

Order approving Exam 2000 for use - Doc. 986.pdf

UNITED STATES DISTRICT COURT
EASTERN DISTRICT OF NEW YORK

-----X

UNITED STATES OF AMERICA,

Plaintiff,

-and-

THE VULCAN SOCIETY, INC., *for itself and on behalf of its members*, JAMEL NICHOLSON, and RUSEBELL WILSON, *individually and on behalf of a subclass of all other victims similarly situated seeking classwide injunctive relief*;

ROGER GREGG, MARCUS HAYWOOD, and KEVIN WALKER, *individually and on behalf of a subclass of all other non-hire victims similarly situated*; and

CANDIDO NUÑEZ *and* KEVIN SIMPKINS, *individually and on behalf of a subclass of all other delayed-hire victims similarly situated*,

Plaintiff-Intervenors,

-against-

THE CITY OF NEW YORK,

Defendant.

-----X

NICHOLAS G. GARAUFGIS, United States District Judge.

Before the court is the City’s motion for approval to use entry-level firefighter Exam 2000 to create an eligible hire list and begin hiring entry-level firefighters from that list. (City Mem. in Support of Mot. (Docket Entry # 977) at 1.)

This motion arrives at this court after substantial work by many parties. In 1973, Judge Edward Weinfeld of the United States District Court for the Southern District of New York found that an entry-level firefighter exam produced an adverse impact and was not job related,

and ordered the City to design a non-discriminatory hiring exam for the FDNY “in accordance with professionally accepted methods of test preparation.” Berkman v. City of New York, 536 F. Supp. 177, 183 (E.D.N.Y. 1982); see also Vulcan Soc. v. Civil Serv. Comm’n, 360 F. Supp. 1265 (S.D.N.Y. 1973). Five mayors, including the incumbent, have presided over the City since then, and the problem remained. In 2007, the United States Department of Justice brought suit against the City because its entry-level firefighter exams still showed the same patterns of disparate impact and unprofessional design as the exam Judge Weinfeld considered. The United States has pursued this case patiently for more than five years. The United States was joined by Plaintiff-Intervenors, principally the Vulcan Society, which have been equally persistent in its litigation of this case. **It is because of the case brought and pursued by these Plaintiffs that the court ordered the City to create a fully professionally-prepared exam that would comply with the dictates of federal, state, and local anti-discrimination laws.** That order was effectuated by the tireless efforts of Special Master Mary Jo White, who oversaw each step of the process, and the work of multiple disinterested testing design specialists who worked for the City, the Plaintiffs, and the Special Master, but whose effort produced a single exam.

The City argues that Exam 2000 does not produce a “practical adverse impact” on minority candidates and, in the alternative, that the use of the exam is job-related or consistent with business necessity, and for those reasons, its use should be approved. (City Mem. in Support of Mot. at 1.) Upon due consideration, the City’s specified pass-fail and rank-order usages of Exam 2000 are approved.

I. BACKGROUND

In 2009, this court ruled that there was no genuine issue of material fact as to the City’s liability for disparate impact violations of Title VII of the Civil Rights Act of 1964, as amended,

42 U.S.C. 2000e et seq. (July 22, 2009 Mem. & Order (Docket Entry # 294).) Part of the standard relief in Title VII cases is compliance relief, i.e., efforts “designed to erase the discriminatory effect of the challenged practice and to assure compliance with Title VII in the future.” Berkman v. City of New York, 705 F.2d 584, 594 (2d Cir. 1983). Compliance relief may include “specifying procedures and standards for a new valid selection procedure.” Guardians Ass’n v. Civil Serv. Comm’n of N.Y., 630 F.2d 79, 109 (2d Cir. 1980). In a subsequent decision, the court ordered the City to begin developing a new selection device that would not violate Title VII (Jan. 13, 2010 Mem. & Order (Docket Entry # 390) at 54-55; see also Feb. 24, 2010 Minute Entry); the court appointed Mary Jo White, a partner at the law firm of Debevoise & Plimpton LLP, to serve as Special Master to oversee this project (Order Appointing Special Master (Docket Entry # 448) at 2). In response, the City hired an independent test development expert company, which began work on this process in frequent consultation with the Plaintiffs and their testing experts, and overseen by Special Master White (who was advised by her own testing expert). (City Mem. in Support of Mot. at 2.) Special Master White reported to the court regularly on the milestones of creating and using Exam 2000: job analysis, test design, validation studies, and test administration and scoring. (See Special Master Reports 1-11 (Docket Entry ## 459, 503, 571, 606, 634, 691, 739, 785, 841, 875, & 935).) Each of these steps is described in detail in the technical report of the City’s testing expert, PSI Services LLC. (FDNY Firefighter Test Development and Validation Report (City Mot. Ex. 1 (Docket Entry # 976-1)).)

