

IN THE UNITED STATES COURT OF APPEALS
FOR THE FIRST CIRCUIT

PEDRO LOPEZ, individually and on behalf of a class of individuals similarly situated; ABEL CANO, individually and on behalf of a class of individuals similarly situated; KEVIN SLEDGE, individually and on behalf of a class of individuals similarly situated; CHARLES DEJESUS, individually and on behalf of a class of individuals similarly situated; RICHARD BROOKS, individually and on behalf of a class of individuals similarly situated; THE MASSACHUSETTS HISPANIC LAW ENFORCEMENT ASSOCIATION, individually and on behalf of a class of individuals similarly situated;

(See inside cover for continuation of caption)

ON APPEAL FROM THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF MASSACHUSETTS
DISTRICT COURT NO. 07-11693-GAO
(HON. GEORGE A. O'TOOLE, JR.)

BRIEF FOR THE UNITED STATES AS AMICUS CURIAE
SUPPORTING PLAINTIFFS-APPELLANTS AND URGING VACATUR

P. DAVID LOPEZ
General Counsel

VANITA GUPTA
Acting Assistant Attorney General

CAROLYN L. WHEELER
Acting Associate General Counsel
Appellate Services

SHARON M. MCGOWAN
BONNIE I. ROBIN-VERGEER
Attorneys

Equal Employment Opportunity
Commission
131 M Street N.E.
Washington, DC 20507
(202) 663-4739

Department of Justice
Civil Rights Division
Appellate Section
Ben Franklin Station
P.O. Box 14403
Washington, DC 20044-4403
(202) 353-2464

(Continuation of caption)

ROBERT ALVAREZ, individually and on behalf of a class of individuals similarly situated; SPENCER TATUM, individually and on behalf of a class of individuals similarly situated; SHUMEAND BENFOLD, individually and on behalf of a class of individuals similarly situated; ANGELA WILLIAMS-MITCHELL, individually and on behalf of a class of individuals similarly situated; GWENDOLYN BROWN, individually and on behalf of a class of individuals similarly situated; LYNETTE PRAILEAU, individually and on behalf of a class of individuals similarly situated; TYRONE SMITH, individually and on behalf of a class of individuals similarly situated; EDDY CHRISPIN, individually and on behalf of a class of individuals similarly situated; DAVID E. MELVIN, individually and on behalf of a class of individuals similarly situated; STEVEN MORGAN, individually and on behalf of a class of individuals similarly situated; WILLIAM E. IRAOLO, individually and on behalf of a class of individuals similarly situated; JOSE LOZANO, individually and on behalf of a class of individuals similarly situated; COURTNEY A. POWELL, individually and on behalf of a class of individuals similarly situated; JAMES L. BROWN, individually and on behalf of a class of individuals similarly situated; GEORGE CARDOZA, individually and on behalf of a class of individuals similarly situated; LARRY ELLISON, individually and on behalf of a class of individuals similarly situated; DAVID SINGLETARY, individually and on behalf of a class of individuals similarly situated; CHARISSE BRITTLE POWELL, individually and on behalf of a class of individuals similarly situated; CATHENIA D. COOPER-PATERSON, individually and on behalf of a class of individuals similarly situated; MOLWYN SHAW, individually and on behalf of a class of individuals similarly situated; LAMONT ANDERSON, individually and on behalf of a class of individuals similarly situated; GLORIA KINKEAD, individually and on behalf of a class of individuals similarly situated; KENNETH GAINES, individually and on behalf of a class of individuals similarly situated; MURPHY GREGORY, individually and on behalf of a class of individuals similarly situated; JULIAN TURNER, individually and on behalf of a class of individuals similarly situated; NEVA GRICE, individually and on behalf of a class of individuals similarly situated; DELORES E. FACEY, individually and on behalf of a class of individuals similarly situated; LISA VENUS, individually and on behalf of a class of individuals similarly situated;

(caption continued)

(Continuation of caption)

