

2000 WL 1790125
United States District Court, E.D. Pennsylvania.

Catherine Natsu LANNING, et al., Plaintiffs,
v.

SOUTHEASTERN PENNSYLVANIA
TRANSPORTATION AUTHORITY, Defendant.
UNITED STATES OF AMERICA, Plaintiff,

v.
SOUTHEASTERN PENNSYLVANIA
TRANSPORTATION AUTHORITY, Defendant.

Nos. CIV.A. 97-0593, CIV.A. 97-1161. | Dec. 7,
2000.

Opinion

FINDINGS OF FACT AND CONCLUSIONS OF LAW

NEWCOMER.

I. BACKGROUND

*1 1. This case is on remand from the United States Court of Appeals for the Third Circuit, which held in *Lanning V. SEPTA*, 181 F.3d 478 (3rd Cir.1999), *cert. denied*, 528 U.S. 1131, 120 S.Ct. 970, 145 L.Ed.2d 840 (2000), that this Court misapplied the business necessity standard enunciated in *Griggs v. Duke Power Co.*, 401 U.S. 424, 91 S.Ct. 849, 28 L.Ed.2d 158 (1971) and codified in the Civil Rights Act of 1991.

2. In its prior opinion, this Court held that SEPTA's use of a 1.5 mile run in 12 minutes, which corresponds to an aerobic capacity of 42.5 ml/kg/min, as a pre-screening device for new applicants was readily justifiable as a business necessity, despite the disparate impact such a standard had on female applicants.

3. In *Lanning*, the Third Circuit enunciated a new standard for business necessity, and remanded the case to this court to determine if SEPTA has carried its burden of establishing that 42.5 ml/kg/min is the minimum aerobic capacity necessary to perform successfully the job of SEPTA transit police officer.

4. In remanding this case, the Third Circuit suggested that this court may wish to exercise its discretion to develop the record further in keeping with the new standard announced.

5. This Court did so exercise its discretion. It permitted a

brief discovery period, and held a supplemental trial on April 6-7, and April 10-12, 2000. From the outset, however, it was made clear to both parties that the Court would not disturb its prior factual findings in this case.¹

¹ While the Third Circuit's Opinion draws heavily from the factual findings of this Court, the Circuit does suggest several concerns it has about the validity of some of the studies relied on in this Court's Opinion, and encourages this Court to take a critical look at these studies, if necessary, on remand. *Lanning* at 493, n. 21. The Court indeed taken a second critical look at these studies as suggested and once again reaffirms their validity consistent with its prior extensive 160 plus page memorandum opinion.

Although the parties' and the Court are intimately familiar with this Courts' prior findings, the Court will present several of these findings here to place this opinion in context, and enhance its readability.

6. In January 1989, Howard Roberts was hired by SEPTA as the Deputy General Manager. As the Deputy General Manager, Mr. Roberts was entrusted with managing the SEPTA Transit Police Department.

7. Shortly after his arrival in 1989, Mr. Roberts became aware of significant problems with the SEPTA Transit Police Department. Most notably, Mr. Roberts noticed that the SEPTA Transit Police Department was unable to control crime on SEPTA property and that problems existed with the physical fitness and capabilities training of its transit police officers.

8. At the time Mr. Roberts arrived at SEPTA, there were no physical fitness standards or physical training programs in place for SEPTA officers. As a result, there were instances where officers were injured, and there were numerous cases of police brutality that were caused by officers retaliating against persons who had previously assaulted physically unfit police officers.

9. Mr. Roberts noted that "crime statistics were very, very bad, officers for the most part arrived at crimes after they had taken place and basically did reports and turned them in." In essence, the SEPTA Transit Police Department was not preventing crime, rather it was merely reporting crime that occurred on SEPTA property.

10. In response to these problems, SEPTA initiated a complete overhaul of the police department under the direction of Mr. Roberts; its goal was 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. SEPTA increased the number of officers from 96 to nearly 200 and introduced a “zone concept” for the area they patrolled.

*2 11. The officers are deployed alone and on foot. When manpower permits, the beats are assigned in overlapping fashion to minimize the distances that officers will have to run to effectuate “officer backups” and “officer assists.” Absent full availability of all zone officers, officer backups or officer assists routinely come from two stations away. There is usually one vehicle patrolling in each zone. However, because of the age of the vehicles and because of other uses of the vehicle, such as the transporting of prisoners, foot patrol officers cannot rely on backup coming from the patrol vehicle.

12. During their tours, SEPTA officers frequently respond to officer assist or officer backup calls. An officer assist call requires other officers to respond immediately to another officer’s call for assistance—the responding officers are expected to use any means to get to the officer requiring assistance. An officer backup call also requires other officers to respond to the officer requesting assistance; however, the officers responding to a backup call do not have to arrive as quickly as they would for an officer assist situation. In essence, an officer assist call indicates that an officer is involved in or about to become involved in a potentially hostile or life- or property-threatening situation.

13. SEPTA officers have only two means by which to respond to officer backup and officer assist calls: (1) ride a train to the location where help is needed, if a train is available; or (2) run to the location where assistance is needed. Backups are run as paced jogs. Assists are paced runs with the goal of maintaining enough reserve energy to engage in any necessary struggling at the location of the call. SEPTA averages about 4 running assist responses per zone per month. Over eight zones, this is approximately 32 running assists per month or approximately 380 running assists per year. SEPTA averages about 20 running backups per zone per month. Over eight zones, this is approximately 160 running backups per month or approximately 1,920 running backups per year.²

² With approximately 190 officers, this means that every SEPTA officer performs a running assist or back-up every month.

14. In 1991, SEPTA hired Dr. Paul Davis to develop and validate a physical fitness test. Dr. Davis is a preeminent expert in the field of physical fitness and employment testing, and he has designed numerous fitness tests for law enforcement agencies, fire departments, armed services personnel and other entities engaged in the protection of

the public.

15. In developing physical abilities testing, Dr. Davis uses a “research design approach,” and applies criterion-related, construct and content validation strategies.³ Dr. Davis believes that the rationale for physical abilities testing is to ensure that there is an appropriate match between the requirements of the job and the individual who is applying for that position.

³ Courts and the psychological profession generally recognize three validation studies: content validity, criterion-related validity and construct validity. *Washington v. Davis*, 426 U.S. 229, 96 S.Ct. 2040, 48 L.Ed.2d 597 (1976). See also *Uniform Guidelines on Employee Selection Procedures*, 29 C.F.R. § 1607, *et seq.*, (“*Uniform Guidelines*”). In general, test validation is the process by which it is determined whether the inferences that the employer draws from results on a selection device are appropriate and meaningful. That is, test validation attempts to determine whether (and the degree to which) persons who are selected by a test will be successful performers on the job, and whether those who are not selected would not have been successful performers on the job.

*3 16. Prior to SEPTA, Dr. Davis had experience with developing physical abilities tests for numerous police and fire departments, approximately 70 different organizations.

17. With regard to SEPTA, Dr. Davis was contacted by Dr. Louis Vanderbeek, the Director of Medical Programs for SEPTA, to develop a physical fitness program for SEPTA. Early in the project, Dr. Davis met with Judith Pierce, the Assistant General Manager of SEPTA, Ronald Sharpe, the Chief of the SEPTA Transit Police Department, and other SEPTA officials to understand exactly what SEPTA’s objective was with respect to developing a physical fitness test.

18. Based upon his meetings with SEPTA officials, Dr. Davis came to understand that SEPTA was trying to enhance the level of fitness, physical vigor and general productivity of its police force; SEPTA also wanted medical criteria from which it could make informed decisions regarding such issues as return to duty, hiring and retirement. From these interviews, Dr. Davis also discovered that crime was rampant on the SEPTA system and that there were questions about safety for the ridership of SEPTA. Davis further learned that SEPTA wanted to remedy this situation and that SEPTA believed that improving the physical fitness of its police force was one of the best methods to achieve such a goal.

19. What distinguishes the essential tasks or functions required of a SEPTA transit officer from the essential

tasks required of police officers from other law enforcement agencies is that all of the activities take place on foot; therefore, the expectation is that SEPTA officers will have to move, run and walk with a higher degree of frequency on a daily basis more than other law enforcement officers. Dr. Davis found that a SEPTA officer would need a “sound, intact, disease-free cardiovascular system” to effectively perform their job. Dr. Davis testified that having such a cardiovascular system translates into aerobic capacity.

20. Dr. Davis also consulted Subject Matter Experts, who described the tasks involved in performing their duties as SEPTA transit officers. The SMEs then determined the relative importance of the tasks. Dr. Davis presented the SMEs with a scale that ranked the criticality of the particular physical task from one to five or six—one being the least critical and five or six being the most critical. Thus, the higher the score provided by a SME, the more critical the task was thought to be.

21. The tasks that were rated as either a one or two are not particularly consequential. Dr. Davis explained that a value of greater than three meant that the officers thought that the particular task was critical. In Davis’ validation study, jogging and running had values of 3.5; based upon the Delphi session, Dr. Davis’ opinion was that these tasks were the most critical tasks.

22. After computing the criticality rankings, Dr. Davis developed a scale regarding the frequency of performance of the tasks. A task which was performed daily was scored as a one; the performance of tasks that occurred weekly was a two; tasks done monthly were scored as a three; yearly tasks were scored a four; and a score of five indicated that the task was rarely performed.

*4 23. Based upon a review of the scales used, Dr. Davis testified that there was a value of greater than five on swimming because the group basically did not do that task. In contrast, jogging received a score of 1.7 that means the SMEs expected jogging to take place almost on a daily basis.

24. Based on a review of the frequency and criticality rankings, Dr. Davis concluded that SEPTA officers walk with high frequency because the officers are predominantly foot-based. Dr. Davis also correctly concluded that SEPTA officers run more frequently than other police departments; he also found that they sprinted more often. In addition, Dr. Davis found SEPTA officers used a baton with more frequency than in other jurisdictions. Overall, Dr. Davis assessed that the SEPTA officers are a more mobile and dynamic law enforcement group than most other law enforcement agencies.

