min standard in large measure rests upon its acceptance of Dr. Siskin's expert opinion. See *Lanning II*, 2000 WL at *7 to *8. Although Dr. Siskin did testify that higher aerobic capacity equated with higher arrests, he did not purport to correlate the 42.5 mL/kg/min cutoff with any satisfactory minimum level of policing. We explained this in *Lanning I* where we stated:

Dr. Siskin testified that in view of the linear relationship between aerobic capacity and the arrest parameters, any cutoff score can be justified since higher aerobic capacity levels will get you more field performance (i.e., "more is better"). See *Lanning*, 1998 WL 341605 at *41. Under the District Court's understanding of business necessity, which requires only that a cutoff score be "readily justifiable," SEPTA, as well as any other employer whose jobs entail any level of physical capability, could employ an unnecessarily high cutoff score on its physical abilities entrance exam in an effort to exclude virtually all women by justifying this facially neutral yet discriminatory practice on the theory that more is better. This result contravenes *Griggs* and demonstrates why, under *Griggs*, a discriminatory cutoff score must be shown to measure the minimum qualifications necessary to perform successfully the job in question.

181 F.3d at 493 (footnote omitted). Yet, the inquiry has advanced no further following remand. I still can not help

but conclude that all of the additional testimony that the District Court heard after *Lanning I*, did nothing more than establish that "more is better." Yet, the majority insists that there is adequate support for this cutoff now. My colleagues argue:

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%.

Majority Op. at 10.

"Reasonable" though it may be, the question remains, what cutoff is *necessary* to ensure public safety and effective policing? Where between the purportedly reasonable level of "5% to 20%" the Majority approves of, and the unreasonable level of 100% it disapproves of, does *Griggs* allow SEPTA to draw a line? Furthermore, why is 5% to 20% reasonable, and why is 100% unreasonable? The questions remain unanswered.

The inquiry would be furthered if SEPTA could point to some objective basis for defining a cutoff or aerobic threshold. However, the 42.5 mL/kg/min threshold seems to have been plucked from the air by Dr. Davis after SEPTA contacted him and asked him to assist in improving the quality of transit police officers. Dr. Davis testified that he initially sought the advice of twenty experienced SEPTA officers (designated as "subject matter experts" or "SMEs") to determine the level of physical exertion they thought was required of transit police officers. The SMEs told him that they thought a SEPTA officer should be able to run a mile in full gear in 11.78 minutes. However, Dr. Davis rejected that cutoff because it was too low. He believed nearly anyone in the general public could satisfy that standard. *Lanning I*, 181 F.3d at 482.¹² He rejected a standard of 50

12. The District Court explained that incumbent officers were older and had a vested interest in establishing a standard that they could meet.

mL/kg/min because it was too high. It would have had a "Draconian effect on women applicants." *Id.* n.3. He therefore apparently decided upon a cutoff of 42.5 mL/kg/min because it was not too high, it was not too low; it was just right. But under *Griggs*, it is not permissible to select a discriminatory employment test in the same manner that Goldilocks chooses which bed to sleep in or which bowl of porridge to eat.

The 42.5 cutoff was not selected based upon the requirements of the job. Rather, it was based upon the expert's (apparently mistaken) conclusion that that level would not exclude women. See FOF ¶ 31. Cf. *Green*, 73 F. Supp. 2d at 200 (describing cutoff score chosen to have less impact than initially selected score as "arbitrarily established"). At the very least, he concluded that the resulting rate of exclusion was acceptable (i.e. "reasonable") because it did not have a "Draconian" impact.

In *Lanning I*, we cautioned:

The danger of allowing an employer to carry its burden by relying simply upon an expert's invalidated judgment as to an appropriate cutoff score in a testing device is illustrated by this case.

181 F.3d at 492 n.19. Accepting the 42.5 mL/kg/min cutoff as being the minimum level consistent with business necessity on this record

disregards the teachings of *Griggs*, *Albemarle* and *Dothard* in which the [Supreme] Court made clear that judgment alone is insufficient to validate an employer's discriminatory practices. . . . [N]owhere in its extensive opinion did the District Court consider whether Dr. Davis' 42.5 mL/kg/min cutoff reflects the minimum aerobic capacity necessary to perform successfully the job of SEPTA transit police officer.

