IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF NEW YORK

UNITED STATES OF AMERICA,

PLAINTIFF,

AND THE VULCAN SOCIETY, INC., ET AL.,

PLAINTIFFS-INTERVENORS,

V.

CITY OF NEW YORK, ET AL.,

DEFENDANTS.

CIV. ACTION NO. 07-CV-2067 (NGG)(RLM)

SERVED MARCH 9, 2009

MEMORANDUM OF PLAINTIFF UNITED STATES IN SUPPORT OF PLAINTIFFS-INTERVENORS' MOTION FOR SUMMARY JUDGMENT REGARDING JOB RELATEDNESS AND BUSINESS NECESSITY

TABLE OF CONTENTS

I. INTRODUCTION 1

II. THE APPLICABLE LAW 2

A. The City Bears the Burden of Proving that the Challenged Practices Are Job Related and Consistent with Business Necessity 2

B. The Second Circuit Has Set Forth in Detail What the City Must Prove to Establish Job Relatedness and Business Necessity 3

C. Proof that an Examination Has Content Validity Is Not Sufficient To Establish that the Employer's Particular Use of the Examination Is Job Related and Consistent with Business Necessity 4

1. The employer must establish that the manner in which it uses the examination is job related and consistent with business necessity 4

2. When an employer uses an examination on a pass/fail basis, it must prove that the cutoff score it uses distinguishes between candidates who can and candidates who cannot perform the job 5

3. When an employer uses scores in rank-order, the employer must prove that small differences in the scores used to rank candidates reflect meaningful differences in job performance 7

III. THE CITY HAS NOT PRESENTED EVIDENCE WHICH, CONSIDERED IN THE LIGHT MOST FAVORABLE TO THE CITY, PROVES THAT ANY OF THE CHALLENGED PRACTICES WAS JOB RELATED AND CONSISTENT WITH BUSINESS NECESSITY 7

A. The City's Evidence Does Not Prove that Written Exams 7029 and 2043 Were Job Related under the Standards Set Forth in Guardians 7

1. The City has produced no evidence that "the test-makers . . . conducted a suitable job analysis" 7

2. The City has produced no evidence to establish that it used "reasonable competence" in constructing the written examinations 9

3. The City has presented no evidence that the content of the written examinations is directly "related to" and "representative of" the content of the job 10

B. The City's Evidence Does Not Prove that the Manner in Which It Used the Examinations Was Job Related and Consistent with Business Necessity 13

1. The City has presented no evidence that the pass/fail cutoff scores the City used on Written Exams 7029 and 2043 distinguish between candidates who can and candidates who cannot perform the job 13

2. The City has presented no evidence that differences in the combined written examination and PPT scores the City used to rank-order candidates reflect meaningful differences in job performance 16

a. The City has presented no evidence that the PPT, as scored by the City, was job related 18

b. The City has presented no evidence to justify the very large differences in ranks that resulted from very small differences in scores on the PPT and written examinations 18

IV. CONCLUSION 20

TABLE OF AUTHORITIES

I. FEDERAL CASES

Bridgeport Guardians, Inc. v. City of Bridgeport, 933 F.2d 1140 (2d Cir.), cert. denied, 502 U.S. 924 (1991) 6, 16

Dothard v. Rawlinson, 433 U.S. 321 (1977) 1

Green v. Town of Hamden, 73 F.Supp.2d 192 (D.Conn. 1999) 5, 6, 13

Griggs v. Duke Power Co., 401 U.S. 424 (1971) 1

Guardians Ass'n of the New York City Police Dept. v. Civil Serv. Comm'n of the City of New York, 630 F.2d 79 (2d Cir. 1980), cert. denied, 452 U.S. 940 (1981) . . . . 3, passim

Gulino v. New York State Educ. Dept., 460 F.3d 361 (2d Cir. 2006), cert. denied, 128 S. Ct. 2986 (2008) 1, 2, 4, 5, 11, 17

Lanning v. SEPTA, 181 F.3d 478 (3d Cir. 1999) 6, 13

Pietras v. Bd. of Fire Comm'rs of the Farmingville Fire Dist., 180 F.3d 468, 475 (2d Cir. 1999) 6, 13

II. FEDERAL STATUTES

42 U.S.C. § 2000e(m) 3

42 U.S.C. § 2000e-2(k) 2

42 U.S.C. § 2000e-2(k)(1)(A)(i) 3

42 U.S.C. § 2000e-7 6

III. FEDERAL REGULATIONS

29 C.F.R. § 1607 3

29 C.F.R. § 1607.5 6

29 C.F.R. § 1607.5H 15

29 C.F.R. § 1607.9A 8

I. INTRODUCTION

As articulated by the Supreme Court in Griggs v. Duke Power Co., 401 U.S. 424 (1971) and codified by the Civil Rights Act of 1991, Title VII prohibits employment practices which, although "fair in form," are "discriminatory in operation." Gulino v. New York State Educ. Dept., 460 F.3d 361, 382 (2d Cir. 2006) (quoting Griggs, 401 U.S. at 431), cert. denied, 128 S. Ct. 2986 (2008). Thus, if an employment practice that results in disparate impact "cannot be shown to be related to job performance," it is prohibited. Id. In determining whether a practice that results in disparate impact is unlawful, "the touchstone is business necessity." Griggs, 401 U.S. at 431 (quoted in Gulino, 460 F.3d at 382). Therefore, a practice that results in disparate impact "must be shown to be necessary to safe and efficient job performance to survive a Title VII challenge." Dothard v. Rawlinson, 433 U.S. 321, 331 n.14 (1977).