RODNEY O. BEST, individually and on behalf of a class of individuals similarly situated; KAREN VANDYKE, individually and on behalf of a class of individuals similarly situated; ROBERT C. YOUNG, individually and on behalf of a class of individuals similarly situated; ROYLINE LAMB, individually and on behalf of a class of individuals similarly situated; LYNN DAVIS, individually and on behalf of a class of individuals similarly situated; JAMES A. JACKSON, individually and on behalf of a class of individuals similarly situated; JUAN ROSARIO, individually and on behalf of a class of individuals similarly situated; LOUIS ROSARIO, JR., individually and on behalf of a class of individuals similarly situated; OBED ALMEYDA, individually and on behalf of a class of individuals similarly situated; DEVON WILLIAMS, individually and on behalf of a class of individuals similarly situated; JULIO M. TOLEDO, individually and on behalf of a class of individuals similarly situated

Plaintiffs-Appellants

MARISOL NOBREGA, individually and on behalf of a class of individuals similarly situated

Plaintiff

v.

CITY OF LAWRENCE, MASSACHUSETTS; CITY OF METHUEN, MASSACHUSETTS; COMMONWEALTH OF MASSACHUSETTS; PAUL DIETL, in his capacity as Personnel Administrator for the Commonwealth of Massachusetts; JOHN MICHAEL SULLIVAN, in his capacity as Mayor of the City of Lawrence, Massachusetts; WILLIAM MANZI, III, in his capacity as Mayor of the City of Methuen, Massachusetts; CITY OF LOWELL; CITY OF WORCESTER, MASSACHUSETTS; APPOINTING AUTHORITY FOR THE CITY OF LOWELL;

(caption continued)

(Continuation of caption)

MICHAEL O'BRIEN, in his capacity as City Manager of the City of Worcester, Massachusetts; CITY OF BOSTON, MA; CITY OF SPRINGFIELD, MA; DOMENIC J. SARNO, JR., in his capacity as Mayor for the City of Springfield; MASSACHUSETTS BAY TRANSPORTATION AUTHORITY; DANIEL GRABAUSKAS, in his capacity as General Manager; BOARD OF TRUSTEES OF THE MASSACHUSETTS BAY TRANSPORTATION AUTHORITY,

Defendants-Appellees

WILLIAM F. MARTIN, in his capacity as Mayor of the City of Lowell, Massachusetts; KONSTANTINA B. LUKES, in her capacity as Mayor of the City of Worcester, Massachusetts,

Defendants

TABLE OF CONTENTS

	PAGE
INTEREST OF THE UNITED STATES	1
QUESTIONS PRESENTED.....	1
STATEMENT OF THE CASE.....	1
SUMMARY OF ARGUMENT	7
ARGUMENT	
I THE DISTRICT COURT DID NOT PROPERLY ANALYZE WHETHER BOSTON’S EXAMINATIONS FOR POLICE SERGEANT, AS CONSTRUCTED AND USED, WERE VALID	12
A. <i>The District Court Failed To Analyze Correctly Whether Boston’s Exams Were Content-Valid</i>	13
B. <i>The District Court Failed To Consider Whether The Selection Of A Cutoff Score And Rank-Order Use Of Exam Results Were Valid</i>	20
II THE COURT DID NOT CORRECTLY ANALYZE WHETHER PLAINTIFFS ESTABLISHED A PRIMA FACIE CASE OF DISPARATE IMPACT AGAINST THE NON-BOSTON DEFENDANTS	24
CONCLUSION	32
CERTIFICATE OF COMPLIANCE	
CERTIFICATE OF SERVICE	
ADDENDUM	