Accordingly, "in Dr. Henderson's opinion it is risky to use incumbent data as a benchmark for establishing entry-level selection devices." 2000 WL 1790125 at *9. That is certainly a logical explanation of why the lower cutoff recommended by the SMEs was rejected. It does not help us to understand how the 42.5 mL/kg/min cutoff was decided upon.

Id., at 491-2. The only explanation of the source of this standard comes from Dr. Davis' testimony during the first trial. It is not reassuring. There, the following exchange occurred between Dr. Davis and counsel for the plaintiffs:

Q. Isn't it correct that you and Dr. Henderson used your own intellectual creativity in coming up with your measures of success?

A. A nice way of saying it, yes.

J.A. Vol. III at A.-639 (emphasis added). None of the *post hoc* justifications for the cutoff suggest that it rests on any firmer ground than that response suggests.

SEPTA made no effort on remand to empirically establish the minimum aerobic capacity necessary to perform as a transit officer. It merely attempted to justify the 42.5 figure arbitrarily selected by Dr. Davis as a "compromise" between a "Draconian" standard that no woman could satisfy on the one hand, and a "Lilliputian" one that anyone could satisfy on the other. As Appellants point out, "[r]ather than attempt to objectively determine a minimum cut score, or whether any was justified . . . [SEPTA's experts] merely sought to measure the effect of the 42.5 standard previously selected by Dr. Davis." Brief of Appellants at 41.

The inquiry was further handicapped on remand by the District Court's refusal to reexamine its original findings. In *Lanning I*, we cited numerous concerns with the various studies that were relied upon to uphold the cutoff, and expressed concern with the District Court's refusal to consider the undisputed evidence that a number of incumbent officers were performing satisfactorily even though they could not meet the 42.5 mL/kg/min standard. We noted:

The District Court rejected as irrelevant the plaintiffs' evidence that incumbent officers had failed the physical fitness test yet successfully performed the job and that other police forces function well without an aerobic capacity admission test. See *Lanning*, 1998 WL 341605 at *68-*70. Under the standard implicit in *Griggs* and incorporated into the Act, this evidence tends to show that SEPTA's cutoff score for aerobic

capacity does not correlate with the minimum qualifications necessary to perform successfully the job of SEPTA transit officer. Accordingly, this evidence is relevant and should be considered by the District Court on remand.

181 F.3d at 494 n.24. We allowed the District Court to exercise its own discretion on remand as to whether to consider additional evidence regarding the necessity of the cutoff. We also noted our concern regarding the court's failure to consider evidence that undermined SEPTA's claim of business necessity.

Nevertheless, on remand, the District Court proclaimed "[f]rom the outset" that it "would not disturb its prior factual findings in this case," although it did allow further development of the record. In other words, it was willing to accept additional testimony to support its conclusion, but was not willing to reexamine the decision that we reversed in light of any new evidence offered on remand. The District Court correctly noted that "the sole question to be resolved [on remand] . . . is whether or not SEPTA has proven that its 42.5 mL/kg/min aerobic capacity standard is the minimum necessary for the successful performance of the job of SEPTA transit police officer." However, it then stated that the evidence at the first trial, both separately and in combination with the evidence adduced on remand, clearly met this test.¹³

The District Court based its conclusion that "any standard less than 42.5 mL/kg/min would result in officers . . . who were a danger to themselves, other officers, and the public at large, [and] who were unable to effectively fight and deter crime"¹⁴ on the following:

Dr. Davis' calculations of the aerobic capacity required to perform essential tasks; the Siskin arrest studies, including the analysis of performance differences between those officers always at 42.5 versus those never at 42.5; Dr. Moffatt's initial study on work output decrements associated with aerobic

13. See *Lanning II*, Conclusions of Law (hereafter "COL") ¶ 4.

14. COL ¶ 32.

capacities below SEPTA's cutpoint; and SEPTA's most recent studies [by Drs. Davis and Henderson, which taken together] more than provides an appropriate empirical basis for demonstrating that [SEPTA's] cutpoint is already set at the minimum.