25. Dr. Davis testified that typical law enforcement officers simply do not engage in the type of activities with

the same frequency as a SEPTA officer. The Court credits this testimony as being accurate. Indeed, the evidence introduced at trial establishes that SEPTA transit officers engage in physical activity more frequently than other law enforcement agencies.

26. The SMEs stated that it was reasonable to expect them to have to run one mile in full gear in 11.78 minutes. Dr. Davis, however, rejected this information when creating the 1.5 mile run as a component of SEPTA’s physical fitness test because the pace that the SMEs established was too low in Dr. Davis’ opinion. Dr. Davis believed that this physical dimension estimate was low because if such a pace was established as a test, it would require an aerobic capacity that almost any person could meet. Thus, if you were to use this estimate as a component of a physical abilities test, this component of the test would have no utility because almost any person could satisfy this minimal requirement. Based on Dr. Davis’ experience and professional medical literature, Dr. Davis rejected this estimate as wholly unrealistic; the Court agrees with this assessment.⁴

⁴ It is not surprising that SME’s would set easily attainable physical goals, as they will subsequently have to attain those goals. Moreover, as these were all older, experienced officers, and as SEPTA management had already determined that the force’s overall physical fitness was lacking, the SME’s were hardly the best source for appropriate standards.

27. Indisputably, the rate at which an officer performs an activity will be a function of the personal fitness level of that officer and that officer’s work pace. The officer who has a high aerobic fitness level will have a greater energy reserve once she arrives at the location of an officer assist or backup call and is going to be able to do something more proficiently vis-a-vis the other officer with a low aerobic capacity who was trying to maintain a pace for which he cannot supply oxygen on an ongoing basis.

28. Ms. Pierce specifically told Dr. Davis that she did not want the SEPTA police department to become the “boneyard” of the Philadelphia Police Department. Mr. Davis understood that Ms. Pierce was not concerned with having a standard that might be perceived as difficult for women to achieve; the job relatedness of the mission came first. In essence, SEPTA wanted to hire individuals who could perform the physical tasks required of a SEPTA officer regardless of whether this person was a man or woman; the Court finds that there certainly is nothing invidious about this goal.

*5 29. Nevertheless, it was Dr. Davis’ opinion that setting an aerobic capacity requirement in a range of 48 to 50 mL/kg/min would have an adverse effect on women because normative data demonstrates that there is a fairly

substantial difference in terms of oxygen uptake and metabolism capabilities on the part of women as compared to men. Based on the normative data, Dr. Davis believed that a standard of 48 to 50 mL/kg/min would present a fairly substantial obstacle for women to seek employment with SEPTA.

30. Consequently, because Dr. Davis believed that the goals of SEPTA could be satisfied by using a 42.5 mL/kg/min standard for aerobic capacity, and because this standard would substantially reduce the adverse impact of a 50 mL/kg/min standard, Dr. Davis recommended to SEPTA that it set its aerobic capacity requirement at 42.5 mL/kg/min.

31. Dr. Davis felt that women could attain a standard of 42.5 mL/kg/min. Dr. Davis based this opinion on a project his company did for St. Paul, Minnesota, in which applicants for the fire department had to successfully run one and one-half miles in 11 minutes and 40 seconds. The aerobic capacity required to complete this run is 45 mL/kg/min. The outcome of the run was that out of the 705 individuals who applied for employment, 585 males and 120 females, 80% of the men passed and 76% of the women passed.

32. Because Dr. Davis wanted to test for an aerobic capacity of 42.5 mL/kg/min, Dr. Davis suggested that SEPTA implement a distance running test whereby applicants would be required to run 1.5 miles in 12 minutes or less. Dr. Davis suggested this distance and time because if an applicant could complete the run in 12 minutes or less, it could be concluded that the successful applicant had an aerobic capacity of at least 42.5 mL/kg/min.

33. In the course of creating the physical abilities test for SEPTA, Dr. Davis was able to link aerobic capacity to the specific critical tasks that he observed SEPTA officers doing on the job. Dr. Davis testified that the link is common sensical in that every job task analysis that has ever been done for any reasonably proactive law enforcement organization finds that running is a critical and essential task. Also, statistical manipulations have been established showing that there exists a correlation between police officer performance and a 1.5 mile run.

34. In sum, the Court finds that Dr. Davis demonstrated that an aerobic capacity of 42.5 mL/kg/min is necessary to successfully perform the functions of a SEPTA transit officer.

35. Scientific studies show that males score higher on tests of VO₂ max and endurance performance than their female counterparts due to physiological differences between men and women. This result is attributable to the well-documented sex differences in body composition and hemoglobin, the iron-containing compound in the blood

responsible for oxygen transport because men have more muscle mass and less fat per unit of body weight than women. The most important factor determining one's capacity for oxygen consumption during exercise is the quantity of muscle mass a person possesses; this is because the site of aerobic metabolism occurs in the active muscles. It is partially because of this difference in the amount of potentially active muscle mass during exercise that men consistently score higher in VO₂ max tests like the 1.5-mile run test administered by SEPTA.

*6 36. Since 1991, SEPTA policy has required that incumbent sworn employees of all ranks in SEPTA's Transit Police Department take and pass a physical fitness test every six months. Despite this policy, there was evidence introduced at trial that incumbents are not always retested every six months.

37. The incumbent physical fitness testing program is based upon the same study relied on by SEPTA for its applicant physical fitness testing program. The components of SEPTA's physical fitness test for applicants that are being challenged in this case are identical to the components of SEPTA's physical fitness test that have been administered to incumbent SEPTA transit police officers since 1991.

38. When incumbent testing was first introduced, SEPTA would discipline incumbent officers for failing to meet their interim goals. However, the patrol officers' union objected to such discipline, claiming that the disciplinary component of SEPTA's physical fitness testing was never the subject of collective bargaining, and thus SEPTA could not unilaterally implement such testing. The union took SEPTA to arbitration over this matter and won. Thus, due to the opposition of the patrol officer's union, SEPTA was precluded from disciplining the patrol officers who failed the incumbent testing.

39. Because SEPTA was unable to discipline officers who failed incumbent fitness testing, Chief Evans attempted to gain compliance with the incumbent fitness standards by offering an incentive whereby officers would receive \$50.00 each time they passed their interim fitness goals, with a maximum of \$200.00 per year. SEPTA additionally offered to reimburse officers for gym memberships. This incentive program for incumbent officers was implemented with the union's concurrence.

40. Given that SEPTA does not have the ability to discipline its incumbents who fail to meet interim fitness goals set by SEPTA, Chief Evans believes that those few officers who repeatedly fail their incumbent testing do so because of a lack of effort, desire or motivation. Chief Evans has elected not to impose discipline on supervisors because he does not believe that half of the police department should be treated differently than the other half—the transit police officers who he cannot discipline.⁵

⁵ 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.

41. Although SEPTA has never taken any steps to determine whether the incumbent officers who have failed the physical fitness test have adversely affected SEPTA's ability to carry out its mission, Chief Evans testified that officers who are not passing their incumbent fitness examinations are not capable of performing all of their policing duties and that a lack of fitness and inability to meet fitness standards has resulted in on-the-job injuries. For example, Chief Evans testified to an incident where a SEPTA officer, who was not meeting her interim fitness goals, was thrown into the track area of a train station by an intoxicated individual. Chief Evans believes that her lack of fitness contributed to her being thrown onto the tracks.

*7 42. Since the implementation of this fitness program, Part I felony offenses, *i.e.*, homicide, rape, robbery, aggravated assault, burglary, theft and auto theft, are down by approximately 70%. Lt. Maslin believes that the fitness program has contributed to this reduction in crime.

43. After this litigation commenced, SEPTA retained statisticians, Bernard Siskin, Ph.D., and David Griffin, Ph.D., to submit expert reports which examine the statistical relationship between the components of SEPTA's physical fitness test on the one hand and the number of arrests and "arrest rates" on the other.⁶

⁶ Dr. Siskin testified at trial as to the results of the studies and reports and the opinions expressed therein. Dr. Griffin only testified as to some of the underlying data.

44. Dr. Siskin's testimony showed, when comparing officers who were always at 42 mL/kg/min or over to officers who were always under 42 mL/kg/min, the higher aerobic capacity group had a 57.1% "arrest rate" advantage in the more serious Part I crimes and 28% greater arrest rate for all offenses. Dr. Siskin also pointed out that the data showed that officers always at 42 mL/kg/min or above made three times (151%) the actual number of Part I arrests and 75% more actual overall arrests when compared to officers who never met the 42 mL/kg/min requirement.

45. Under his regression analysis, Dr. Siskin demonstrated that for the period of 1991 through 1996, SEPTA could have achieved 470 additional arrests—70

of which were Part I arrests for serious crimes—if the aerobic capacity of all the officers was 42 mL/kg/min or above for this time period. These findings reflect a 10% increase in Part I arrests and a 4% increase in the overall arrest rate. This analysis was based on a regression analysis that took into account all relevant variables, including rank, zone and tour and assignments to special units. Dr. Siskin testified that taking these variables into account, the statistical relationship and predictive nature of aerobic capacity remained significant and demonstrates that meeting SEPTA's aerobic capacity standard of 42 mL/kg/min consistently predicted higher arrests and arrest rates for Part I offenses.

46. Dr. Siskin also studied 953 perpetrators who had been arrested for committing Part I crimes in order to determine their aerobic capacity. The analysis was based upon the sex, race and age of the perpetrators. Dr. Siskin utilized a study (the "Vogel Study") provided by one of defendant's experts, Dr. Moffatt, in order to develop a statistical prediction of the aerobic capacity levels of the 953 perpetrators who were apprehended during the years 1991–1996. Based on his analysis, Dr. Siskin was able to provide an estimate of the aerobic capacity of the 953 perpetrators who were caught or apprehended. The mean age of the arrested perpetrators was 26.3 yrs.