On February 2, 2009, plaintiffs-intervenors served on defendants a motion for summary judgment with respect to liability. Also on February 2, the United States served on defendant City of New York (the "City") a motion for partial summary judgment with respect to the issue of disparate impact. Prior to that date, despite extensive discovery, including numerous written discovery requests, requests for supplementation and Rule 30(b)(6) and expert depositions, the City had not set forth clearly the evidence or theory on which it intended to rely to prove that the four practices challenged in this case were job related and consistent with business necessity, as required by Title VII.

On February 23, 2009, the City served on the other parties Defendants' Memorandum of Law In Opposition to Plaintiff-Intervenors' Motion for Summary Judgment (referred to herein as "the City's brief" and cited as "Def. Mem."). The City's brief makes clear for the first time that the City has not, and cannot, raise any genuine issue of material fact with respect to job

relatedness and business necessity. The arguments raised by the City in its brief are deficient as a matter of law in two respects. First, the City – not the United States or plaintiffs-intervenors – bears the burden of proof on these issues. Second, clear precedent in this Circuit explains exactly what the City must prove to meet its burden and why the City's proof fails as a matter of law. Accordingly, the United States joins in plaintiffs-intervenors' motion for summary judgment to the extent that it requests that the Court find that the City has not produced evidence sufficient to prove – as the City must – that any of the four practices challenged by the United States and plaintiffs-intervenors were job related and consistent with business necessity.[1]

II. THE APPLICABLE LAW

A. The City Bears the Burden of Proving that the Challenged Practices Are Job Related and Consistent with Business Necessity.

The burdens of proof applicable to disparate impact cases such as this were codified by the 1991 amendments to Title VII. See 42 U.S.C. § 2000e-2(k); Gulino, 460 F.3d at 382. Once the United States establishes that the challenged practices have resulted in a disparate impact, the burden of proof shifts to the City, which must demonstrate that its use of each of the challenged

---

[1] In joining plaintiffs-intervenors' motion, the United States submits no additional evidence and relies in this memorandum only on materials already submitted by plaintiffs-intervenors or the City. In that regard, it should be noted that the United States' reference to various materials (e.g., the report of Frank Landy and the declarations of Catherine Cline and F. Mark Schemmer attached as Exhibits 5, 4 and 2, respectively, to the declaration of the City's counsel ("Fraenkel Decl.")) should not be interpreted as waiving objections to the admission of those materials into evidence. Plaintiffs-intervenors have moved to strike the Cline and Schemmer declarations, and the United States agrees that they are inadmissible. The United States refers to the declarations and other inadmissible materials in this memorandum only to demonstrate that, considering all of the evidence relied upon by the City (admissible or not), the City has failed to raise a genuine issue of material fact. However, if the Court does not grant the motion for summary judgment, the United States intends to file a motion in limine prior to trial requesting that the Court exclude certain documents and testimony.

practices was "job related for the position in question and consistent with business necessity." 42 U.S.C. § 2000e- 2(k)(1)(A)(i). To make it impossible to misinterpret the burdens of proof, the 1991 amendments to Title VII also added to the statute a provision defining "demonstrates" as "meets the burdens of production and persuasion." 42 U.S.C. § 2000e(m).

Remarkably, the City nonetheless misstates the applicable burdens throughout its brief. For example, the City argues that "Plaintiffs have not and cannot demonstrate that the examinations are invalid"[2] and "plaintiffs have not disproven the validity of the examinations." Def. Mem., pp. 13, 15. Similarly, the City asserts, "Unless . . . invalidity is established, a cut score based on hiring needs is permissible." Id., p. 13. Thus, the City attempts to convince the Court that the City can prevail even though it cannot prove that the challenged practices are job related if the United States and plaintiffs-intervenors do not prove that they are not job related. To the contrary, the City bears the burden of proving job relatedness and business necessity.

B. The Second Circuit Has Set Forth in Detail What the City Must Prove to Establish Job Relatedness and Business Necessity.

The City contends that the practices challenged in this case are job related and consistent with business necessity because Written Exams 7029 and 2043 have "content validity." See Def. Mem., p. 4. Nearly thirty years ago, in Guardians, 630 F.2d at 94-99, the Second Circuit explained in detail what the City must prove in order to establish content validity. Less than three years ago, the Second Circuit reaffirmed that courts in "this Circuit remain[] bound by the

---

[2] Employment testing professionals use the terms "job relatedness" and "validity" interchangeably. Such professionals and the Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. § 1607, recognize three methods of proving job relatedness: content validation, criterion-related validation and construct validation. Guardians Ass'n of the New York City Police Dept. v. Civil Serv. Comm'n of the City of New York, 630 F.2d 79, 91 (2d Cir. 1980), cert. denied, 452 U.S. 940 (1981).

validation requirements expressed in earlier Supreme Court precedent . . . as interpreted by Guardians." Gulino, 460 F.3d at 386. Thus, in cases challenging the use of an examination, Guardians continues to "delineate the appropriate standard for assessing job relatedness." Id. Under Guardians, to establish that an examination has content validity, the employer must prove, inter alia, that: (1) "the test-makers . . . conducted a suitable job analysis;" (2) "they . . . . used reasonable competence in constructing the test;" (3) "the content of the test [is] related to the content of the job;" and (4) "the content of the test [is] representative of the content of the job." 630 F.2d at 95.[3] See also Gulino, 460 F.3d at 384, 386 (restating the Guardians requirements).