The City now requests court approval to use the results of Exam 2000.¹ Exam 2000 has been graded with a pass-fail cutoff score, scaled to a 70 on a 100-point scale; if a candidate

¹ Exam 2000 is the same test as promotional Exam 2500, which was administered to emergency medical technicians who already work at the FDNY and who are seeking to join the ranks of firefighters.

scores below that score, he or she fails and is not considered further in the hiring process. The scores of the candidates who have received a score greater than the cutoff score were then combined with any applicable bonus points, and these candidates were placed on a rank-ordered (from highest score to lowest passing score) list of candidates for consideration as entry-level firefighters.² (See FDNY Firefighter Test Development and Validation Report at 74-78.) Special Master Mary Jo White recommends that the court approve the use of this exam in these ways. (Report and Recommendation (Docket Entry # 984).) The City does not request a fairness hearing for the court to evaluate third-party collateral challenges to the exam. (See Sept. 20, 2012 Status Conf. Tr.)

II. DISCUSSION

In evaluating a disparate impact claim, the court undertakes a multi-step process. The first step is evaluating whether there is an adverse impact on minority candidates, and if it does have an adverse impact, the second step is evaluating whether the practice is defensible as job-related. Gulino v. N.Y. State Educ. Dep't, 460 F.3d 361, 383-84 (2d Cir. 2006). Disparate impact case law requires analysis of the adverse impact of precise employment practices. Lewis v. City of Chicago, 130 S. Ct. 2191, 2197 (2010). Here, the court must consider the City's use of a pass-fail score and the City's rank-ordering of candidates based on their score and bonus points.

All the parties agree that the use of Exam 2000 score results in rank-ordering of candidates as the City requests would produce little to no difference in hiring between minority and white candidates. (City Mem. in Support of Mot. at 20-21; Pl.-Intervenors Ltr. (Docket

² Bonus points in this context refer to additional points given for New York residents, certain military veterans, and certain relatives of police officers and firefighters who died in the line of duty. These bonus points have been included due to long-standing New York State laws and City policy. See, e.g., N.Y. Civ. Serv. Law §§ 85-85-c (McKinney 2006). Their use has not been challenged in this lawsuit.

Entry # 980) at 1; United States Ltr. (Docket Entry # 982) at 1-2.) The City's expert report shows that over the expected four-year life of the exam, minority candidates are projected to be hired at close to the rate at which white candidates will be hired.³ (FDNY Firefighter Test Development and Validation Report at 79-80.) Minority candidates from all but one minority group will be certified for further processing at a rate that is at least eighty percent of the rate of success of white candidates for each year of the expected life of the exam-generated list. (Id.) See also EEOC v. Joint Apprenticeship Comm. of Joint Indus. Bd. of Elec. Indus., 186 F.3d 110, 118 (2d Cir. 1998) (four-fifths or 80% rule is "looked to [] for guidance in gauging the significance of statistical evidence of discrimination"). Candidates from one minority group, Native Americans, have a projected success rate of less than eighty percent of the white candidates for two of the four years that the Exam 2000 list is anticipated to be active. (FDNY Firefighter Test Development and Validation Report at 80.) However, standard deviation analysis suggests no adverse impact, as there is less than two units of standard deviation for the disparity for these years. (Id.) See 29 C.F.R. § 1607.4(D) ("[D]ifferences [greater than four-fifths] in selection rate may not constitute adverse impact where the differences are based on small numbers and are not statistically significant . . ."); See generally Waisome v. Port Auth. of New York & New Jersey, 948 F.2d 1370, 1376 (2d Cir. 1991) ("Standard deviation analysis measures the probability that a result is a random deviation from the predicted result—the more standard deviations the lower the probability the result is a random one."); Smith v. Xerox Corp., 96 F.3d 358, 365 (2d Cir. 1999), overruled on other grounds by Meacham v. Knolls Atomic Power Lab., 461 F.3d 134, 141 (2d Cir. 2006) (the Second Circuit has "looked to whether the plaintiff can show a statistically significant disparity of two standard deviations" in making a