47. Dr. Siskin's analysis showed that 51.9% of the perpetrators were estimated to have an aerobic capacity of 48 mL/kg/min, and only 27% of the perpetrators were estimated at or below 42 mL/kg/min.

*8 48. Dr. Siskin also conducted a study of the aerobic capacity of the SEPTA officers that apprehended perpetrators of Part I crimes in the SEPTA transit system.

49. Dr. Siskin studied 382 Part I arrests for the period of 1994–1996. Dr. Siskin found that the arresting SEPTA transit police officers maintained a mean aerobic capacity of 46.8 mL/kg/min; whereas, the aerobic capacity of the SEPTA transit patrol officer population was approximately 43.9 mL/kg/min. The aerobic capacity of the SEPTA transit police officers who apprehended the Part I criminals during the years of 1994 through 1996 was found to be statistically significantly higher (at the 0.01 level) than the general SEPTA patrol officer population. Furthermore, 94% of the arresting patrol officers in this study maintained an aerobic capacity that exceeded 42 mL/kg/min. Only SEPTA patrol officers who made arrests were studied. Therefore, of 382 possible matches between a perpetrator and an arresting officer, there were 281 cases of SEPTA transit patrol officers making the arrests.

50. Subsequent to the filing of the *Lanning* administrative charges with the PHRC and the EEOC, SEPTA retained Robert Moffatt, Ph.D., an exercise physiologist, to defend SEPTA's physical fitness test.

51. During his two tours of the SEPTA system, Dr. Moffatt observed dramatic differences between the job duties of a SEPTA officer and those of other law enforcement officers with whom he had worked—the Citrus County, Florida Sheriff's Office and the Metropolitan Dade County, Florida Sheriff's Office. Dr. Moffatt noted that the SEPTA transit police force is predominately on foot patrol and arrives at various locations on foot. The SEPTA officers patrol alone and traverse a wide number of steps during their shifts.

52. In interviews with the SEPTA officers, Dr. Moffatt was told that one of the critical tasks of a SEPTA officer is running from one station to the next for officer assist calls. The officers also told Dr. Moffatt that they had to be prepared to fight or subdue a perpetrator upon arrival. Because this scenario was deemed a critical task, Dr. Moffatt decided to test for the amount of aerobic capacity that would be necessary to successfully engage in this task.

53. Dr. Moffatt wanted to determine through a simulation of a typical SEPTA backup/assist call how long it would take the officers to run from point A to point B. Protocols were devised for the testing of SEPTA transit police officers from which Dr. Moffatt could establish a pace for use in laboratory testing.

54. From the simulations in Philadelphia, Dr. Moffatt was able to establish an average assist response pace of 187 seconds. Laboratory simulations were then setup with a treadmill and a bench stepping device where Dr. Moffatt could control the work performed and measure the amount of oxygen consumed, as well as the energy expenditure for that work. Dr. Moffatt made sure that the laboratory simulation modeled the concourse that was run in Philadelphia with respect to the distances, angles and number of steps.

55. Based on his studies, Dr. Moffatt believes that SEPTA's aerobic capacity standard of 42.5 mL/kg/min as it relates to transit police officer work is very conservative. Indeed, Dr. Moffatt believes that the aerobic capacity cutoff for SEPTA transit police officers should be 45 mL/kg/min.

*9 56. The practical significance of Dr. Moffatt's studies is that a SEPTA transit police officer with an aerobic capacity less than 45 mL/kg/min has to run 3–5 blocks working at maximal effort and may not arrive in a reasonable time period, and if they do arrive in a timely fashion, their ability to do anaerobic work drops off so significantly that they may be ineffective upon arrival.

57. Dr. Henderson, another of SEPTA's experts, contends that there has historically been difficulty in using performance evaluations as a criterion for measuring

police work due to potential bias that may exist in such subjective evaluations.

58. Dr. Henderson testified that the fact that incumbent transit police officers have failed incumbent aerobic capacity tests or muscular strength and endurance tests is irrelevant to the validity of the test developed as a selection device. Dr. Henderson testified that using incumbents as a benchmark to determine whether a selection device is valid is dangerous for several reasons. Initially, a selection device is not designed to be an absolutely perfect predictor for all members of a company. Also, the incumbent argument incorrectly assumes that the incumbent population will necessarily match the applicant population. Incumbents are generally older individuals than those who a selection device is being used on for new hiring. Moreover, a second fallacious assumption is that the incumbent population is performing well; admittedly, there will be considerable variation in effectiveness of workers already on a job. Generally, applicants train for a test where incumbents will basically walk in and take a test without any preparation. Therefore, in Dr. Henderson's opinion it is risky to use incumbent data as a benchmark for establishing entry-level selection devices.⁷

⁷ Of particular significance to the Court counseling against relying on incumbent data to set entry standards is the fact that a stated goal of SEPTA management is to increase the fitness level of its force. It would be absurd for SEPTA to set physical fitness standards designed to increase the fitness level of their force based on admittedly unfit officers' performance.

With those already established facts serving as a background, the Court now turns its attention to the evidence adduced at the supplemental trial. After considering the testimony of the witnesses, the admitted exhibits, the arguments of counsel, the parties' post-trial submissions, the Court's prior factual findings, and the standards as set forth by the Third Circuit, the Court makes the following findings of fact and conclusions of law.

II. FINDINGS OF FACT

59. SEPTA's experts conducted additional studies which focused on officer pursuits and running-related critical tasks that SEPTA's aerobic capacity test was designed to test applicants on: the running tasks included the ability to pursue and apprehend fleeing suspects; the ability to track and identify fleeing suspects in order to aid in the apprehension of fleeing criminals; the ability to respond, on foot, in an acceptable time to emergency assist calls from officers; and the ability to perform arduous tasks once the officer either apprehends the suspect or arrives at

the scene of an emergency assist. SEPTA's additional studies were carried out in order to assess whether individuals failing SEPTA's 1.5 mile running test were capable of meeting minimum performance requirements for the essential tasks required of SEPTA transit police officers.

60. SEPTA's expert, Dr. Henderson, also reviewed SEPTA's prior evidence to determine whether at the initial trial of this matter SEPTA had demonstrated that its running test measured the minimum aerobic capacity required of SEPTA officers. Dr. Henderson opined that two lines of evidence in the original trial clearly demonstrated that SEPTA had already proven that its aerobic capacity test was set at the minimum aerobic capacity required to perform the transit officer position. The evidence that Dr. Henderson noted included the comprehensive arrest data and analyses conducted by Drs. Siskin and Griffin, and the physiological evidence presented by Dr. Moffatt.

*10 61. Dr. Henderson testified that the arrest rates and actual arrests of those who passed SEPTA's aerobic capacity test were markedly and significantly higher than those that failed SEPTA's test. Likewise, Dr. Henderson noted that Dr. Moffatt's evidence concerning the precipitous decline in work capacity for those individuals who could not meet SEPTA's aerobic capacity test clearly showed that SEPTA's aerobic capacity test was set at the minimum aerobic capacity required to perform the SEPTA transit officer job.

62. Furthermore, Dr. Henderson noted that the original evidence of Dr. Siskin, which compared the performance of officers *always* above 42.5 mL/kg/min and officers *always* below 42.5 mL/kg/min, showed a significant difference in arrest rates, actual arrests, commendations and a number of other objective field measures from the SEPTA data. Dr. Henderson testified that the field performance differences of officers always at 42.5 mL/kg/min and those never at 42.5 mL/kg/min demonstrated that SEPTA's aerobic capacity test was already set at the minimum aerobic capacity required of SEPTA transit police officers.

63. Furthermore, Dr. Henderson noted that SEPTA's loss of approximately 470 arrests (due to some officers failing to maintain the minimum aerobic capacity) was further evidence that SEPTA was already at the minimum aerobic capacity required to successfully perform the transit officer position.

64. Thus, it was Dr. Henderson's opinion that the initial evidence demonstrated that SEPTA had already proven that it was testing for the minimum aerobic capacity required, and this was borne out, in part, by the loss of the 470 additional arrests in the transit system.

65. It is axiomatic that such a loss in arrests is a significant threat to the public safety.

66. Dr. Henderson also re-analyzed some of Dr. Moffatt's data from the first trial in a more detailed and elaborate manner than had been originally offered at the first trial. Dr. Henderson undertook this re-analysis in order to determine whether any of that data further addressed the issue of whether 42.5 mL/kg/min was the minimum aerobic capacity required of SEPTA transit officers; in other words, was there more to be yielded from the research that Dr. Moffatt undertook.

67. Dr. Henderson's re-analysis of Dr. Moffatt's data examined the comparative success of four groups of subjects who differed on aerobic capacity levels. Those individuals below SEPTA's aerobic capacity cut-point only had a 33% probability of arriving at an emergency assist call in a timely manner. In striking contrast to those individuals who were below SEPTA's aerobic capacity cut-point, individuals meeting SEPTA's aerobic capacity requirement had an 80% to 90% success rate in carrying out an emergency assist; this finding was consistent with respect to an even more relaxed time standard of 200 seconds for carrying out an emergency assist.

68. Dr. Henderson opined that those individuals failing SEPTA's test would not be capable of carrying out critical emergency assists. Dr. Henderson testified that the success rates of those individuals below 42.5 mL/kg/min was far too low for the requirements of the SEPTA transit officer position on various job standards.

*11 69. Thus, re-analyzing Dr. Moffatt's data, Dr. Henderson compared the pass rate for those individuals who passed SEPTA's aerobic capacity test to the pass rates of those individuals that failed SEPTA's aerobic capacity test on five job standards: 1) percent of group meeting the 188 second assist standard; 2) percent of group meeting a 200 second assist standard; 3) average post-run work output; 4) average percent decrease on post-run work output; and 5) percent of group exceeding 740 kg. of work output.