TABLE OF AUTHORITIES

CASES:	PAGE
<i>Bazile v. City of Houston</i> , 858 F. Supp. 2d 718 (S.D. Tex. 2012)	29
<i>Boston Chapter, NAACP, Inc. v. Beecher</i> , 371 F. Supp. 507 (D. Mass.), aff'd, 504 F.2d 1017 (1st Cir. 1974).....	14
<i>Boston Chapter, NAACP, Inc. v. Beecher</i> , 504 F.2d 1017 (1st Cir. 1974).....	<i>passim</i>
<i>Boston Police Superior Officers Fed'n v. Civil Serv. Comm'n</i> , !35 Mass. App. Ct. 688, 624 N.E.2d 617 (1993).....	15
<i>Bouman v. Block</i> , 940 F.2d 1211 (9th Cir. 1991).....	30
<i>Bradley v. City of Lynn</i> , 443 F. Supp. 2d 145 (D. Mass. 2006)	21
<i>Bridgeport Guardians, Inc. v. City of Bridgeport</i> , 933 F.2d 1140 (2d Cir. 1991)	24, 26
<i>EEOC v. Steamship Clerks Union</i> , 48 F.3d 594 (1st Cir. 1995)	10-11, 25, 31
<i>Firefighters Inst. for Racial Equal. v. City of St. Louis</i> , 549 F.2d 506 (8th Cir. 1977)	14
<i>Fudge v. City of Providence Fire Dep't</i> , 766 F.2d 650 (1st Cir. 1985)	10, 26
<i>Green v. Town of Hamden</i> , 73 F. Supp. 2d 192 (D. Conn. 1999).....	21
<i>Guardians Ass'n of New York City Police Dep't v. Civil Serv.</i> <i>Comm'n of City of New York</i> , 630 F.2d 79 (2d Cir. 1980).....	<i>passim</i>
<i>Guardians Ass'n of New York City Police Dep't v. Civil Serv.</i> <i>Comm'n of City of New York</i> , 633 F.2d 232 (2d Cir. 1980).....	14
<i>International Bhd. of Teamsters v. United States</i> , 431 U.S. 324 (1977).....	25

CASES (continued):	PAGE
<i>Isabel v. City of Memphis</i> , 404 F.3d 404 (6th Cir. 2005)	21, 26
<i>Johnson v. City of Memphis</i> , 770 F.3d 464 (6th Cir. 2014)	22
<i>Jones v. City of Boston</i> , 752 F.3d 38 (1st Cir. 2014).....	11, 27, 31
<i>Lanning v. Southeastern Pa. Transp. Auth.</i> , 181 F.3d 478 (3d Cir. 1999).....	9, 21
<i>Legault v. aRusso</i> , 842 F. Supp. 1479 (D.N.H. 1994).....	14, 16-17, 19
<i>Lopez v. Massachusetts</i> , 588 F.3d 69 (1st Cir. 2009).....	2
<i>Paige v. California</i> , 291 F.3d 1141 (9th Cir. 2002)	28, 30
<i>Ricci v. DeStefano</i> , 557 U.S. 557 (2009).....	12
<i>United States v. City of New York</i> , 637 F. Supp. 2d 77 (E.D.N.Y. 2009)	14, 22
<i>United States v. City of New York</i> , 731 F. Supp. 2d 291 (E.D.N.Y. 2010).....	16, 20
<i>United States v. Massachusetts</i> , 781 F. Supp. 2d 1 (D. Mass. 2011).....	17
<i>Vulcan Pioneers, Inc. v. New Jersey Dep't of Civil Serv.</i> , 625 F. Supp. 527 (D.N.J. 1985), aff'd, 832 F.2d 811 (3d Cir. 1987)	28

STATUTES:

Title VII of the Civil Rights Act of 1964, 42 U.S.C. 2000e <i>et seq.</i>	1
42 U.S.C. 2000e-2(k)(1)(A)	2
42 U.S.C. 2000e-2(k)(1)(A)(i).....	5, 7, 12
42 U.S.C. 2000e-5(a)	1
42 U.S.C. 2000e-5(f)(1).....	1
42 U.S.C. 2000e-7	24
42 U.S.C. 2000e-16	1