C. Proof that an Examination Has Content Validity Is Not Sufficient to Establish that the Employer's Particular Use of the Examination Is Job Related and Consistent with Business Necessity.

1. The employer must establish that the manner in which it uses the examination is job related and consistent with business necessity.

The Second Circuit also held in Guardians that an employer cannot prevail merely by

---

[3] The Second Circuit further explained that a suitable job analysis must identify the work behaviors required to perform the job and the "knowledges, skills or abilities" needed for "the effective performance of [the job] tasks." Id. Moreover, "the relationship of abilities to tasks" must be "clearly set forth" in the job analysis. Id. at 96. With regard to the second requirement for content validation, "competence in constructing the test," Guardians explained that "employment testing is a task of sufficient difficulty to suggest that an employer dispenses with expert assistance at [its] peril," and the decision to forgo expert assistance "should require a Court to give the resulting test careful scrutiny." Id. The Court also cautioned against allowing job incumbents (here firefighters) to draft examination questions, explaining that they may have "expertise in identifying tasks involved in their job but [they are] amateurs in the art of test construction." Id. In addition, if the "test construction process" does not "succeed in meeting even its own goal of testing for all the identified abilities" – i.e., if the questions do not measure what they are supposed to – the requirement of competence in test construction has not been met. Id. The Court also explained that the third requirement of content validity involves proof that the examination questions resulting from the test construction process are directly related to the job tasks. Id. at 97-98. Finally, to fulfill the fourth (representativeness) requirement, the Court stated that the employer must show that the examination has not needlessly eliminated "some significant part of the job's requirements." Id. at 99.

showing that the examination at issue has some level of validity. Thus, the Court of Appeals stated, as a fifth requirement, that there must be "a scoring system that usefully selects from among the applicants those who can better perform the job." Guardians, 630 F.2d at 95, quoted in Gulino, 460 F.3d at 384-85. As Guardians explains, when an employer's use of an examination has resulted in a disparate impact, the employer must prove that the manner in which the employer used the examination was job related and consistent with business necessity.

2. When an employer uses an examination on a pass/fail basis, it must prove that the cutoff score it uses distinguishes between candidates who can and candidates who cannot perform the job.

The Second Circuit explained in Guardians that when an employer uses an examination on a pass/fail basis, the "scoring system" requirement means that even if the examination is job related it may nonetheless violate Title VII because of the cutoff score used. Guardians, 630 F.2d at 105. Guardians involved the City's use of a written examination as a pass/fail screening device (as well as for rank-order selection) for entry level police officers. Although the Second Circuit found that, arguably, the City had established that its examination had some level of content validity, the Court nonetheless affirmed the trial court's judgment against the City. As the Court explained:

> No matter how valid the exam, it is the cutoff score that ultimately determines whether a person passes or fails. A cutoff score unrelated to job performance may well lead to the rejection of applicants who were fully capable of performing the job. When a cutoff score unrelated to job performance produces disparate racial results, Title VII is violated.

Guardians, 630 F.2d at 105, quoted in Green v. Town of Hamden, 73 F.Supp.2d 192, 199-201 (D. Conn., 1999). In order to justify its pass/fail use of the written examination at issue, the City was required to prove that the cutoff score it used was "set so as to be . . . consistent with normal

expectations of acceptable proficiency within the work force." Guardians, 630 F.2d at 105 (quoting 29 C.F.R. § 1607.5). Thus, to justify its pass/fail use of Written Exams 7029 and 2043 in this case, the City must prove that the cutoff scores it used (84.705 on Written Exam 7029 and 70 on Written Exam 2043) "reasonably measured the minimum entry skill level necessary for an applicant to be a competent fire fighter." Green, 73 F.Supp.2d at 199. See also id. at 200 (employer must prove cutoff "reasonably delineat[ed] between applicants who were competent to perform the entry level job and those that were not"); Lanning v. SEPTA, 181 F.3d 478 (3d Cir. 1999) (cutoff score must be shown to measure the "minimum qualifications necessary for successful performance of the job"); Pietras v. Board of Fire Commissioners of the Farmingville Fire District, 180 F.3d 468, 472 n.5 (2d Cir.1999) (appeal of liability judgment had no merit because arbitrary cutoff did not reflect the needs of the volunteer firefighter job).

Moreover, in Guardians, 630 F.2d at 105, the Court of Appeals specifically rejected the City's argument that the City was justified in setting its cutoff score based solely on the number of candidates necessary to meet its expected hiring needs.[4] In addition, the Second Circuit rejected reliance on civil service rules as a justification for using an examination in a manner that results in disparate impact, noting that "Title VII explicitly relieves employers from any duty to observe a state hiring provision 'which purports to require or permit' any discriminatory employment practice." Guardians, 630 F.2d at 104-105 (citing 42 U.S.C. § 2000e-7). See also Bridgeport Guardians v. City of Bridgeport, 933 F.2d 1140, 1148 (2d Cir. 1991).

---

[4] As discussed infra, although Guardians acknowledged that, if an employer proves that its practice of rank-order selection on the basis of scores on an examination is appropriate, then it may be permissible to use a cutoff score on that examination based on hiring needs, those circumstances do not exist in this case. See section III.B.1., pp. 13-16, infra.

3. When an employer uses scores in rank-order, the employer must prove that small differences in the scores used to rank candidates reflect meaningful differences in job performance.