COL ¶ 31.¹⁵

Whether the analyses and studies provided by SEPTA's experts are sufficient to support this finding turns on whether they were reasonably calculated to measure minimum requirements, *i.e.*, whether they fairly corresponded to actual job requirements and whether they fairly determined minimum performance standards for those requirements.¹⁶ In assessing this, we must remember

15. Two other conclusions of particular import are that the evidence "clearly demonstrates that 42.5 is the minimum aerobic capacity necessary . . . given the abysmal success rate on critical job tasks of those that failed . . . when compared to those that passed . . ." and that "[t]he testimony and studies of Drs. Davis and Henderson conclusively demonstrate that . . . 42.5 mL/kg/min is the minimum required . . ." COL ¶¶ 9-10.

However, as we observed in *Lanning I*,

[i]t is unlikely that such a study could validate rankhiring with a discriminatory impact based upon physical attributes in complex jobs such as that of police officer in which qualities such as intelligence, judgment, and experience surely play a critical role. This is especially true in SEPTA's case, when the record indicates that SEPTA patrol officers encounter 'running assists', the most strenuous task upon which SEPTA's aerobic capacity testing predominately was justified, at an average rate of only twice per year.

Id. 181 F.3d at 493 n.23.

16. This standard is in accordance with Congress's restoration of the pre-*Wards Cove* conception of "business necessity" in which the employer is required to justify not only the legitimacy of the ends, but also the *necessity* of the means. More specifically, the two-part standard follows directly from the Supreme Court's holdings in *Albemarle Paper Co. v. Moody*, 422 U.S. at 431 (holding that "discriminatory tests are impermissible unless shown, by professionally acceptable methods, to be 'predictive of or significantly correlated with *important elements* of work behavior which are relevant to the job . . . for which candidates are being

that the employer, SEPTA, has the burden of proof as a matter of express congressional policy, and that that policy is based upon a recognition of the very real social harm resulting from the disparity that the employer must justify.¹⁷ Consequently, the federal courts have a heightened responsibility to scrutinize the employer's proffered justification.¹⁸

Finally, the Equal Employment Opportunity Commission's Guidelines also call for heightened scrutiny where there is a severe disparate impact, low statistical correlation, or over-emphasis on limited aspects of job performance. Each of these are present here. See 29 C.F.R. § 1607.14B(6). The District Court's failure to heed those

evaluated'") (emphasis added) and *Dothard* 433 U.S. at 332 (indicating that "a discriminatory employment practice" such as a discriminatory cutoff score on an entry level exam "must be shown to be necessary to safe and efficient job performance to survive a Title VII challenge") (emphasis in original). See also *Griggs*, 401 U.S. at 432 (stating that "any given requirement must have a manifest relationship to the employment in question").

17. See Pub. L. 102-166 § 3(4) (1991) (Civil Rights Act was passed "to respond to recent decisions of the Supreme Court by expanding the scope of relevant civil rights statutes in order to provide adequate protection to victims of discrimination").

Cf. *Griggs*, 401 U.S. at 431 (holding that Title VII requires "the removal of artificial, arbitrary, and unnecessary barriers to employment when the barriers operate invidiously to discriminate on the basis of racial or other impermissible classification"); *Lanning I*, 181 F.3d at 489 ("[O]nly by requiring employers to demonstrate that their discriminatory cutoff score measures the minimum qualifications necessary . . . can we be certain to eliminate the use of excessive cutoff scores that have a disparate impact on minorities as a method of imposing unnecessary barriers to employment opportunities.").

18. Cf. *J.E.B. v. Alabama ex rel. T.B.*, 511 U.S. 127, 135 (1994) (applying "heightened" or "intermediate" scrutiny to peremptory strikes based on sex); *Clark v. Jeter*, 486 U.S. 456, 460 (1988) (observing that "intermediate scrutiny . . . generally has been applied to discriminatory classifications based on sex") (citing *Mississippi University for Women v. Hogan*, 458 U.S. 718, 723-24 and n. 9 (1982); *Mills v. Habluetzel*, 456 U.S. 91, 99 (1982); *Craig v. Boren*, 429 U.S. 190, 197 (1976); *Mathews v. Lucas*, 427 U.S. 495 (1976)).