REGULATIONS:	PAGE
29 C.F.R. 1607.4(D).....	<i>passim</i>
29 C.F.R. 1607.5(A).....	5
29 C.F.R. 1607.5(B).....	5, 13
29 C.F.R. 1607.5(G).....	20
29 C.F.R. 1607.5(H).....	9, 20
29 C.F.R. 1607.14(C).....	5
29 C.F.R. 1607.14(C)(1)	13
29 C.F.R. 1607.14(C)(2)	14
29 C.F.R. 1607.14(C)(4)	8, 13, 16
29 C.F.R. 1607.14(C)(6)	18
29 C.F.R. 1607.14(C)(9)	7, 9, 20, 22
29 C.F.R. 1607.15(C)(4)	8, 13, 16
29 C.F.R. 1607.16(D).....	13
Questions and Answers on the Federal Executive Agency Guidelines on Employee Selection Procedures, 44 Fed. Reg. 11,996 (1979).....	28
 MISCELLANEOUS:	
David C. Baldus & James W. L. Cole, <i>Statistical Proof of Discrimination</i> § 7.1 (1987 Supp.)	30

INTEREST OF THE UNITED STATES

The United States has a substantial interest in this appeal, which involves the application of Title VII of the Civil Rights Act of 1964, 42 U.S.C. 2000e *et seq.*, to law enforcement promotion practices. The Attorney General enforces Title VII against public employers, 42 U.S.C. 2000e-5(f)(1), and the Equal Employment Opportunity Commission (EEOC) enforces the statute against private employers, 42 U.S.C. 2000e-5(a) and (f)(1). In addition, Title VII applies to the United States in its capacity as the Nation's largest employer. 42 U.S.C. 2000e-16.

The United States files this brief pursuant to Federal Rule of Appellate Procedure 29(a).

QUESTIONS PRESENTED

The United States will address the following questions:

1. Whether the district court erred in analyzing whether an employment exam used to select police sergeants in Boston was valid.
2. Whether the district court erred in assessing the statistical evidence presented by plaintiffs to establish a prima facie case of disparate impact in the non-Boston jurisdictions.

STATEMENT OF THE CASE

1. In 2007, black and Hispanic police officers sued their employers—Boston, several smaller municipalities, and the Massachusetts Bay Transportation

Authority (MBTA). They claimed that defendants discriminated against them on the basis of race or ethnicity in violation of the disparate impact provisions of Title VII of the Civil Rights Act of 1964, 42 U.S.C. 2000e-2(k)(1)(A), through their administration of promotional exams for the position of police sergeant. Record Appendix (RA) 89-90; *Lopez v. Massachusetts*, 588 F.3d 69, 72-73 (1st Cir. 2009).

Between 2005 and 2008, all defendants administered one or two police sergeant promotional examinations developed by Massachusetts' Human Resources Division (HRD).¹ *Lopez*, 588 F.3d at 77; RA97; see also RA109-115. In any given year, the same HRD test was used statewide, except for Boston's exams, which included Boston-specific questions. RA97, 121-122.

HRD's examinations consisted of two components: an 80-question, closed-book, written, multiple-choice, knowledge test, and an "education and experience" rating (E&E). RA97. The written test accounted for 80% of the final score; the E&E, which assigned point values to work and academic experience, accounted for the remaining 20%. RA97, 123. Based on a 100-point scale, 70 was the passing score on the written test. RA97.² Applicants were ranked according to their

¹ Plaintiffs also sued Massachusetts and HRD, but in a previous appeal, this Court dismissed the state defendants. *Lopez*, 588 F.3d at 90.

² The passing score on Boston's 2008 written test was 66. RA2695-2703.

combined score. *Ibid.* Nearly every defendant's practice was to make selections in strict rank order. RA98.

2. In the summer of 2010, the district court presided over an 18-day bench trial, focusing solely on liability. On September 5, 2014, the court issued its decision entering judgment in favor of all defendants. RA135.

a. The court first examined whether plaintiffs met their initial burden of proving the exams had a disparate impact on minority police officers. Boston conceded its 2005 and 2008 exams had such an impact. RA108. Accordingly, the court addressed whether plaintiffs had demonstrated a prima facie case against the other defendants. RA102-116.