Guardians also addressed rank-order selection, rejecting the notion that an employer can justify selecting candidates in rank-order of examination scores merely by showing that the examination has some level of validity. Guardians, 630 F.2d at 101. Thus, the Court's "conclusion that the test itself may have had" some validity "d[id] not . . . lead to approval of using its results for rank-ordered selections." Id. The Court of Appeals explained that a "test may have enough validity for making gross distinctions between those qualified and unqualified for a job" but "be totally inadequate to yield passing grades" that justify rank-ordering. Id. at 100. The Court therefore rejected the City's contention "that if the appropriate abilities [are] tested for, it makes eminent sense to select candidates strictly on the basis of ranked scores." Id. It held instead that "[p]ermissible use of rank-ordering requires a demonstration of such substantial test validity that it is reasonable to expect one- or two-point differences in scores to reflect differences in job performance." Id. at 100-101.

III. THE CITY HAS NOT PRESENTED EVIDENCE WHICH, CONSIDERED IN THE LIGHT MOST FAVORABLE TO THE CITY, PROVES THAT ANY OF THE CHALLENGED PRACTICES WAS JOB RELATED AND CONSISTENT WITH BUSINESS NECESSITY.

A. The City's Evidence Does Not Prove that Written Exams 7029 and 2043 Were Job Related under the Standards Set Forth in Guardians.

1. The City has produced no evidence that "the test-makers . . . conducted a suitable job analysis."

According to the City, it has met the first requirement of Guardians, the requirement of a "suitable job analysis." Def. Mem., p. 4. However, the City has produced no evidence to support that contention. The City apparently recognizes that the "job analysis" work performed by the

City's Department of Citywide Administrative Services ("DCAS") in the course of developing Written Exam 7029 was insufficient.[5] Thus, the City admits that it did not conduct a "full scale" job analysis prior to creating Exam 7029." Id., p. 7. Indeed, the City admits that "in terms of documentation, procedures and the mechanics of development, the development of exams 7029 and 2043 were far from ideal." Id., p. 6. According to the City, however, DCAS "built on" and "updated" a job analysis purportedly performed by a consultant, Dr. Frank Landy, in connection with an earlier examination. Id., p. 7. The City says nothing about the quality of the "updating" done by DCAS – other than admitting that it was "far from ideal" and "lacking." Id., p. 6.

The City nonetheless asserts that it has established that a suitable job analysis was conducted, fulfilling the first requirement of Guardians, because one of its experts, Dr. Schemmer, "declares that given Dr. Landy's stature in the field it would be hard to imagine that one of Dr. Landy's studies would possess substantial defects." Def. Mem., p. 5. The Uniform Guidelines specifically reject such statements as evidence of validity: "Under no circumstances will the general reputation of a test . . . its author or its publisher, or casual reports of its validity be accepted in lieu of evidence of validity." 29 C.F.R. § 1607.9A. Regardless of Dr. Landy's "stature," the bare statement that "it is hard to imagine" that a job analysis he conducted would have "substantial defects" is a far cry from evidence that he in fact conducted an appropriate job analysis – much less that the "less than ideal" updating performed by DCAS was appropriate. In fact, Dr. Landy's own report regarding the work he did for the City (which is labeled on its front cover a "Draft") explicitly states that he was not able to complete a job analysis because the

[5] The City admits that it did not conduct a job analysis in connection with the development of Written Exam 2043. Id., p. 5 (asserting that "it was reasonable for DCAS . . . to not conduct a new job analysis for the development of Exam 2043.")

firefighters' union refused to cooperate, and only 217 of the 5,500 job analysis questionnaires Dr. Landy sent to FDNY firefighters were completed and returned. Fraenkel Decl., Exh. 5, pp. 1-2, 14.[6] Moreover, of the 217, only 184 were considered usable by Dr. Landy. Id., pp. 14-15.

2. The City has produced no evidence to establish that it used "reasonable competence" in constructing the written examinations.

As stated above, under Guardians, the second element of content validity that the City must prove is that it "used reasonable competence in constructing" the examinations. The City's brief says nothing about how the examinations were constructed. This is understandable because the City constructed Written Exams 7029 and 2043 using a practice that Guardians cautioned against. Specifically, the City asked firefighters, "who may have had expertise in identifying tasks involved in their job but were amateurs in the art of test construction," Guardians, 630 F.2d at 96, to write examination questions to measure nine abilities that the City contends are the "content" of Written Exams 7029 and 2043 (Written Comprehension, Written Expression, Memorization, Problem Sensitivity, Deductive Reasoning, Inductive Reasoning, Information Ordering, Spatial Orientation, and Visualization). Fraenkel Decl., Exh. 22, pp. 339:19-340:13, 354:8-355:7; Declaration of Richard A. Levy ("Levy Decl."), submitted in support of plaintiffs-intervenors' motion for summary judgment , App. M Resp. to Request 28. The City nonetheless argues that it has fulfilled the second requirement of Guardians because Dr. Schemmer states in

[6] Later in its brief, the City elevates the Landy "job analysis" to a "validity study." See, e.g., Def. Mem., p. 14. However, while Dr. Landy's report states that he planned to conduct (and even began) a criterion-related validity study, apparently he did not complete it. Fraenkel Decl., Exh. 5, p. 30. No report of such a validity study was ever produced in discovery.

his declaration (though not in his expert report or at deposition)[7] that "Exams 7029 and 2043, are in content and substance very representative of entry level firefighter selection exams which used more rigorous methods and which were thoroughly documented." Def. Mem., p. 6. Of course, the City has presented no evidence that these other "firefighter selection exams" were valid. More importantly, Dr. Schemmer's bare statement that Written Exams 7029 and 2043 are like other firefighter examinations is an insufficient basis for a finding that the City must have used reasonable competence in creating its written examinations.