70. Dr. Henderson opined that the data demonstrated consistently that 80% of those passing SEPTA's aerobic capacity test could meet each of these minimum job standards, whereas, at best, only 33% of those failing SEPTA's aerobic capacity test could meet minimum job requirements. Dr. Henderson testified that these marked differences demonstrated that those failing SEPTA's aerobic capacity test could not meet minimum job standards reflecting essential transit officer functions.

71. Dr. Henderson testified that re-analysis of Dr. Moffatt's original data showed that individuals failing Dr. Davis' test are unable to meet minimum standards and their ability to perform job criteria are drastically low and

a threat to the system.

72. Dr. Henderson testified that after conducting the re-analysis of Dr. Moffatt's original data, he extended this research to pursuit activities that were required of SEPTA officers. Dr. Henderson and Dr. Davis conducted a further study to determine how well the 12 minute cut-off point distinguished between people who were capable of apprehending or tracking perpetrators, and what success rate SEPTA could expect from those that passed its aerobic capacity test, compared to those that failed its aerobic capacity test.

73. To implement this study, Drs. Davis and Henderson designed a course that was similar to what Dr. Moffatt had originally designed, since both station-to-station emergency assists and many pursuit runs require SEPTA officers to run at least 3–5 blocks.

74. Drs. Davis and Henderson recruited 86 test subjects from the area surrounding the University of Maryland. The first group of subjects was representative of applicants who apply to become SEPTA officers. The applicant subjects were recruited from a population age-matched to SEPTA's applicant pool, with an average age of 26 years.

75. SEPTA's experts then determined the aerobic capacity of each of the 86 subjects by timing each of the 86 individuals on a 1.5 mile run. The 1.5 mile run took place on day one of their research. On day two, the 86 applicant subjects were tested on a .35 mile pursuit course wearing a 20 pound vest to simulate the total weight that SEPTA officers carry in the course of their duties. Drs. Henderson and Davis recorded the time that it took each of the 86 individuals to complete the .35 mile pursuit course. At the conclusion of the .35 mile pursuit course, each individual was required to drag a dummy weighing 175 pounds a distance of 30 feet. The time it took each subject to drag the dummy 30 feet also was recorded.

76. Thus, for each of the 86 applicant subjects, SEPTA's experts obtained their 1.5 mile run time; their aerobic capacity; the time it took for each of the simulated applicants to run the .35 mile pursuit course; and the amount of time it took to drag a dummy (weighing 175 pounds) 30 feet.

*12 77. From this data, Dr. Henderson determined that there was a high correlation between running SEPTA's 1.5 mile run in 12 minutes and the officer's ability to run the .35 mile pursuit course. The correlation coefficient was $r=.87$, and was highly statistically significant with a "p" value of .001.

78. Further, Dr. Henderson determined that there was a high correlation between successfully completing the 1.5 mile run in twelve minutes or less and short sprints, such

as one-tenth of a mile. The one-tenth of a mile run was representative of the distance an officer would run chasing a perpetrator out of a subway station, up a set of stairs and onto the street.

79. The correlation coefficient between passing the 1.5 mile run and success on short pursuits was $r=.81$. The pursuit course SEPTA's experts developed included a grade, turns and multiple steps the applicant subjects would have to run. The short pursuit distance replicated a crime occurring on a SEPTA platform and a pursuit by a SEPTA officer across the platform and up a set of steps.

80. Much was made at trial about the presence of several University of Maryland track runners and the impact said runners may or may not have had on the outcome of the study, or the study's reliability. Dr. Henderson calculated the correlation coefficient without the presence of any Maryland track team members in the simulated applicant pool. This recalculation showed that the correlation coefficient between completing the 1.5 mile run and the .35 mile pursuit time remained the same and, in fact, the correlation was even higher once the Maryland track team runners were taken out of the sample. Thus, the Court is satisfied that the integrity of the study was not compromised, and accepts its validity.

81. Drs. Henderson and Davis set minimum benchmarks on various job standards representing critical tasks to measure the performance of the simulated applicants in order to determine whether SEPTA's 1.5 mile running test was set at the minimum aerobic capacity required to perform the transit officer job. The benchmarks of minimum performance on various job standards were developed to assess the performance differences between the test subjects that passed SEPTA's 1.5 mile test and the test subjects that failed SEPTA's 1.5 mile test.

82. To develop pursuit benchmarks, Drs. Davis and Henderson recruited a group of 31 individuals who simulated fleeing suspects, or perpetrators. The simulated perpetrators ran the .35 mile pursuit course without the weight carried by the applicant subjects. The purpose was to develop a distribution of running times, and from those times, ascertain a minimum standard that a SEPTA officers would have to meet in order to pursue and apprehend fleeing suspects. Consequently, data was collected on 31 subjects that ran the .35 mile pursuit course without the burden of 20 pounds that SEPTA officers carry on a routine basis on their patrols.

83. The subjects who simulated the perpetrator group consisted of 25 males and 6 females, approximately 20 years of age. Although the perpetrator sample was somewhat younger (average age 20 years) than the mean age of actual SEPTA perpetrators (26 years), SEPTA's experts testified that the sample was representative with respect to aerobic capacity, since aerobic capacity is

essentially flat (unchanging) from age 8 to age 30.

160 seconds and 185 seconds.

*13 84. Consequently, SEPTA's experts testified that the age difference had no impact on the representativeness of the perpetrator sample. Moreover, SEPTA's experts, Dr. Henderson and Dr. Davis, noted that it was the younger segment of the perpetrator population that was most likely to run and, thus, the perpetrator group was taken from the segment of perpetrators most likely to initiate a foot pursuit.

85. From the collected data, SEPTA's experts determined that: (1) 50% of the perpetrator group completed the .35 mile pursuit course in less than 133 seconds; (2) 75% of the simulated perpetrator group ran the pursuit course in less than 163 seconds; and (3) 90% of the perpetrator group ran the pursuit course in 248 seconds or less.

86. Dr. Henderson testified that there is a wide and extensive body of literature demonstrating a significant relationship between arrest rates and suppression of crime. Authoritative studies have demonstrated that once a police department achieved a 20% arrest rate or higher, suppression and deterrence of crime was quite strong. Moreover, the higher the arrest rate, the more suppression and deterrence of crime was achieved. This direct and significant effect was found to be linear from 20% arrest rates through 50% arrest rates; simply put, the higher the arrest rates, the greater the suppression and deterrence of crime.

87. SEPTA's experts determined that a minimum acceptable standard for a pursuit run was the ability to apprehend the slowest 20% to 25% (bottom quarter) of fleeing suspects, *i.e.*, those simulated perpetrators running the .35 mile pursuit course in 163 seconds or slower. This number is directly related to achieving a 20% arrest rate, which is significant in deterring crime.

88. SEPTA's experts concluded that applicants without that capability to pursue and apprehend the slowest 20% to 25% of the fleeing suspects compromised SEPTA's transit police department's ability to effectively combat and deter crime. A benchmark of 165 seconds was set for the pursuit course. The applicant subjects in the study would have to complete the pursuit course in 165 seconds in order to meet a minimally acceptable pursuit standard. Thus, SEPTA's experts chose 163 seconds (rounded to 165 seconds) as a minimally acceptable running time on the pursuit course, based on the need to apprehend at least the slowest 20% to 25% of a fleeing perpetrator sample.⁸

⁸ Dr. Henderson further testified that it makes no difference where the cutoff score is set with respect to the pursuit criterion of the study. This is so because of the large discrepancy in success rate between the passing and failing group at all cutoff points between

89. SEPTA's experts also set an additional minimum benchmark that would allow a SEPTA transit officer to track a fleeing suspect and, thus, maintain visual contact in order to aid in the apprehension of a fleeing felon. Tracking time was set at 175 seconds, 10 seconds longer than what would be required to overtake a fleeing perpetrator from the slowest quarter of simulated perpetrator times. SEPTA's experts chose this as a reasonable tracking time, since ten seconds translated to approximately 45 feet of distance between the chasing officer and a fleeing criminal.

90. At the initial trial of this matter, this Court found that SEPTA officers routinely run after and pursue fleeing criminals and are required to apprehend the criminal, many times alone and without assistance. Officer pursuits typically span three to eight city blocks and may entail from 3 to 10 minutes of running followed by the need to engage in arduous activity.

*14 91. Regardless of the actual percentage of arrests that are a result of aerobic encounters, it is a significant part of the job of SEPTA transit officer, and considering the nature of the officer's duties, as described herein, these are critical tasks.

92. In sum, SEPTA's standard of 165 seconds would allow SEPTA officers to apprehend and catch only 20–25% of the slowest perpetrators—a minimally acceptable standard that would serve to deter and suppress crime. Similarly, a standard of 175 seconds was set as a minimally acceptable standard to maintain visual contact with the slowest 20–25% of the perpetrator population.

93. Both of SEPTA's experts testified that it would be desirable to catch 50% of the fleeing perpetrators, but in light of the aerobic capacity of the fleeing perpetrators (75% greater than 42.5), this was not feasible. At the first trial, SEPTA established that the aerobic capacity of the perpetrators in the SEPTA system averaged 48 mL/kg/min, significantly higher than SEPTA's aerobic capacity cut-point of 42.5 mL/kg/min. Furthermore, evidence from the first trial indicated that at least 25% of the perpetrators in the SEPTA system have an aerobic capacity greater than 48 mL/kg/min, and 75% of the perpetrators had an aerobic capacity greater than 42.5 mL/kg/min.

94. However, apprehending 20–25% of fleeing perpetrators provides a meaningful deterrence, even where the apprehension includes tracking and maintaining visual contact with the fleeing felon in order to effectuate the arrest.