The court began by reciting the EEOC's 4/5 "rule of thumb," which provides that "[a] selection rate for any race . . . which is less than four-fifths (4/5) (or eighty percent) of the rate for the group with the highest rate will generally be regarded . . . as evidence of disparate impact." RA102-103 (quoting 29 C.F.R. 1607.4(D)).³ The court recognized that adverse impact can be assessed at various points, including not only selection rates, but also passing scores, effective passing

³ Part 1607 of 29 C.F.R. sets forth the Uniform Guidelines on Employee Selection Procedures (Guidelines). These regulations were promulgated jointly by the EEOC, the Civil Service Commission, the Department of Labor, and the Department of Justice.

scores (the score at which, as a practical matter, one is eligible for promotion), or mean group scores (comparing how well groups perform as a whole). RA103.

To overcome the difficulty of small sample sizes for the non-Boston defendants, plaintiffs presented several types of aggregated data (differences in selection rates, passing rates, and mean test scores) both statewide and over time for individual defendants. RA4307-4310. In addition, plaintiffs presented disaggregated data for each jurisdiction that demonstrated particular exam years' adverse impacts for promotion rates, passing rates, effective passing rates, and mean scores. RA4311-4327.⁴

While acknowledging that “[a]ggregation * * * may be appropriate in some cases,” the court refused to allow plaintiffs to rely on aggregated data to establish their prima facie case. RA106. The court offered three reasons for rejecting data aggregated across jurisdictions: first, because each city could promote only its own officers to sergeant; second, because there might be differences among cities' candidate pools due to variations in original selection procedures and training; and third, because aggregation, in some circumstances, can produce statistical anomalies yielding false positives. RA106-107. As to aggregating data over time for particular jurisdictions, the court found the

⁴ Plaintiffs' aggregated statewide statistics excluded Boston.

technique unreliable because repeat test-takers could potentially distort the results. RA107-108.

Having rejected aggregated data, the court examined disaggregated statistics for the individual jurisdictions, but focused only on promotion-rate data. RA109-116.⁵ The court concluded that “[w]ithout aggregation, the plaintiffs’ statistical case against the jurisdictions other than Boston falls apart” (RA109), and ruled in favor of all non-Boston defendants (RA116).

b. The district court next turned to whether Boston had carried its burden of demonstrating that its exams in 2005 and 2008, notwithstanding their adverse impact, were “job related” and “consistent with business necessity” (RA116 (citing 42 U.S.C. 2000e-2(k)(1)(A)(i))), and therefore “valid” for the selection process, *ibid.* The Guidelines address the validity of selection instruments and recognize criterion, content, or construct validity. RA95 (citing 29 C.F.R. 1607.5(A)). Because Boston relied on the content-validity method to justify its test, it had the burden of producing “data showing that the content of the selection procedure is representative of important aspects of performance on the job for which the candidates are to be evaluated.” *Ibid.* (citing 29 C.F.R. 1607.5(B), 1607.14(C)).

⁵ The only exception is that the court examined passing rates for MBTA because the plaintiffs employed there had failed the exam. RA115.

The court found that, in preparing the 2005 and 2008 Boston exams, HRD relied heavily on a 1991 validation report for a sergeant exam conducted by its predecessor, the Department of Personnel Administration. RA118. The court cited a “testability analysis” included in the 1991 report that identified knowledge, skills, and abilities (KSAs) important to the police sergeant job that could be assessed under the written test or the E&E component. RA120. According to the court, that analysis showed “that more than half of the KSAs identified as pertinent to the job of sergeant were tested.” *Ibid.* (citing RA3927-3933, included in the Addendum to this brief as Attachment A). In the court’s view, “[t]his was sufficient to meet the ‘representative sample’ requirement of the Uniform Guidelines.” *Ibid.*

Defendants’ expert, Dr. Outtz, an industrial psychologist, conceded that the written job-knowledge test by itself was not valid “because it could not measure some skills and abilities (as distinguished from knowledge) essential to the position, such as leadership, decision-making, interpersonal relations, and the like.” RA123-124; see also RA2111-2112, 2185. According to the court, however, Dr. Outtz thought such skills and abilities were addressed by the E&E rating; therefore, the court found the exam as a whole to be “minimally valid,” satisfying Title VII standards. RA124.