3. The City has presented no evidence that the content of the examinations is directly "related to" and "representative of" the content of the job.

With respect to the third and fourth elements of a showing of content validity, the City has presented no evidence that the content of Written Exams 7029 and 2043 is directly related to the content of the firefighter job or that the content of the written examinations is representative of the content of the job. Instead, the City asserts that the United States and plaintiffs-intervenors have not established that the content of the examinations is not directly related to and representative of the content of the firefighter job. Def. Mem., p. 8 ("Plaintiffs make several unavailing attempts to disprove the content relatedness of Exams 7029 and 2043.") This, of course, is yet another example of the City misstating the burdens of proof. The City has simply produced no evidence that would allow the Court to find that the City has proved, as it must, that the content of the examinations is related to and representative of the content of the job.

---

[7] Indeed, in Dr. Schemmer's deposition, which was taken over five months after his report was completed, he could not offer an opinion regarding the content or "substance" of the written examinations. Dr. Schemmer testified that he did not know whether the items on the written examinations tested for the abilities they were intended to measure and that he had looked at perhaps only 10-20 of the 170 questions on the two written examinations. Fraenkel Decl., Exh. 8, pp. 341:3-342:9.

Indeed, given the unrebutted statistical analysis conducted by one of the United States' experts, Dr. Bernard Siskin, the City cannot even establish what the content of the test is. According to the City, Written Exams 7029 and 2043 measured nine abilities that are important for performance of the firefighter job. See p. 9, supra. (listing the nine abilities). Clearly, if Written Exams 7029 and 2043 do not measure those abilities, the content of the examinations is unknown and the City cannot prove that it is related to and representative of the content of the job.[8] As plaintiffs-intervenors explained in their memorandum, Dr. Siskin demonstrated in his report that Written Exams 7029 and 2043 "did not . . . succeed in meeting even [the City's] own goal of testing for . . . the [nine] identified abilities." Guardians at 96; see plaintiffs-intervenors' Memorandum of Law in Support of Motion for Summary Judgment ("Pl.-Int. Mem."), p. 28.

Dr. Siskin explained that one can statistically determine the extent to which the City succeeded in measuring the nine abilities by studying the patterns of correlations between the items (i.e., questions) on the examinations. Levy Decl., App. V, p. 5. In this context, a correlation is a measure of the relationship between examination questions. Correlation values range from 0 (indicating no relationship) to 1 (indicating perfect relationship). Id. Dr. Siskin explained that, if the items on Written Exam 7029 or 2043 measure nine different, distinguishable abilities as the City intended, then the items that, according to the City, measure the same ability should relate more closely to each other than to items that measure a different ability. Id. For example, the items that were intended to measure Deductive Reasoning should

---

[8] The City does not, and cannot, argue that, even though the written examinations did not measure the nine abilities they were intended to measure, they are content valid because they measured general cognitive ability. The Second Circuit made clear in Guardians – and recently reaffirmed in Gulino – that "attempts to measure general qualities such as intelligence" cannot be justified using a content validation strategy. Guardians, 630 F.2d at 93; Gulino, 460 F.3d at 384.

correlate more highly with other items intended to measure Deductive Reasoning than with items intended to measure Spatial Orientation. To the contrary, however, Dr. Siskin's correlation analyses showed that, with respect to Written Exam 7029, for eight of the nine abilities, the items correlate as or more highly with items intended to measure a different ability than they do with items intended to measure the same ability.[9] Id., pp. 5-6 ¶¶ 8-9. Similarly, for Written Exam 2043, items that purport to measure four of the nine abilities correlate more highly with items intended to measure a different ability than with items intended to measure the same ability.[10] Id., p. 6 ¶ 9 Although the City attempted to rebut another analysis done by Dr. Siskin, a "factor analysis," Def. Mem., p. 9, the City understandably made no attempt to rebut Dr. Siskin's correlations analysis. At deposition, both the City's designated expert, Dr. Philip Bobko, and Dr. Cline agreed that, if the items on the written examinations had measured nine different, distinguishable abilities, then the patterns of correlations that Dr. Siskin found should not have occurred. Fraenkel Decl., Exh. 15, pp. 188:17-189:18 and Exh. 14, pp. 322:16-23, 325:6-22.

---

[9] For example, items that are supposed to measure Deductive Reasoning correlate more highly with items that supposedly measure Written Expression (correlation = .153), Written Comprehension (.148) and Spatial Orientation (.147) than they do with other items intended to measure Deductive Reasoning (.139). Id., pp. 5-6 ¶ 8.

[10] For the other five abilities, the correlation with items intended to measure the same ability was barely above the correlation with items intended to measure different abilities. Id. For example, the correlation between Inductive Reasoning items was .222, and the correlation between Inductive Reasoning and Information Ordering items was .215. Id., pp. 6-7 ¶ 9.

B. The City's Evidence Does Not Prove that the Manner in Which it Used the Examinations Was Job Related and Consistent with Business Necessity.

1. The City has presented no evidence that the pass/fail cutoff scores the City used on Written Exams 7029 and 2043 distinguish between candidates who can and candidates who cannot perform the job.