95. Plaintiffs noted that nine members of the University of Maryland Track Team were part of the simulated perpetrator group of thirty-one individuals. Dr. Henderson testified that he removed the nine members and recalculated for the remaining twenty-two simulated perpetrators pertinent running times: (1) the mean (average) running time; (2) the 25th percentile running time; (3) the 50th percentile running time and the 75th percentile running time. Dr. Henderson produced these calculations and testified that the slowest quartile running time was, as before, approximately 163 seconds, precisely where the original cut point for the lowest 20th to 25th percentile was originally established.

96. Plaintiffs also complained that members of a high school track team were part of the perpetrator sample. Dr. Henderson testified that there was no significant difference in the pursuit course running time of high school track members that were part of the twenty-two simulated perpetrators. Plaintiffs presented no evidence or calculations which demonstrated that high school track members affected in any way the validity of the pursuit times identified in Drs. Henderson and Davis' study. Thus, the Court accepts the validity of the running times, and rejects plaintiff's attempts to discredit them.

97. SEPTA's experts, based on the average performance of the SME's, and their own expert judgment, set 185 seconds as the time to complete the critical emergency assist tasks. SEPTA's experts also set completing an emergency assist within 200 seconds as an extreme lower bound performance for this critical job task.

***15** 98. SEPTA's experts also set standards for dragging a dummy within 30—35 seconds; completing a pursuit run plus a dummy drag within 200 seconds; completing a pursuit run plus dummy drag within 214 seconds; completing a 165 second pursuit run and a 30 second dummy drag; and completing a 175 second pursuit run and 35 second dummy drag. Dr. Henderson testified that the combined pursuit run and dummy drag most aptly describes the job since once the officers arrive, arduous and dangerous police work may and does occur.

99. The benchmarks for the emergency assist runs in the Henderson/Davis study were derived from the data gathered by Dr. Moffatt for use at the original trial of this matter. Therein, Dr. Moffatt conducted a simulated assist run with eleven SEPTA SME's, and determined the average assist response pace to be 187.6 seconds. Dr. Moffatt's finding with respect to the average assist time was adopted by this Court following the original trial of this matter. The normative basis for the use of the dummy drag in the studies conducted by SEPTA's experts at the remand hearing stems from the fact that 75% of the individuals in the applicant pool were able to drag the dummy 30 feet in 30 seconds.

100. The Court finds no merit in plaintiff's attacks on the reliability of SEPTA's studies based on the presence of any athletes, as they were effectively removed with no impact on the results.

101. Notably, only 14% of the failing group could carry out the emergency assist in 185 seconds, whereas 84% of the passing group⁹ could meet this critical job task. By way of further comparison, when the emergency assist standard is set at 188 seconds, only 29% of the failing group can perform the task whereas 88% of the passing group is able to successfully accomplish this critical component of the SEPTA transit officer position.

⁹ The term "passing group" as used herein is defined as the group of individuals which completed the 1.5 mile run in twelve minutes or less, translating to an aerobic capacity of at least 42.5 ml/kg/min. "Failing group" is defined as those individuals that did not complete the 1.5 mile run in less than twelve minutes, translating to an aerobic capacity of less than 42.5 ml/kg/min.

102. Exhibit D-7 also contained a comparison between those that passed SEPTA's test and those that failed SEPTA's test on the combined tasks of running the pursuit course in 165 seconds and pulling a dummy in 30-seconds, and running the pursuit course in 175 seconds and pulling the dummy within 35 seconds. Like all other performance measures, those that passed SEPTA's aerobic capacity test significantly outperformed those that failed on these performances measures. As shown on D-7, those individuals passing SEPTA's aerobic capacity test were able to meet the combined tasks over 50% of the time, whereas the failing group could only meet these job standards 4% of the time.

103. Dr. Henderson testified that the performance scores of the passing and failing groups on Exhibit D-7 demonstrated that persons failing the SEPTA 1.5 mile run could not meet the minimum performance requirements of a SEPTA transit officer.

104. Defendants also introduced Exhibit D-40 which was a graphic depiction of the performance difference of those that were able to pass SEPTA's 1.5 mile run and those subjects that could not. Exhibit D-40 showed striking differences on each of the job standards; Dr. Henderson, as before, developed D-40 to exclude Maryland track team members. Once again the differences in performance between the passing and failing groups on SEPTA's aerobic capacity test were even greater. Exhibit D-40 depicts the extreme performance differences on critical job standards of those passing SEPTA's 1.5 mile run and those failing SEPTA's 1.5 mile run through a wide range of job standards.

***16** 105. Moreover, Dr. Henderson testified that the

performance differences between the passing groups and the failing groups were significantly different, and were not driven or determined by comparing high aerobic capacity individuals to low aerobic capacity individuals. Dr. Henderson testified that even if one took out the very low aerobic capacity individuals and the very high aerobic capacity individuals, the performance differences remained the same with the failing group unable to meet the performance standards required of SEPTA transit officers.

106. Thus, even if one looks at only the marginal passing and marginal failing groups, the differences between these closely defined groups is still statistically significant well below the maximum “p” value of .05 that the Uniform Guidelines recommend.

107. Dr. Henderson testified, and Exhibit D-43 showed, that even when one looked at the marginal passing group’s performance on the many job criterions and compared their performance to the marginal failing group’s performance, the marginal passing group still significantly outscored the marginal failing group on each of the job standards.

108. Nonetheless, Dr. Henderson testified that the proper analysis was to compare the performance of the entire passing group on SEPTA’s test with the entire failing group on SEPTA’s test. This was shown in Exhibit D-7.

109. Exhibit D-41, a series of bar graphs, also compares the performance of test subjects that passed SEPTA’s test to the performance of test subjects that failed SEPTA’s test. The bar graphs in D-41 unmistakably demonstrate that on 12 job standards, the failing group on SEPTA’s aerobic capacity test did not meet any of the minimum job standards required of SEPTA officers.

110. A review of Exhibit D-41 shows that the failing group’s success rate across 12 job standards ranges from approximately 5% to 20%, whereas, the passing group’s success rate ranges from 70% to 90% on the 12 job standards that were set at minimally acceptable performance levels.

111. Dr. Henderson testified that SEPTA’s most recent studies, as described in Exhibits D-7, D-40 and D-41, further demonstrated that individuals failing SEPTA’s aerobic capacity test cannot meet minimum job requirements that are essential to the successful performance of the transit officer position.

112. Further, Dr. Henderson testified that the failing group’s performance on the job standards was dangerously low, unacceptable and they would not be able to successfully do the job.

*17 113. Dr. Davis also testified that SEPTA’s recent

studies demonstrated both statistically and empirically that 42.5 was the minimum aerobic capacity below which individuals could not perform the essential functions of the SEPTA transit officer position. Dr. Davis testified that the applicant subjects below 42.5 were virtually incapable of meeting minimum job standards, whereas individuals meeting 42.5 were five-to-six times more likely to be able to perform the essential and critical tasks of the transit officer position. The Court agrees with both Drs. Davis and Henderson that individuals below 42.5 cannot perform the essential and critical tasks required of a SEPTA transit police officer and credits their testimony.

114. During their case-in-chief, plaintiffs called Michael O’Connor to testify. Mr. O’Connor is the former Chief of the New York City Transit Police. Mr. O’Connor’s testimony served to enlarge the already expansive list of reasons that SEPTA’s transit police department is unique from other police forces throughout the country, thereby rendering the use of physical fitness requirements of another police department impossible.

115. The Plaintiffs’ reliance on Mr. O’Connor’s testimony for the proposition that SEPTA’s aerobic capacity requirement is not necessary to perform the transit officer position duties is not credited by this Court. Mr. O’Connor conceded that he did not interview any of the 25 transit authorities surveyed by Dr. Landy. Dr. Landy’s survey showed that of the 25 transit authorities surveyed, 14 responded to the question “what is the approximate typical distance between stations they might have to get to on foot” with a completely negative response, *i.e.*, 14 of the 20 transit authorities indicated that their officers do not travel or patrol between stations on foot.¹⁰

¹⁰ Five of the twenty-five transit authorities surveyed by Dr. Landy provided no response to this question.

116. Furthermore, of the 6 remaining transit authorities responding to this query, none indicated that the distances between stations were nearly as far as the distances that SEPTA has to patrol on foot. As noted at the first trial, SEPTA officers are expected to run five to eight blocks in order to carry out a station-to-station emergency assist call. None of the 25 transit authorities surveyed came even close to requiring its officers to travel this distance on foot.

117. Thus, this Court finds that SEPTA, is, indeed, unique in this critical task as the Court had found in the original trial. Mr. O’Connor’s testimony that the transit police officer position is the “same” everywhere simply does not withstand analysis.

118. Similarly, Mr. O’Connor admitted that the size and

resources dictate the management practices of a transit police department. Mr. O'Connor testified that in 1995, the New York City Transit Police Department had 50 vehicles at its disposal. It was proven at time of the original trial that SEPTA had only four vehicles at its disposal, and many times the vehicles were not working. Moreover, Mr. O'Connor testified that the number of vehicles in New York City increased to "a few thousand" after the New York City Transit Police Department was integrated into the New York City Police Department.

119. Further, Mr. O'Connor admitted that there were significant differences in resources between the New York City Transit Police Department and SEPTA's Transit Police Department, thus, rendering his comparison of New York City to Philadelphia useless to this Court. Noteworthy in this regard is that unlike the New York City Transit Police Department, SEPTA's Transit Police Department has not been integrated into the Philadelphia Police Department. This Court finds the lack of integration to be a significant difference which also weighs against crediting Mr. O'Connor's testimony as to the applicability of the New York City transit system experience to the SEPTA transit environment, problems and difficulties encountered in Philadelphia.