Although the court recognized that “[t]he use of a ranking device requires a separate demonstration that there is a relationship between higher scores and better job performance” (RA95-96 (citing 29 C.F.R. 1607.14(C)(9))), the court did not address whether Boston’s exams were valid for rank-ordering. Instead, the court’s only analysis of rank-order selection came in the context of discussing alternative employment practices, when the court held that plaintiffs failed to carry their burden of demonstrating that “banding”—treating scores within a certain number of points the same for promotion purposes—was a realistically available alternative, given state-law restrictions. RA132-134. Likewise, the court did not address whether the passing score was valid.

c. Finally, the court ruled that plaintiffs failed to establish the existence of an available “alternative employment practice” that was equally valid and would have had a lower discriminatory impact. RA134-135.

SUMMARY OF ARGUMENT

1. a. Having conceded that its 2005 and 2008 police sergeant exams had an adverse impact on minorities, Boston was required to demonstrate that the exams were “job related” and “consistent with business necessity,” 42 U.S.C. 2000e-2(k)(1)(A)(i), and therefore “valid” for the selection process. The district court failed to analyze correctly whether Boston’s exams were valid because it treated the assessment of whether the exams were representative of job content as a mere

quantitative exercise of counting up how many knowledge, skills, and abilities required by the job supposedly were tested. But the representativeness requirement is a qualitative one that demands that the exam test the most important KSAs needed for the job and that it actually measure those it purports to test.

The court agreed that certain skills and abilities that the written, multiple-choice test did not assess, such as “leadership, decision-making, interpersonal relations, and the like,” were essential to the sergeant position. RA123-124. Indeed, the court accepted that the sergeant exam was “minimally valid” only because it understood defendants’ expert to testify that such essential skills and abilities were “attested to by the E&E component.” *Ibid.* But although an employer must demonstrate that its exam *actually tests* the skills and abilities required for successful job performance, 29 C.F.R. 1607.14(C)(4), 1607.15(C)(4), the court did not require such a showing of Boston. And it is difficult to see how the E&E, which is just a rating sheet indicating work experience, academic degrees, and courses taught, could measure the KSAs attributed to it. Furthermore, without a better understanding of which KSAs were actually tested by the E&E and of their relative importance, the court had no way to assess whether the weight assigned to the E&E matched its importance. The court’s failure to evaluate whether the E&E adequately measured the KSAs assigned to it—much less

essential skills and abilities—renders suspect its conclusion that the exam as a whole was representative and therefore valid.

b. Even assuming the exams were “minimally valid” in content, the court failed to evaluate whether they were *used* in a valid manner—here, as a rank-ordering device with a passing cutoff score. See *Boston Chapter, NAACP, Inc. v. Beecher*, 504 F.2d 1017, 1025 n.14 (1st Cir. 1974). A discriminatory passing score is impermissible unless the employer shows that it measures “the minimum qualifications necessary for successful performance” of the job. *Lanning v. Southeastern Pa. Transp. Auth.*, 181 F.3d 478, 489 (3d Cir. 1999); see also 29 C.F.R. 1607.5(H). Here, the court failed to require a showing that the passing score was anything more than an arbitrary cutoff. Similarly, although the court recognized that the use of a ranking device “requires a separate demonstration that there is a relationship between higher scores and better job performance” (RA95-96 (citing 29 C.F.R. 1607.14(C)(9))), the court never addressed whether Boston’s use of exam results to rank applicants was valid.