Even if the City could prove that Written Exams 7029 and 2043 had some (unquantified) level of job relatedness, summary judgment still would be appropriate because the City has presented no evidence establishing, as it must, that its pass/fail use of the written examinations with the cutoff scores it used was job related and consistent with business necessity. Again misstating both Title VII's burdens of proof and the law in this Circuit, the City asserts, without citation, that "[u]nless and until invalidity is established, a cut score based on hiring needs is permissible." Def. Mem., p. 13. As explained above, however, even assuming the City could prove that the written examinations were valid, the City also must establish that the cutoff scores correspond to the minimum level of the tested abilities necessary to perform the firefighter job successfully. Guardians, 630 F.2d at 105; Green, 73 F.Supp.2d at 199-200; Lanning, 181 F.3d at 489; Pietras, 180 F.3d at 472 n.5. But the City does not even contend that either the cutoff score it used on Written Exam 7029 (84.705) or the cutoff score it used on Written Exam 2043 (70) corresponds to the minimum level of the tested abilities that is necessary to perform the firefighter job successfully. Levy Decl., App. N (Resp. to Interrogatory 30).

Indeed, the City has admitted that the cutoff scores were not selected based on any connection between the scores and the ability to perform the job. It is undisputed that the City selected the cutoff score of 84.705 on Written Exam 7029 – as it had selected the cutoff score the Second Circuit found unjustified in Guardians, 630 F.2d at 105 – solely based on "hiring needs," to allow a given number of candidates to pass the written examination and continue on in the

selection process.[11] Def. Mem., p. 12. Guardians did acknowledge that, if an employer proves

---

[11] Ignoring Guardians, the City argues that Dr. Cline and one of the United States' experts, Dr. Jones, endorse the "hiring needs" method of setting a cutoff score. Def. Mem., pp. 11-12. Given Guardians, whether these individuals approve of such a method is not material. Nonetheless, it must be noted that the City relies on what can most generously be described as misrepresentations as to the deposition testimony of Drs. Jones and Cline. For example, the City states, "According to Dr. Jones one method for establishing a cut-score, which has business justification, is to base[] it on hiring needs." Id., p. 11. In context, however, Dr. Jones' testimony is to the contrary:

> Q: What would be wrong with using the . . . anticipated need of hires or future potential candidates as the basis for establishing the cut-off score?
>
> A: Well, I think it can be a consideration but using it to set the cut-off . . . bears no relationship to the standard that people in my profession like to see, and that is one of job-relatedness. To say we need to have X number of people available is not [a] statement of job-relatedness at all. . . . All we know is that we set the bar so that we get enough people through the funnel, and that's not what I would consider accepted practice in our discipline.

Fraenkel Decl., Exh. 13 pp. 80:24-81:16. Moments later, Dr. Jones also testified: "If you set a qualifying standard such that you run out of candidates in the first two months you use the exam and you can't administer another exam for two years, well, of course, you've created a problem for yourself. . . . So I think business, and business needs always have some influence on what's done. But when qualifying standards are set and, in particular, when those standards have an adverse impact . . . I think there needs to be a f[ar] sounder rationale for them than 'business need.'" Id., pp. 81:17-82:9. See also, id., pp. 79:22-80:23. Similarly, according to the City, Dr. Cline "say[s] that hiring needs is an appropriate basis for setting the cut score." Id., p. 12. In fact, Dr. Cline testified: "Well, . . . even if you put people on the list, they're not going to get hired. So if they only have a certain demand, then only a certain percentage of the candidate population is going to be hired." Fraenkel Decl., Exh.14, pp. 458:17-459:6. The following exchange then occurred:

> Q: Apart from the market forces and the numbers that you might want to hire . . ., the question was . . ., from a performance standpoint, does it make sense to you . . . that you would have a higher . . . test standard for firefighters than for police officers given their respective responsibilities and roles?
>
> A: Given their respective responsibilities and roles, it doesn't make any sense, but it's what happens in the supply and demand. . . . So the firefighters are really overqualified. They really probably are smarter than they need to be in some ways. I shouldn't have said that. That was all speculation, too.

that rank order selection on the basis of the examination at issue is justified, then the employer may use a cutoff score higher than that corresponding to the minimum level of the tested abilities needed to perform the job successfully.[12] This is consistent with the Uniform Guidelines provision cited by the City, Def. Mem., p. 12, which provides: "Where applicants are ranked on the basis of properly validated selection procedures and those applicants scoring below a higher cutoff score than appropriate in light of such expectations have little or no chance of being selected for employment, the higher cutoff score may be appropriate, but the degree of adverse impact should be considered." 29 C.F.R. § 1607.5H. However, the City has produced no evidence that candidates who scored below the cutoff of 84.705 on Written Exam 7029 had "little or no chance of being selected." Indeed, as the United States' statistical expert, Dr. Siskin,

---

Id., pp. 459:11-460:10.

[12] However, Guardians made clear that where, as here, a claim that rank-ordering would have been justified is based on a content validity argument, that claim "must be closely scrutinized." 630 F.2d at 100. As the Second Circuit explained:

> This close scrutiny is required because rank-ordering makes such a refined use of the test's basic power to distinguish between those who are qualified to perform the job and those who are not. If a test is content valid, it may be reasonable to infer that the test scores make some useful gross distinctions between candidates. . . . But [that does not] provide[] confidence for inferring that one-point increments . . . are a valid basis for making job-related hiring decisions. . . . A test may have enough validity for making gross distinctions between those qualified and unqualified for a job, yet may be totally inadequate to yield passing grades that show positive correlation with job performance."