Guidelines, or to require that SEPTA's studies conform to specific validation criteria, provides an independent justification for reversing the District Court.¹⁹

19. The District Court found that, in developing its physical abilities testing, SEPTA's expert had "applie[d] criterion-related, construct and content validation strategies." FOF ¶ 15. However, it is clear from the record that no real attempt was made to establish either criterion or construct validity for SEPTA's test because no empirical data was submitted to show the required correlation between tested running times and ultimate job success. The only attempt to establish a correlation to actual job performance was an arrest analysis prepared by Dr. Siskin. That analysis neither encompassed a representative spectrum of SEPTA transit officer job duties nor evidenced any unsatisfactory performance by those officers failing to meet the cutoff. Compare, e.g., *Guardians Association of New York City Police Dept., Inc. v. Civil Service Commission*, 630 F.2d 79, 92 (2d Cir. 1980) (construct validation requires criterion-related study, which in turn requires "a demonstration from empirical data that the test successfully predicts job performance") (citing EEOC Guidelines); *Berkman v. City of New York*, 705 F.2d 584, 588 (2d Cir. 1983) (rejecting proffered justification of discriminatory selection procedure where city "failed to produce . . . [inter alia] 'empirical data demonstrating that the selection procedure [was] predictive of or significantly correlated with important elements of job performance' . . . (called 'criterion validation')"); *Williams v. Ford Motor Co.*, 187 F.3d 533, 539-41 (6th Cir. 1999) (criterion studies examine "whether performance on [a] test adequately correlates with performance on the job"); *United States v. City of Chicago*, 573 F.2d 416, 426 n.10, 427 (7th Cir. 1978) (fire department was required to show that promotional examinations were "predictive of successful performance in the jobs being tested for"; where there was "no attempt to determine from a job analysis what traits are necessary for job performance", tests at issue "fail[ed] to demonstrate either criterion validity or construct validity because they fail[ed] to predict performance . . .") (citing EEOC Guidelines); *Melendez v. Illinois Bell Telephone Co.*, 79 F.3d 661, 669 (7th Cir. 1996) (test invalid where there was no correlation between performance on challenged test and performance in job for which test was given, and little or no support for validity of test in predicting core areas of job performance); *Firefighters Institute For Racial Equality v. United States*, 549 F.2d 506, 510-11 (8th Cir. 1977) ("Criterion-related studies involve the correlation of job performance with success on an examination.").

VII.

Given all that I have set forth, I do not think there is any need to dwell on the specifics of the various statistical studies or methodologies that have been summoned to support the 42.5 mL/kg/min cutoff. After all has been said and done, after all the studies have been tabulated, all of the regression analyses analyzed, and all of the numbers fed into all of the number-crunching software on all of the computers, on all of the experts' desks; one unassailable fact remains. The 42.5 mL/kg/min aerobic capacity is not required of transit officers before or after they begin policing. Rather, it is used to disqualify applicants even though the application may be submitted as much as two and one-half years before graduation from the police academy and employment as a police officer. Moreover, as I have noted at length above, once officers actually begin protecting SEPTA's riders and property, they do not have to demonstrate or maintain *any* minimum aerobic capacity. Yet, SEPTA insists upon arguing that this aerobic cutoff has a very precise, demonstrable, correlation to public safety, and that it therefore supports the gender discrimination that SEPTA admits results from it.

Even though I do not think it necessary to attempt a statistical exegesis to illustrate why SEPTA has not demonstrated a business necessity for using the cutoff as an application requirement, I can not conclude without commenting on one particularly fascinating (and perhaps illuminating) study that SEPTA commissioned to support the 42.5 mL/kg/min cutoff. Although my colleagues criticize an analysis that "poke[s] a hole here or there in one or more of the District Court's extensive findings of fact and conclusions of law[.]" Majority Op. at 4, I can not allow the sheer extent of the District Court's opinion to substitute for the proof required under *Griggs*. Accordingly, I must take one specific "poke" that I believe exposes the hole that SEPTA's business justification should have fallen through.

As noted above, SEPTA wanted to determine the aerobic capacity of "perpetrators" that transit police might have to pursue in order to determine the validity of the 42.5 mL/kg/min aerobic cutoff. As part of its inquiry it attempted to measure the aerobic output necessary to

apprehend the average perpetrator. Of course, any good study begins with a representative sample, and the perpetrator study was no exception. Accordingly, Dr. Davis recruited a group of 31 subjects in order to get an idea of the aerobic capacity necessary for transit police officers to chase and apprehend them. However, there was one little problem with the study. It seems that Dr. Davis recruited his subjects at the University of Maryland track while the University's NCAA Division I track team was there.