2. The district court ruled that plaintiffs failed to establish a *prima facie* case of disparate impact against the non-Boston jurisdictions. In reaching that conclusion, the court rejected plaintiffs’ statistics aggregating data both across jurisdictions and over time, even though the challenged exams offered by the smaller jurisdictions were identical.

a. This Court has recognized that “[t]he utility of statistical evidence ‘depends on all of the surrounding facts and circumstances.’” *EEOC v. Steamship Clerks Union*, 48 F.3d 594, 604-605 (1st Cir. 1995) (*Steamship*) (citation omitted). Here, the court engaged in a highly artificial analysis of whether plaintiffs established a prima facie case against the smaller jurisdictions that essentially ignored the fact that these same tests (albeit with Boston-specific questions) had an enormous, statistically significant disparate impact in Boston where sample sizes were large.

Moreover, while the court focused on perceived flaws in plaintiffs’ statewide promotion-rate data, plaintiffs presented *other* kinds of aggregated statistics (including differences in mean scores and passing rates), both statewide and over time for individual defendants, that demonstrated adverse impact in many instances. RA4307-4308, 4310. Plaintiffs also presented disaggregated data for each jurisdiction, in addition to promotion-rate data, that reflected a statistically significant adverse impact for certain cities in certain measures. RA4311, 4314-4327. Add to that the Boston adverse-impact evidence, and plaintiffs’ showing “sufficiently diminish[es] chance as a likely explanation” for discriminatory disparities. *Fudge v. City of Providence Fire Dep’t*, 766 F.2d 650, 658 (1st Cir. 1985).

b. The court's reasons for disallowing aggregation do not withstand scrutiny in any event. The Guidelines endorse aggregated statistics where sample sizes are small and a selection procedure has been "used in the same manner in similar circumstances elsewhere" and "over a longer period of time." 29 C.F.R. 1607.4(D). Here the *same* exams were used in the *same manner* in each jurisdiction. Speculative differences among cities in original selection procedures or in-service training provide no basis to reject plaintiffs' prima facie case.

The court's concern that combining different sample sizes and selection rates would create false positives, a phenomenon known as Simpson's Paradox, is equally inapposite. Defendants presented no proof that Simpson's Paradox was actually a factor in this case. The court's rejection of data aggregated over time because of the possible impact of repeat test-takers is similarly problematic where defendants made no showing that repeat test-takers skewed the results. See *Jones v. City of Boston*, 752 F.3d 38, 48 (1st Cir. 2014); *Steamship*, 48 F.3d at 604.

For these reasons, this Court should vacate and remand for the district court to reconsider, under correct legal principles, both the validity of the Boston exams and whether plaintiffs established a prima facie case against the non-Boston defendants.

ARGUMENT

I

THE DISTRICT COURT DID NOT PROPERLY ANALYZE WHETHER BOSTON'S EXAMINATIONS FOR POLICE SERGEANT, AS CONSTRUCTED AND USED, WERE VALID

Under Title VII, once a plaintiff shows that “a particular employment practice” has a disparate impact, the employer must show that the “challenged practice is job related * * * and consistent with business necessity.” 42 U.S.C. 2000e-2(k)(1)(A)(i). See *Ricci v. DeStefano*, 557 U.S. 557, 578 (2009).

Having conceded that its 2005 and 2008 exams had a disparate impact on minorities, Boston’s “major task” was “to show a substantial relationship between the test results and job performance.” *Boston Chapter, NAACP, Inc. v. Beecher*, 504 F.2d 1017, 1021 (1st Cir. 1974). Because of flaws in the district court’s understanding of the “representativeness” requirement and gaps in its analysis of the E&E component, its conclusion that the exams were content-valid lacks foundation. As well, the court failed to ask whether the exams, even if “minimally valid” in terms of content, were *used* in a valid manner—here, as a rank-ordering device with a passing cutoff score. Therefore, for either reason, the court lacked a sufficient basis to conclude that Boston’s administration of the sergeant examinations in 2005 and 2008 comported with Title VII.