Id. In this case, as in Guardians, the City has "never established a valid basis for rank-ordering" on the basis of the written examination scores. Id. at 105. Indeed, as explained above, the City did not rank-order candidates' based on written examination scores. It ranked them based on a combination of their written examination and physical performance test (PPT) scores, implicitly acknowledging that, given the importance of physical abilities to the firefighter job, ranking based on candidates' scores on Written Exam 7029 or 2043 would not be justified.

showed in his report, a candidate who scored 83.529 on Written Exam 7029 and would have scored 100 on the PPT (if the City had allowed him to take it) would have a combined score of 83.652, much higher than that of a candidate who scored 84.705 and received a score of 75 on the PPT. Pl.-Int. Rule 56.1 Stmt., ¶ 65; Def. Resp. to Pl.-Int. Rule 56.1 Stmt., ¶ 65. The latter candidate would have received a combined score of 70.000. Id. Thus, the candidate who scored 83.529 and failed Written Exam 7029 would have been ranked substantially higher on the eligibility list and had a substantially better chance of being hired than the candidate who scored 84.705 and passed Written Exam 7029. Where candidates who failed the written examination would have had a better chance of hire than candidates who passed it, the City cannot logically justify its pass/fail cutoff on the theory that those who failed would have had "little or no chance of being selected." Thus, the City cannot justify its use of a cutoff score of 84.705 on Written Exam 7029 based on "hiring needs."

With respect to Written Exam 2043, the City did not select the cutoff score of 70 for that examination based on hiring needs. It selected that cutoff solely because 70 is the "default" cutoff score under the City's Civil Service Rules. Pl.-Int. Rule 56.1 Stmt., ¶ 63; Def. Resp. to Pl.-Int. Rule 56.1 Stmt., ¶ 63. As stated previously, in Guardians, 630 F.2d at 104-105, the Second Circuit rejected reliance on civil service rules as a justification for the use of a cutoff score that results in a disparate impact. See Bridgeport Guardians, 933 F.2d at 1148.

2. The City has presented no evidence that differences in the combined written examination and PPT scores the City used to rank-order candidates reflect meaningful differences in job performance.

Finally, the City has not even attempted to justify the manner in which it actually ranked

candidates on the Exam 7029 and 2043 eligibility lists.[13] As stated previously, see n..12, supra, the City ranked candidates based on a combination of their written examination and PPT scores. The Second Circuit made clear in Gulino that the City cannot establish that its use of a combination of written examination and PPT scores was job related merely by showing that one component (i.e., the written examination) of the combined scores was job related. Gulino, 460 F.3d at 387. See also Fraenkel Decl., Exh. 8, p. 247:5-22 (agreeing). Yet the City has presented no evidence that the combination of written examination and PPT scores it used to rank candidates was valid.[14]

---

[13] As explained above, Guardians, 630 F.2d at 100-01, held that proof of validity is not sufficient to justify rank order processing and selection. Therefore, the City's citation to the testimony of Dr. Hough, one of the United States' experts, in support of its assertion that "[i]f an examination is valid, rank ordering and selection of candidates is permissible," Def. Mem., p. 13, is not material. Nonetheless, it must be noted that the City's characterization of Dr. Hough's testimony is not accurate. According to the City, it is Dr. Hough's opinion that, "given a valid examination, rank ordering can be used." Def. Mem., p. 13 (citing Dr. Hough's deposition at 148:7-19). What Dr. Hough actually said, however, was limited to tests that (unlike the combined scores at issue here) have been shown to have a sufficient level of criterion-related validity – and then only under some circumstances: "I'd say assuming that the test is valid, . . . that there is a relationship, in fact, between the predictor and the criterion, and it's a good level of validity such that you have confidence in the accuracy of your predictions that a rank-order has merit [under] many circumstances." Fraenkel Decl., Exh. 17 pp. 147:24-148:12.

[14] The City does attempt to rely on the deposition testimony of three witnesses, asserting that "they believe that people higher on the list are likely to perform better than those ranked lower down." Def. Mem., p. 16. Even if this testimony were admissible, it would be insufficient to meet the standard of Guardians. Nonetheless, it should be noted that these witnesses are not qualified to present expert opinions and were not designated by the City as experts. Two of the witnesses testified that their opinions were not based on personal knowledge. See Fraenkel Decl., Exh. 27, pp. 107:23-109:25 (one of the witnesses had not "actually seen" that people higher on the list performed better "in [his] experience of working with firefighters" because "[a] new guy walks in the door tomorrow, I don't know if he's on the top of the list or the bottom of the list"); Fraenkel Decl., Exh. 28, pp. 72:13-73:12 (when the witness worked in the firehouse he did not "experience that those people who had scored high on the test were better firefighters than those who had not scored as high" because "[n]o one in the firehouse . . . knew anybody's test number"). Only former Chief Daniel Nigro testified that he had actually observed some

a. The City has presented no evidence that the PPT, as scored by the City, was job related.