*18 120. Notably, Mr. O'Connor testified that he was better able to deal with crime in New York City after he received \$118 million for capital improvements. This money was spent on better equipment, such as radios and public address systems. Nonetheless, Mr. O'Connor admitted that he had seen nothing in the record which would indicate that SEPTA had received anywhere near that level of funding or, for that matter, any funding. Furthermore, Mr. O'Connor admitted that he had no way of knowing if any of the other twenty-five transit authorities that Dr. Landy surveyed (identified on Exhibit D-54) received several hundred million dollars for capital improvements, as had the New York City transit authority.

121. Mr. O'Connor testified that New York City received approximately 1,000 new officers, representing a 25% increase in their forces from approx. 3000 to approx. 4000.

122. Mr. O'Connor agreed with defense counsel that the emergency assist call is the most serious call an officer can receive. He went on to testify that this critical task should be accomplished on foot as quickly as possible. Further, Mr. O'Connor believes that if an officer has to respond to an assist call by running from one station to another, it is important to get there on time with some level of reserve strength.

123. During the original trial of this matter this Court found that SEPTA officers are required to provide emergency assistance to other officers approximately 400

times per year. This is a critical task that Mr. O'Connor acknowledged must be handled in the most expeditious fashion possible.

124. Mr. O'Connor also agreed that a backup call, which SEPTA does nearly 2,000 times per year, can turn into an emergency assist call at any moment.

124. Mr. O'Connor was questioned whether New York transit police officers chased fleeing felons. Mr. O'Connor opined that for serious crimes, this was a priority task. Mr. O'Connor stated that it was his recollection that New York City transit authority officers engaged in foot pursuits approximately 10% of the time, notwithstanding that they generally did not pursue criminals. Mr. O'Connor admitted that it was standing New York City policy not to pursue every criminal, although Mr. O'Connor did concede that New York City transit police officers were expected to have the ability to catch fleeing felons.

125. When Mr. O'Connor was questioned on the propriety of using arrest rates as a measure of police performance, he conceded high crime areas in New York City did evaluate its officers on this basis. Curiously, however, Mr. O'Connor testified that he did not believe that arrests deterred crime beyond the incarceration of the criminal. When SEPTA's counsel and this Court pressed Mr. O'Connor, he then testified that he had no opinion as to whether or not arresting a perpetrator has a deterrent effect on other criminals in the locale. This incredible response lead this Court to question Mr. O'Connor's qualifications as an expert witness.

*19 126. The Court subsequently questioned Dr. Henderson why he believed that a person like Mr. O'Connor, with his transit police background, would have no knowledge of the deterrent effect that apprehensions have on the at-large criminal community. Dr. Henderson testified that Subject Matter Experts were not reliable for determining the efficacy of certain activities such as the effect arrests have on deterrence. To answer this question, the Court would have to look to other areas of expertise, particularly studies from criminologists that have analyzed apprehension rates and the ensuing deterrent or suppression effect on crime, such as the D'Alessio Study and others that Dr. Henderson cited. This Court concludes that Mr. O'Connor's testimony that improving arrest rates does not suppress or deter crime is unsupported and not credible.

127. Plaintiffs called as an expert witness Dr. Frank J. Landy, an industrial and organizational psychologist. The thrust of Dr. Landy's testimony was to criticize SEPTA's studies on remand. As will be discussed, *infra.*, this Court finds that Dr. Landy's criticisms are without merit and the Court does not credit his testimony.

128. Dr. Landy explained that he typically sets a cut-point for a physical abilities test one standard deviation below the average performance of incumbents. Using that method, one standard deviation below the average typically put the cut-point between the 16th and 20th percentile of incumbent performance. Put another way, 84% of the incumbents are able to exceed the performance level that Dr. Landy sets as an applicant test cut-point. For an applicant who is taking the test, he or she only need perform at the bottom 16th–20th percentile of an incumbent group.¹¹

¹¹ Among the numerous problems with this methodology is the obvious and fatal defect, at least regarding SEPTA's stated and valid goals, of relying on already underperforming incumbents to set bench marks.

129. Using this methodology, Dr. Landy typically passed 90% or more of the male applicants on his test. As an example of this methodology, counsel for SEPTA pointed out that in Austin, Texas, where Dr. Landy devised a police test, the male pass rate was 99.1% on the obstacle course. Similarly, the male pass rate on the dummy drag was 100% and the male pass rate on the bench press was 94.2%. It was also shown in the Austin Police test that the female pass rate on the obstacle course was 79.3% on the dummy drag was 92.9% and on the bench press the female pass rate was 89.3%. In light of these high pass rates, it is indisputable that Dr. Landy's test has little or no utility in the selection of either males or females.

130. Confronted with this cut-point method and the results of Austin and other jurisdictions where Dr. Landy worked, he acknowledged that although he insulated those cities from an adverse impact challenge, his test had low, little or no utility in the selection of males. Dr. Landy's methodology, while assuredly one that avoids an adverse impact challenge, has little use in the selection of candidates for the performance of safety-related tasks. This is certainly true in light of Dr. Landy's admission that the very high pass rates means that the test has no or very little utility.

***20** 131. Dr. Landy was pressed on providing some basis for this most peculiar methodology of setting cut-points. Under cross-examination, Dr. Landy was unable to identify a single other researcher that used such a method. Specifically, counsel elicited from Dr. Landy that the references that he described in support of his cut-point methodology did not, in fact, provide any authoritative support for setting a cut-point at the lowest 16th or 20th percent of an existing population.

132. Dr. Landy admitted he never went back to any jurisdiction to validate the cut-point that he had set based upon the 16th to 20th percentile of incumbent

performance. Dr. Landy acknowledged that his model for setting cut-points would ultimately lower the average incumbent performance time.

133. Dr. Landy could not point to any evidence, empirical or otherwise, to support the development of a cut-point which of necessity must over time reduce the overall work performance of the incumbent group. In this regard, Dr. Landy was cross-examined on how his methodology worked. For example, with the Cincinnati Fire Department, it was shown that incumbent firefighters who were experienced in handling hoses, ladders, sledgehammers and the tools of their profession, were outperformed by inexperienced candidates. The mean score for incumbents on a stepmill test was 316.33 seconds. However the time it took the candidates to complete the stepmill test was 176 seconds. Dr. Landy agreed that the incumbents did much worse than the applicants on this test.

134. But rather than accept the incumbent's time as the cut off, Dr. Landy admitted that he disregarded it and, *based on his expert judgment*, reset the cut-off score for the stepmill event to 180 seconds. This is precisely what SEPTA's experts had done as well with regard to the SME's in the instant case. As was demonstrated at the first trial, Plaintiff's first expert, Dr. Zedeck, and now Dr. Landy exercise judgment when setting a cut-point.

135. Curiously, Dr. Landy conceded that he had identified fitness standards for the Columbus Police Department that set a *desirable* aerobic capacity for officers at 44.8/ml. (males, age 20–29); and *acceptable* at 40.8 (males, aged 20–29). Furthermore, Dr. Landy made these same recommendations to the Chicago Police Department. Once again, Dr. Landy identified for the City of Chicago a *desirable* aerobic fitness capacity for male officers to be 44.8; and again, an *acceptable* aerobic capacity to be 40.8.

136. These fitness recommendations are at sharp odds with Dr. Landy's present testimony since, in this case, Dr. Landy deemed an appropriate standard of aerobic capacity for SEPTA officers to be somewhere between 30 and 33.5 mL/kg/min. Dr. Landy relied upon what is known as the Sothman study to determine that SEPTA's officers need no more than 33.5 ml. of aerobic capacity to perform their jobs. This logic was flawed and expressly contradicted by portions of the Sothman study. While this study was thoroughly litigated at the first trial, it is worthy to mention that the Sothman study specifically stated the purpose of choosing 33.5 was to "displace" as few firefighters as possible. When Dr. Landy was confronted with this point he admitted that 33.5 ml. would not allow a firefighter to extricate himself or save someone from a rigorous blaze.

***21** 137. Notwithstanding that Dr. Landy claims that the

“Sothman” article was the basis upon which he determined that a 33.5 ml. aerobic capacity was adequate for SEPTA, the article itself indicates that 33.5 ml. does not permit a firefighter to actively engage in an arduous fire, a high rise fire, or even extricate himself or someone else from danger. In fact, the “Sothman” article went on to state that 41 ml. was a more desirable VO_2 max standard and attributed that recommendation to SEPTA’s expert, Dr. Davis.

138. Dr. Landy admitted that a firefighter’s ability to extricate himself or someone else from a fire was a critical task. In a similar fashion, Dr. Landy admitted that a critical task of a SEPTA transit officer was to carry out an emergency assist as fast as they could and arrive at the scene with the ability to handle arduous police work.

139. Dr. Landy did not dispute that SEPTA had the right to test for the ability of an individual to complete the task of carrying out an emergency assist. To Dr. Landy, this was no less of an imperative than a firefighter being able to extricate himself or a victim from a fire.

140. Clearly, Dr. Landy’s use of 33.5 mL/kg/min is inapplicable to the critical tasks of a firefighter; nor does his recommendation of 33.5 mL/kg/min address the critical demands that is required of a SEPTA transit officer. Dr. Landy did, however, admit that making an emergency assist may save an officer from injury or prevent death, both of which are critical to the transit officer position.

141. Dr. Landy criticized Drs. Davis and Henderson’s work on the basis that there may have been classification errors in their aerobic capacity groupings. However, when confronted with his own work in Sothman, (Table 6), Dr. Landy admitted that it was possible that his classifications could likewise be in error with respect to any individual simply because measurement error pertains to any study. When pressed, Dr. Landy conceded that in his work on the Sothman study, not even a laboratory test is 100% accurate as to an individual’s aerobic capacity.

142. Measurement error does not establish a correlation; rather, it serves to suppress or understate it. Thus, when Drs. Davis and Henderson established a correlation between running the 1.5 mile run and success on the pursuit course, any measurement error would only serve to mask the already high correlation. In affirmation of this widely accepted principle, Dr. Landy admitted that peer review publications accept for publications studies where measurement error occurs. It appears that Dr. Landy dismisses measurement error in his studies, yet criticizes Drs. Davis and Henderson for the potential of measurement error.