The resulting "sample" contained 9 University of Maryland track "stars," 10 high school track "stars," 2 other college athletes, and 2 high school running backs who had been accepted to play football at major colleges.²⁰ After plaintiffs' counsel complained about the skewed sample, the results for the 9 members of the University of Maryland track team were factored out, and the results were analyzed without including their times. This reduced the sample to 22 subjects of whom "only" 14 (a mere 64%) were trained athletes, and four were nonathletes.²¹ Moreover, the high school runners were not your average high school "track star." The group remaining after "correcting" for the members of the University of Maryland track team included high school runners who were "winners of regional, state and sometimes *national* races." J.A. IV at A-916 (emphasis added). Dr. Landy characterized their times as "staggering." His reaction to suggesting that the aerobic performance of such a group represented the

20. It may stretch the limits of judicial notice, but I would ask the Majority to accept as fact that major college running backs are generally not noted for an absence of speed, quickness, or aerobic capacity.

21. The plaintiffs were unable to get any information about the remaining four members of the group. Dr. Landy, one of plaintiffs' experts testified about the information he was able to gather pertaining to the 22 subjects who had been contacted.

[W]e know that of the 22, 10 are high school track members but not just track members as if they just wander out every afternoon and run a little bit. These 10 are acknowledged track stars. . . .

J.A. Vol. IV at A-917. He further explained that one of the two football players "has been recruited by the University of Georgia and will go there. He is a running back and a linebacker and a defensive end." *Id.*

population that SEPTA police might have to chase was appropriate. He stated: "[i]t just stretches anything that I could imagine, to imagine that 22 of 31 SEPTA perpetrators are world class athletes." *Id.* at A-917.

Although this is not your usual "representative sample," it was representative enough to form part of the justification for the 42.5 mL/kg/min cutoff. Incredibly, the only justification Dr. Davis offered for basing his study on such an athletically talented group of trained athletes was that athletes are part of the general population and therefore part of the population that uses SEPTA. See Appellants' Br. at 38.²²

It is difficult to imagine a more graphic demonstration of what can result when studies seek to justify a standard rather than objectively define one. Yet, that is what happened here.²³ Of course, this, by itself, is not fatal to the 42.5 mL/kg/min cutoff. It is, however, illustrative of the concern we expressed in *Lanning I*, and it is also illustrative of why the current cutoff, though perhaps "reasonable" in the eyes of some, can not be seen as "necessary" when viewed through the lens of Title VII.

VIII.

Prior to today's decision, it was established in this Circuit, as it remains established in others, that a job

22. Though the record does not include any objective study of the ridership of SEPTA, I have resided in Philadelphia for over a quarter of a century, and I have serious doubts that approximately 64% of the people using public transportation in Philadelphia so closely resemble regional and national track stars, and running backs from major colleges. Yet, that is the justification that is offered for considering this study.

23. I do not doubt that Doctors Davis and Henderson are very learned experts who are knowledgeable and respected in their respective fields. I am nevertheless troubled by their propensity or willingness to include college athletes — indeed, superior athletes from Division I schools — in a study of aerobic capacity of the *average* perpetrator using public transportation in Philadelphia. I am even more troubled by the extent to which it suggests the defensive nature of SEPTA's attempt to justify its cutoff, and the District Court's acceptance of it.

requirement that has a disparate impact based upon gender could only be upheld if the relationship between the discriminatory requirement was so closely related to the essential of a given job that it could be justified as a business necessity. Today, in upholding a discriminatory application process based only upon a colorable claim of business necessity, we retreat from that standard while purporting to apply it. Yet, in enacting the Civil Rights Act of 1991, I believe Congress meant exactly what it said; discrimination in the name of "business necessity" must truly be *necessary*. No such necessity has been established here.

SEPTA can not "have its cake and eat it too." The aerobic cutoff is either so vital to public safety that it is a business necessity; in which case, SEPTA can not require it only of applicants without showing that the failure to require it of incumbents has had a negative impact on their performance as police officers. If, on the other hand, the relationship between public safety and the 42.5 mL/kg/min cutoff is as tenuous as suggested by SEPTA's failure to consider it when making job offers, promotions or commendations, then SEPTA can not continue to use justifiable concerns about public safety as a boogeyman to support the admittedly discriminatory cutoff it uses to screen applicants. Accordingly, I must respectfully dissent from the decision of my colleagues.

A True Copy:
Teste:

*Clerk of the United States Court of Appeals
for the Third Circuit*