Indeed, the City does not even argue that the PPT was job related. The PPT consisted of eight events, each with its own pass/fail cutoff. Pl.-Int. Rule 56.1 Stmt., ¶ 19; Def. Resp. to Pl.-Int. Rule 56.1 Stmt., ¶ 19. Candidates were required to pass at least six of the eight events in order to pass the PPT and be ranked on the eligibility list. Id. Candidates who passed the PPT received a score of 75, 87.5 or 100, depending on whether they passed six, seven or all eight events. Id. There is no dispute about the scoring method the City used. Id. Given the scoring method, a candidate who excelled on all eight PPT events received the same score (100) as one who barely passed all eight. Moreover, a candidate who failed one event (even by one second) would score 12.5 points (the difference between a score of 100 and a score of 87.5) below a candidate who passed all eight events – even if the candidate who failed one event performed much better on the other seven events than the candidate who passed all eight. The City has not even attempted to explain how, given the scoring method it used, candidates' scores on the PPT related in any meaningful way to their expected job performance.

b. The City has presented no evidence to justify the very large differences in ranks that resulted from very small differences in candidates' scores on the PPT and written examinations.

Moreover, as the United States' statistical expert, Dr. Siskin, explained in his report, given the way in which the City scored and combined the written examinations and the PPT, very

---

difference between candidates hired in the early years of a list (i.e., those ranked at the top) and those hired in the later years. However, Chief Nigro explained that the difference was due to the delay caused by the challenged practice itself. Fraenkel Decl., Exh. 29, pp. 107:15-109:20 ("[I]t's sometimes more difficult at the bottom because the people are starting to get older and getting more set in their ways of doing things. . . . So someone that's been in the workforce longer and is older is harder to train . . . than somebody who's 21 years old . . . .")

small differences in performance on the PPT resulted in very large differences in ranks on both the Exam 7029 and 2043 eligibility lists. Levy Decl., App. V, pp. 22-24 ¶¶ 33-35. For example, Dr. Siskin showed that when candidates are ranked based on their Written Exam 7029 and PPT scores, a candidate who passed all eight of the PPT events could rank as high as first, but a candidate who failed one event (even by one second) could not rank above 3,603 on the list. Id., p. 22 ¶ 33. A candidate who failed two PPT events could not rank above 5,471. Id. Likewise, for Exam 2043, a candidate who passed all eight PPT events could rank as high as first, but a candidate who failed one event (even by one second) could not rank above 3,439. Id., p. 23 ¶ 35. A candidate who failed two PPT events could not rank above 5,776 on the Exam 2043 eligibility list. Id.

Similarly, very small differences in candidates' scores on the written examination had a very large effect on candidates' ranks on the eligibility lists. Id., pp. 22-24 ¶¶ 33-35. For example, Dr. Siskin showed that when candidates are ranked based on their written examination and PPT scores, a candidate who answered all of the questions on Written Exam 7029 correctly could rank as high as first, but a candidate who answered four questions incorrectly (receiving a written examination score of 95.294) on Written Exam 7029 could not rank above 2,454. Id., p. 23 ¶ 34. A candidate who answered four questions on Written Exam 2043 incorrectly could not rank above 1,713. Id., p. 23 ¶ 35. The City does not dispute that differences of four questions on the written examinations are within the range of variability that can be expected due solely to chance.[15] See Pl.-Int. Rule 56.1 Stmt., ¶ 69; Def. Resp. to Pl.-Int. Rule 56.1 Stmt., ¶ 69. In other

[15] There is no evidence regarding the standard error of measurement – the "margin of error" – of the PPT. Obviously, however, differences of seconds could easily be due to chance.

words, chance reasonably could have accounted for a difference of thousands of places between two candidates on the eligibility lists.[16] Thus, the City cannot bear its burden of proving that its practice of rank-order processing and selection of candidates from either the Exam 7029 and 2043 eligibility list was job related and consistent with business necessity.

IV. CONCLUSION

For the reasons set forth above, in the United States' Memorandum of Law in Support of Motion for Partial Summary Judgment and Reply Memorandum in Further Support of Motion for Partial Summary Judgment (dated February 2 and March 9, 2009, respectively), there is no need for trial. There is no question that each of the practices at issue has resulted in a disparate impact upon blacks and Hispanics – indeed, the City has conceded most of the prima facie case. Moreover, the City has presented no evidence that would allow it to carry its burden of proving that the practices were job related and consistent with business necessity. Accordingly, the Court should hold that the City's use of the challenged practices has resulted in an unlawful disparate impact upon black and Hispanic candidates for the position of entry-level firefighter and grant summary judgment in favor of the United States and plaintiffs-intervenors with respect to liability.

Dated: March 9, 2009

Respectfully submitted,

LORETTA KING
Acting Assistant Attorney General

---

[16] In response to this undisputed fact, the City merely asserts that, "[u]nless an exam is perfect it will have some margin of error. . . there will always be a risk of choosing between candidates whose scores, though different, may fall within the margin of error." Def. Mem., p. 15. While true, the City's assertion does not justify scoring the examinations and combining candidates' written examination and PPT scores in a way that causes minor differences in written examination or PPT scores to result in differences of thousands of places on the eligibility lists.

Sharon A. Seeley
Special Litigation Counsel
(202) 514-4761
Clare Geller
Varda Hussain
Carolyn Weiss
David N. Reese
Trial Attorneys
United States Department of Justice
Civil Rights Division
950 Pennsylvania Avenue, N.W.
Patrick Henry Building, Room 4500
Washington, D.C. 20530
Telephone: (202) 514-4761
Facsimile: (202) 514-1005

BENTON J. CAMPBELL
United States Attorney

Elliot M. Schachner
David Eskew
Assistant United States Attorneys
Eastern District of New York
271 Cadman Plaza East
Brooklyn, New York 11201-1820
Telephone: (718) 254-6053
Facsimile: (718) 254-6479