³ See N.Y. Civ. Serv. Law § 56 (1) ("The duration of an eligible list shall be fixed at not less than one year nor more than four years . . .")

prima facie showing of disparate impact.). Standard deviation analysis also confirms that there is no adverse impact on any other minority (ethnicity or gender) from the rank-order usage of the exam. (See FDNY Firefighter Test Development and Validation Report at 80.) Based on the agreement of the parties, the recommendation of the Special Master, and the un-rebutted expert evidence presented, the court approves the rank-order use of the results of Exam 2000 as the City requests.

The court next turns to the pass-fail usage of the Exam. The statistical analysis of the pass-fail rates of candidates who took Exam 2000 is mixed, and the results are different depending on the method of analysis used. Although the pass rate of all minority candidates is close to the pass rate of white candidates at the pass-fail score of 70 (with minority passage rates between ninety-seven and ninety-nine percent of those of white candidates), the difference is statistically significant (i.e., greater than two units of standard deviation) for the candidates of all minority groups other than women. (FDNY Firefighter Test Development and Validation Report at 80.) This court has previously noted that reliance on the four-fifths rule by itself is inappropriate, and that evidence of statistical significance is also important to determining whether there is an adverse impact. (July 22, 2009 Mem. & Order at 33-34.) Thus, the statistically significant discrepancy in the pass-fail rates of minority candidates is of concern to the court.

Nonetheless, the City and Plaintiff-Intervenors believe that the court should approve the Exam because it has no practical adverse impact, in that the list resulting from Exam 2000 will be exhausted well before anyone who earned a scaled score below a 70 would be considered even if there were no cutoff score. (City Mem. in Support of Mot. at 20-22; Pl.-Intervenors Ltr. at 1.) The City and Plaintiff-Intervenors argue that, given the four-year life of the exam and the

City's estimated hiring needs over those four years, it is extremely unlikely that anyone who is ranked below a 97 or 96 on the scaled exam will be called for further consideration even if there were no pass-fail score, and, as discussed above, there is no statistically significant difference between the rates of white and minority candidates being ranked at 97 or above. (See City Mem. in Support of Mot. at 20-22; FDNY Firefighter Test Development and Validation Report at 78-80.) The court has found no legal support for this argument. **The cutoff score is formally a dispositive step in the hiring process; thus, it must be evaluated for its possible adverse impact regardless of whether further steps in the hiring process would lead to cancelling its adverse impact at the "bottom line." Connecticut v. Teal, 457 U.S. 440, 450-52 (1982).**

Noting the factual issue of adverse impact at the cutoff score, the court considers the evidence of the job-relatedness of the use of that score. The City argues this as an alternative ground for concluding that the cutoff score is consistent with Title VII, offering expert evidence that the Exam was studied for criterion validity, content validity, and construct validity. (See City Mem. in Support of Mot. at 11-14.) These concepts are defined in the Equal Employment Opportunity Commission's Uniform Guidance on Employee Selection Procedures, 29 C.F.R. §§ 1607.1-1607.18, a source that the Second Circuit considers "perhaps the most important source of guidance" in evaluating the evidence of expert test creation professionals, and the "primary yardstick in which [a court] measures" a defendant's attempt to show job-relatedness. Gulino, 460 F.3d at 383-84. The Guidelines state that employers "may rely upon criterion-related validity studies, content validity studies or construct validity studies" to "validate," i.e., demonstrate, the job-relatedness of selection procedures. 29 C.F.R. § 1607.5(A) (emphasis added); see also Banos v. City of Chicago, 398 F.3d 889, 892 (7th Cir. 2005) ("The City can show that its process is 'job-related' by any one of three tests: criterion-related, content validity,

or construct validity.”). A study of criterion validity is a comparison of how examinees perform on an exam with how they perform in the job itself. See 29 C.F.R. § 1607.14(B). A criterion-related validity study “should consist of empirical data demonstrating that the selection procedure is predictive of or significantly correlated with important elements of job performance.” 29 C.F.R. § 1607.5(B). Such a study can be difficult to perform but, if done properly, is very useful in considering job-relatedness in part because it provides evidence of how well an exam is testing constructs (that is, relatively abstract qualities of an applicant that may not be directly observable but can be crucial to job performance). See Guardians Ass’n, 630 F.2d at 92.