143. Dr. Landy also quibbled with one of the tables that contained four aerobic capacity classifications that Dr.

Henderson established. More to the point, however, Dr. Landy did not claim that classifying individuals as either *failing SEPTA’s test* or *passing SEPTA’s test* is incorrect or is an arbitrary classification. Consistent with Dr. Landy’s approach of setting pass points at very low levels he admitted that in the Sothman study, he set the passing point on a firefighter test at two standard deviations below the mean of incumbent performance. Consequently, 97% of the incumbent group could pass the cut-point. Dr. Landy’s model on the “Sothman” article was to set the cut-point at the lowest 3rd percentile.

*22 144. Dr. Landy’s method is to consistently lower standards without any rational demonstration or support for setting cut-points at the low level that he now advocates. For example, notwithstanding that the Sothman article does not support setting a firefighters aerobic capacity at 33.5 ml., if one expects a firefighter to be able to meet demanding and dangerous tasks, Dr. Landy went on to assert that a VO_2 max value of 33.5 is considerably more compelling as a minimum threshold than 42.5 for police officers. This statement was made in Dr. Landy’s rebuttal report.

145. At the first trial of this matter, this Court rejected Plaintiffs’ first expert’s recommendation of 36 ml. and, now, Dr. Landy is offering an even lower aerobic capacity cut off, despite the fact that in other jurisdictions his recommendations for police officers were entirely consistent with SEPTA’s 42.5 ml. Dr. Landy’s recommendations for the Chicago and Columbus Police Departments were even higher than Dr. Davis 42.5 since Dr. Landy defined a desirable aerobic capacity at 44.8 ml. This inconsistency raises serious questions about the credibility of Dr. Landy’s testimony. Indeed, one cannot understand Dr. Landy’s 33.5 ml. recommendation in light of his earlier identification of a fitness standard of 44.5 as acceptable for police officers in Chicago and Columbus.¹²

¹² That this recommendation was for male officers ages 20—29 does nothing to enhance the credibility or consistency of Dr. Landy’s testimony.

146. At the trial of this matter, Defendant’s Exhibit 58 was submitted into evidence which compared the performance of those meeting Landy’s 33.5 cut-off to those individuals passing SEPTA’s test. Exhibit D–58 reveals that the Landy Group performed markedly worse on all job standards than those passing SEPTA’s aerobic capacity test. The differences were statistically significant and across all job standards.

147. Confronted with significant performance deficiencies of those that passed the Landy proposed cut-point of 33.5 relative to those that passed SEPTA’s aerobic capacity test, Dr. Landy conceded that the Landy group could not

outperform those passing SEPTA's test on the relevant job criterion.

148. Dr. Landy did not challenge Drs. Davis and Henderson's choice of a .35 mile pursuit course, particularly in light of Dr. Landy's choice of a 540 yard, or .31 mile obstacle course for officer candidates in Colorado Springs. Dr. Landy readily admitted that SEPTA officers regularly ran 3 to 5 and even 8 blocks. Dr. Landy admitted that SEPTA officers ran underground and encountered various obstacles, such as turnstiles, people and barriers when they have to pursue and chase fleeing felons.

149. At the initial trial, this Court agreed with Dr. Davis' rejection of the SME's estimate of the time to run 1 mile in full gear, and agreed with Dr. Davis' use of his judgment to reject the SME's estimate. Similarly, Dr. Zedeck agreed that a researcher must exercise judgment and once again, Plaintiffs' expert, Dr. Landy can be found to be exercising a professional judgment (albeit questionable) when setting a cut-point.

*23 150. Plaintiffs' claims that SEPTA's experts cannot exercise judgment is wide of the mark and misplaced. This is especially true where SEPTA's experts have relied on extensive empirical data developed in the first trial and developed here on remand. It is noteworthy that when counsel for SEPTA pressed Dr. Landy, he admitted that although he looked at data from other cities before he exercised his judgment to reset the stepmill test in Cincinnati, he did not do a transportability study that is required by the Uniform Guidelines.

151. Dr. Landy was also presented with the test results for the Madison Fire Department where he established cut-points for physical abilities tests. There, it was found that in seven physical ability events the applicants outperformed experienced firefighters on Dr. Landy's tests. Confronted with these findings and faced with this anomaly, Dr. Landy nonetheless agreed that the individuals who were the incumbents had more experience on the seven physical abilities tests than the applicants. Dr. Landy admitted that on each of the 7 physical ability tests the incumbents who had much more experience did worse than the applicants.

152. The degree of encouragement provided by either SEPTA or the expert's in this case to runners is irrelevant, and has no bearing on the results. Dr. Landy candidly admitted that he could not say how this allegedly affected the outcome of the study.

153. Based on the foregoing, the Court finds that the testimony of Dr. Landy is not credible or reliable, and thus assigns it no weight.

III. CONCLUSIONS OF LAW

1. At the original trial of this action, this Court found that it was more likely than not that applicants that pass the 1.5 mile run component of SEPTA's physical fitness test will be successful performers on the job, whereas it is highly probable that those officers who do not pass the 1.5 mile run component of SEPTA's test will not be successful performers on the job because they lack the aerobic capacity necessary to fulfill the demanding obligations of a SEPTA transit officer. Additionally, this Court found that SEPTA's aerobic requirement was readily justifiable as a business necessity.

2. As stated previously, the Third Circuit found that this Court did not employ the proper business necessity standard, and thus remanded the case back to this Court to determine whether or not SEPTA met its burden of proving that 42.5 ml/kg/min is the minimum aerobic capacity necessary to successfully perform the job of SEPTA transit officer.

3. Although this Court has unequivocally stated that it will not disturb its prior factual findings, it did state that it would reconsider those findings in light of the new standard articulated by the Third Circuit. Moreover, the Court permitted the parties to further develop the record tailored to the newly articulated standard.

4. Thus, the sole question to be addressed on remand, and the sole question to be resolved in these conclusions of law, 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. As will be explained more fully herein, the Court finds that SEPTA's evidence adduced at the first trial, both separately and particularly in combination with the evidence adduced at the remand hearing, clearly demonstrates that its aerobic capacity requirement is the minimum for successful performance of the job of SEPTA transit police officer.

*24 5. As the testimony of virtually all witnesses at both trials in this case establishes, apprehending perpetrators, deterring crime, assisting fellow officers in emergency situations, and backing up fellow officers are critical components of the job of SEPTA transit police officer.

6. An inability to proficiently perform any of these tasks would compromise the effectiveness of the SEPTA transit police, as well as compromise the safety of the officer, his or her fellow officers, and the public.

7. Thus, a SEPTA officer would not be satisfactorily performing his or her duties if he were unable to perform any of these critical tasks to an appropriate level of proficiency.

8. As demonstrated at both the original trial and the

remand hearing, individuals below an aerobic capacity of 42.5 ml/kg/min are unable to satisfactorily perform the critical tasks necessary for a successful SEPTA officer. This is borne out by the facts found in this case, including the following.

- a. SEPTA officers are a part of a unique, foot based patrol unlike any other transit force.
- b. Each SEPTA officer must engage in at least one aerobic encounter during the course of his or her duties every month, either as an emergency assist, or a running backup. As already established, these are critical components of the duties of an officer.
- c. Dr. Davis has established that running is a critical and essential task, and that there exists a significant correlation between police officer performance and a 1.5 mile run.
- d. Crime has been dramatically reduced since implementation of the fitness program. Lt. Maslin believes that the fitness program has contributed to this reduction.
- e. Dr. Siskin, relying on data provided by Dr. Moffatt, determined that only 27% of the perpetrators arrested had an aerobic capacity below 42 ml/kg/min.
- f. Dr. Moffatt's studies demonstrate that a SEPTA transit police officer with an aerobic capacity of less than 45 ml/kg/min may not arrive in a timely fashion to an assist or backup, and their ability to do work drops off so dramatically, they will likely be ineffective upon arrival.
- g. Dr. Henderson's reanalysis of this data demonstrated that those individuals below SEPTA's aerobic capacity cut point had only a 33% chance of arriving at an emergency assist in a timely manner, vs. 80% to 90% for those who meet the standard.
- h. On five job standards, 80% of those that met SEPTA's aerobic capacity test could meet the minimum; only 33% of those who failed could meet the minimum.
- i. As established by the testimony of Dr. Henderson, there is a significant relationship between arrest rates and the deterrence of crime. Arrest rates of at least 20% serve to suppress and deter crime.
- *25 j. Drs. Davis and Henderson conducted additional research for the remand trial, focusing on pursuit and apprehension of perpetrators.
- k. The standards set by Drs. Davis and Henderson focused on the ability of officers to apprehend and catch the slowest 20–25% of the perpetrators as being

minimally acceptable, as this level provides meaningful deterrence.

- l. The performance of those not meeting SEPTA's standard of 42.5 ml/kg/min was again abysmal. Only 14% of the failing group could meet the emergency assist standard (vs. 84% of the passing group); On the pursuit and dummy drug test, those passing SEPTA's standard could successfully perform over 50% of the time, those failing were able to successfully perform only 4% of the time.
 - m. On all job standards, the failing group did not achieve even a minimally acceptable performance level, with success rates ranging from only 5% to 20%.
9. The evidence presented at the first trial and on remand clearly demonstrates that 42.5 is the minimum aerobic capacity necessary to successfully perform the job, given the abysmal success rate on critical job tasks of those that failed SEPTA's 1.5 mile running test when compared to those that passed the 1.5 mile running test.
 10. This Court credits the testimony of Defendant's experts, Dr. Paul O. Davis and Dr. Norman D. Henderson. The testimony and studies of Drs. Davis and Henderson conclusively demonstrates that SEPTA's aerobic capacity requirement of 42.5 mL/kg/min is the minimum required to perform the critical tasks of a SEPTA transit police officer.
 11. Moreover, this Court has already credited the studies of Drs. Moffatt and Siskin with respect to the work each performed in preparation for the original trial.
 12. Based on the evidence presented at both trials, the input of SEPTA management, the contributions of the Subject Matter Experts, and the comprehensive studies of Drs. Siskin, Moffatt, Henderson and Davis, this Court concludes that meeting SEPTA's aerobic capacity standard is clearly the minimum required to perform the critical tasks of the job such as pursuits, officer back-ups, officer assists and arrests. Any lesser requirement simply would not satisfy the minimum qualifications for the job of SEPTA transit police officer and would endanger the public and undermine deterrence of crime and apprehension of criminals.
 13. Despite rejecting the holding of *Spurlock v. United Air Lines, Inc.*, 475 F.2d 216 (10th Cir.1972) (as not having been specifically endorsed by Congress in the 1991 Act) the Third Circuit does suggest in Footnote 16 that public safety is a legitimate consideration, and is encompassed in the business necessity standard articulated by the Court. Therefore, public safety is a factor this Court must consider when weighing all of the evidence in this case, although public safety alone likely does not justify otherwise discriminatory employment

practices.

14. As established at the first trial, SEPTA could have had an additional 70 part I arrests, and 470 additional overall arrests, had all of its officers met the aerobic standard. As this Court stated in its original findings, this Court is not unmindful of the significance of the additional 470 overall arrests and additional 70 Part I arrests that would be obtained if SEPTA's less-fit officers met SEPTA's aerobic capacity standard. For many of the 470 additional arrests, there would be fewer criminals in the SEPTA transit system left to prey on and victimize the riding public. 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.

***26** 15. These lost arrests have a significant impact on the public safety.

16. In footnote 24, the Court of Appeals instructed this Court that plaintiff's evidence of incumbent officers that had failed the physical fitness tests yet successfully performed the job, and that other police forces function well without an aerobic capacity requirement was relevant evidence and should be considered. The Court has done so, and finds the evidence itself wholly lacking on its own, and in comparison with the overwhelming evidence in support of a 42.5 ml/kg/min cutoff, utterly unpersuasive.

17. During the course of the trial, and again on remand, Plaintiffs presented evidence regarding physical fitness tests from other transit authorities and police jurisdictions, and argued that these tests, which have lower standards than SEPTA's test, should be adopted by SEPTA. This Court originally found that invalidated tests from dissimilar law enforcement agencies were not an acceptable alternative to SEPTA's validated test.

18. On remand, Plaintiffs have again failed to present any evidence indicating that the physical fitness tests of unrelated law enforcement agencies are appropriate for SEPTA or will equally serve SEPTA's needs. As this Court has already found, SEPTA is a unique foot based patrol, with demands placed on its officers unlike those placed on other officers in other police departments.

19. Moreover, as was demonstrated at remand, many of the fitness requirements that have been adopted by other forces have been developed with the stated goal of displacing as few incumbents as possible, and to avoid Title VII challenges. These were not the goals of SEPTA, as it was instead concerned with improving the quality of its police force.

20. Based on these factors, the Court assigns very little weight to this evidence. Moreover, regarding incumbent performance, the Court assigns little weight to that evidence as well, for many of the same reasons the Court of Appeals criticized this Court's reliance on commendations in the first trial-the subjective nature of such evidence. There are simply too many factors that determine incumbent performance ratings to make such evidence reliable. Not only are there subjective supervisory evaluations, as this Court noted in the first trial, there are significant issues with the officers' collective bargaining unit regarding evaluations and performance requirements. Additionally, basing requirements on an incumbent force that was by all accounts insufficient is hardly the best method by which to measure performance and improve standards. This Court will not accept the proposition that employers are restricted from raising standards and that they are bound in their hiring by the level of performance of its incumbent work force.

21. For these reasons, the Court determines that this evidence is entitled to little weight.

***27** 22. Even so, had the Court fully credited this evidence, it would fall far short of changing the outcome of this opinion in the face of the overwhelming empirical evidence credited and previously detailed by this Court in support of 42.5 ml/kg/min being the minimum aerobic capacity necessary for successful performance of job of SEPTA transit police officer.

23. At the remand hearing, Plaintiffs also argued that Dr. Davis improperly ignored the work experience of the subject matter experts ("SMEs") when setting the 42.5 mL/kg/min aerobic capacity standard for SEPTA officers. This argument was fully considered and subsequently rejected by this Court at the original trial of this matter, and again in this Court's findings of fact based on the supplemental hearing. Dr. Davis did not ignore the work experience of the SME's. In fact, he incorporated virtually all of their collective wisdom into designing his tests, with one exception, their estimate of the time to run one mile. The pace suggested by the SME's translates to an aerobic capacity of 33.5, which is less than the aerobic capacity maintained by a sedentary female. A level of 33.5 is hardly appropriate for a safety-sensitive job that entails critical, life-and-death tasks dependent on aerobic capacity ranging from 42.5 to 54. Moreover, these admittedly older and questionably fit officers have every incentive to suggest a fitness standard that they could meet with little difficulty.

24. Therefore Dr. Davis, like all experts in this field including plaintiffs' own Dr. Landy, exercised his judgment, and decided that a VO₂ max of 42.5 ml/kg/min would be more appropriate. As at the first trial, the Court

agrees with this assessment.

25. As stated by this Court in the first trial, Dr. Davis' decision to require 42.5 mL/kg/min of aerobic capacity was supported both empirically and by his considerable experience in developing tests for law enforcement agencies. As the SIOP Principles acknowledge:

[J]udgment is necessary in setting any critical or cutoff score. A fully defensible empirical basis for setting a critical score is seldom, if ever, available. The only justification that can be demanded is that critical scores be determined on the basis of a rationale which may include such factors as estimated cost-benefit ratio, number of openings and selection ratio, success ratio, social policies of the organization, or judgments as to require knowledge, skill or ability on the job. If critical scores are used as a basis for rejecting applicants, their rational or justification should be made known to the users.

SIOP Principles at 32–22 (emphasis added). Dr. Davis' validation study satisfies this standard in that it articulates a justification for using a cutoff score of 42.5 mL/kg/min on SEPTA's physical fitness test. Additionally, as was seen at trial, all experts in this field use their judgment at some point in setting benchmarks.

26. Additionally, while expert judgment was certainly a component of determining these cut-off scores, as was already clearly demonstrated, the 42.5 ml/kg/min was also directly related to the input from the SME's, as well as SEPTA management's stated goals of improving the overall fitness of its force, increasing SEPTA's ability to control crime on SEPTA property and enhance ridership safety.

*28 27. Finally, although it was Dr. Davis who initially suggested the 42.5 ml/kg/min standard as a compromise to avoid a draconian effect on women, it was hardly an arbitrary figure, as Dr. Davis had recommended a similar number to the Anne Arundel police force, and had a familiarity with that standard and the capabilities and limitations of such a standard as a result of work he had done for firefighters in Minnesota. It was hardly a random or arbitrary figure, as the evidence at both the first and subsequent trials has abundantly borne out.

28. Plaintiffs also argue that the cutoff scores on each of the job standards used by Drs. Davis and Henderson were arbitrary, and based solely on expert judgment. Plaintiffs base this argument on footnote 19 of the Third Circuit's opinion. Initially, the Court notes that, as was seen throughout, the standards set were so low, that no party can credibly complain about a standard that only requires a participant (prospective SEPTA transit officer) to be successful approximately 25% of the time. Moreover, the evidence adduced at both the first and second trials was not simply the judgment of one expert. Defendants' evidence was developed, and their benchmarks were set,

as a result of the mandate of SEPTA management to improve the crime fighting ability of SEPTA's force, and the fitness of its officers. Subject Matter Experts were consulted, and their input was incorporated into virtually every aspect of designing SEPTA's fitness requirements, and the tests designed to test for those abilities in officers, with an appropriate amount of expert judgment brought to bear on these issues. As Plaintiffs' expert in test development, Dr. Sheldon Zedeck testified at the first trial, the test developer's judgment can and should be exercised by the person validating the test.

29. Additionally, these various standards were not merely the product of one expert, but based on the data collected, and expertise of Drs. Henderson, Davis, Moffatt, and Siskin.

30. Thus, it is this Court's finding on remand that Drs. Davis and Henderson developed strong empirical evidence to support their conclusion that an aerobic capacity of 42.5 was the minimum aerobic capacity required to perform essential transit officer tasks.

31. Furthermore, at the initial trial, SEPTA offered 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 capacities below SEPTA's cutpoint; and SEPTA's most recent studies more than provides an appropriate empirical basis for demonstrating that its cutpoint is already set at the minimum.

32. As the evidence in this case makes abundantly clear, any standard less than 42.5 ml/kg/min would result in officers unable to arrive in a timely fashion to help a fellow officer in an assist or back-up, and officers unable to apprehend perpetrators. Thus, SEPTA would be a police force with officers who were a danger to themselves, other officers, and the public at large, who were unable to effectively fight and deter crime. This is exactly the situation that existed in the late 1980's and early 1990's, and it would be unconscionable for this or any court to lower standards that would inevitably result in a degradation of law enforcement back to such a dangerous time.

*29 33. In short, the result of a standard below 42.5 ml/kg/min would be officers unable to successfully perform the job of SEPTA transit police officer.

34. Based on the foregoing Findings of Fact and Conclusions of Law, this Court finds that SEPTA has met its burden of establishing the business necessity of its aerobic capacity standard as articulated by the United States Court of Appeals for the Third Circuit. Thus, this Court will enter judgment in favor of defendant, and

Lanning v. Southeastern Transp. Authority, Not Reported in F.Supp.2d (2000)

against plaintiffs.

AN APPROPRIATE ORDER FOLLOWS.

Parallel Citations

84 Fair Empl.Prac.Cas. (BNA) 1012