The independent testing consultant designing Exam 2000 performed a criterion validity study of Exam 2000 and its cutoff score. It did so by administering a draft version of Exam 2000 to a large number of FDNY firefighters and comparing the score of each incumbent firefighter on Exam 2000 to that firefighter’s scores in Fire Academy examinations and on specially-created evaluations from the supervisors of those same firefighters. (FDNY Firefighter Test Development and Validation Report at 47-54.) The data showed a significant correlation between success on academy scores and success on the exam, and between success on some of the test components and candidates’ performance reviews. (Id. at 56.) Based on the results of this study, the exam was assembled. (Id.) Moreover, this study was used to determine the cutoff score. The consultant determined the empirical relationship between exam score and job performance (as ranked by incumbent firefighters’ supervisors). (Id. at 70.) This mathematical relationship was used to determine the exam score that corresponded to a minimally acceptable performance ranking. (Id.) This exam score became the raw cutoff score that was scaled to a 70 in the final score system. (Id. at 70 & App. AB-AD (Docket Entry # 976-3).)

The City argues that this criterion validation study is evidence that the exam, and specifically its cutoff score, is job-related. (City Mem. in Support of Mot. at 22-24.) The United States agrees, and represents that its experts agree as well. (United States Ltr. at 1-2.) The court is also persuaded. Although the court is fully aware that a criterion study is not the only way to show job-relatedness, see Guardians, 630 F.2d at 105, it is among the most valuable methods in showing a connection between the use of a particular employment selection device and success on the job, which is the goal of the job-relatedness defense, see Albemarle Paper Co. v. Moody, 422 U.S. 405, 431 (1975) (to be job-related, test must be “shown, by professionally acceptable methods, to be predictive of or significantly correlated with important elements of work behavior which comprise or are relevant to the job”). The court notes that Plaintiff-Intervenors have not conceded that the Exam’s pass-fail score is job-related. (See Pl.-Intervenors Ltr. at 2 n.1 (“If [] the court were to reach the question of the proposed cut scores compliance with Title VII, state and city law, Plaintiff-Intervenors submit that there exist substantially equally valid alternatives with less adverse impact. [C]onsideration of that question would be an academic exercise.”) However, they have not provided a specific explanation of why they believe it is not, let alone provided evidence to rebut the City’s technical report. (See id. at 2.) In light of the un-rebutted evidence that the pass-fail cutoff score was selected through a robust criterion validation study, one of several sufficient methods of demonstrating job-relatedness, the court accepts the City’s argument that the pass-fail usage of the exam is job-related. Accordingly, the court approves its use as the City has requested.

The court notes that this exam and resulting list are the product of a City-led, collaborative effort to improve on the City’s past record of non-compliance with Title VII. While the court appreciates each party’s contribution to this effort, it is most grateful for the pro

bono efforts of Special Master Mary Jo White and her colleagues at Debevoise & Plimpton LLP, including her law partner Mary Beth Hogan and her associates Philip A. Fortino and Chinelo E. Dike-Minor. Their work in supervising this effort has been tireless and thorough.

The court further observes that its grant of the City's motion to use Exam 2000 as it requests does not end this case. The City remains under court order to, inter alia, create and implement a non-discriminatory post-examination process (for all candidates for the job of firefighter) under the guidance of Court Monitor Mark S. Cohen. (See Remedial Order and Partial Judgment, Permanent Injunction, & Order Appointing Court Monitor (Docket Entry # 765) ¶¶ 37-46.) This process must of course comply with Title VII and New York State and City law as well.

III. CONCLUSION

The City's motion to use Exam 2000 in the manner described to create a list of eligible candidates for the position of entry-level firefighter is APPROVED. Mary Jo White is discharged from the role of Special Master with the court's thanks. Any duties that remain under the terms of the previous Order Appointing Special Master (Docket Entry # 448) are assigned to Court Monitor Mark S. Cohen.

SO ORDERED.

Dated: Brooklyn, New York
September 28, 2012

s/Nicholas G. Garaufis

NICHOLAS G. GARAUFI
United States District Judge