A. *The District Court Failed To Analyze Correctly Whether Boston's Exams Were Content-Valid*

The court relied on the 1991 “testability analysis” in concluding that a representative sample of KSAs was measured by the exams. RA120; see Attachment A. The testability analysis reveals that, of the 156 KSAs required for the sergeant job, 60 were deemed testable through the “Written Test,” 24 as testable through the E&E component, and 72 as “Not Tested.” Attachment A. Notably, nothing in the testability analysis indicates which KSAs were *actually* tested; it merely classified whether KSAs *could* be tested by either the written test or the E&E, or were not testable by either device. The court’s conclusion that “more than half” of the pertinent KSAs were tested (RA120, 124) was therefore correct only if the E&E actually measured most of the 24 KSAs assigned to it.

But the court’s focus was misdirected. Under the Guidelines, evidence of an exam’s content validity should “consist of data showing that the content of the selection procedure is representative of important aspects of performance on the job for which the candidates are to be evaluated.” 29 C.F.R. 1607.5(B); see also 29 C.F.R. 1607.14(C)(1) and (4), 1607.15(C)(4), 1607.16(D). Yet the court treated the assessment of an exam’s representativeness as a mere quantitative exercise of counting up how many KSAs required by the job purportedly were tested. Instead, the analysis is a qualitative one necessitating that the exam test the most important KSAs needed for the job and that it actually measure those it purports to test.

Content validity demands “that an important and distinguishing attribute be tested in some manner to find the best qualified applicants.” *Firefighters Inst. for Racial Equal. v. City of St. Louis*, 549 F.2d 506, 512 (8th Cir. 1977); see also 29 C.F.R. 1607.14(C)(2) (work behaviors selected for measurement should be “critical work behavior(s) and/or important work behavior(s) constituting most of the job”). Furthermore, a content-valid exam must not only test required skills and aptitudes, “it must test them ‘in proportion to their relative importance on the job.’” *Guardians Ass’n of New York City Police Dep’t v. Civil Serv. Comm’n of City of New York*, 633 F.2d 232, 243-244 (2d Cir. 1980) (citation omitted); accord *Legault v. aRusso*, 842 F. Supp. 1479, 1488 (D.N.H. 1994); *Boston Chapter, NAACP, Inc. v. Beecher*, 371 F. Supp. 507, 515-516 (D. Mass.), aff’d, 504 F.2d 1017 (1st Cir. 1974). Representativeness is required “to prevent either the use of some minor aspect of the job as the basis for the selection procedure or the needless elimination of some significant part of the job’s requirements from the selection process entirely.” *Guardians Ass’n of New York City Police Dep’t v. Civil Serv. Comm’n of City of New York*, 630 F.2d 79, 99 (2d Cir. 1980) (*Guardians*); see also *United States v. City of New York*, 637 F. Supp. 2d 77, 118-119 (E.D.N.Y. 2009) (city failed to demonstrate extent to which “important abilities” of firefighter were tested on exams).

The court agreed that skills and abilities such as “leadership, decision-making, interpersonal relations, and the like” were essential to the sergeant position. RA123-124. Evidence at trial established what common sense teaches: that a police sergeant, a supervisor, needs not only to “know and understand relevant law” (RA118), but also to have strong leadership and supervisory abilities, “command presence,” interpersonal skills, and oral communications skills. *E.g.*, RA213, 669, 875-876, 1244-1245. Defendants maintained that the written knowledge test assessed supervisory ability, but there is an obvious difference between testing an applicant’s knowledge (or rote memorization) of supervisory practices according to sources on a reading list, and testing an applicant’s ability to put that knowledge into practice and actually supervise. This Court made this point in *Beecher* when it recognized, with respect to a written, multiple-choice, firefighter exam, that “there is a difference between memorizing * * * fire fighting terminology and being a good fire fighter. If the Boston Red Sox recruited players on the basis of their knowledge of baseball history and vocabulary, the team might acquire authorities like John Kieran but who could not bat, pitch or catch.” 504 F.2d at 1023; cf. *Boston Police Superior Officers Fed’n v. Civil Serv. Comm’n*, 35 Mass. App. Ct. 688, 695, 624 N.E.2d 617, 621 (1993) (Commission properly found that “the multiple choice and training and experience components

alone failed to constitute a fair test of supervisory skills and ability” for Boston police lieutenant).⁶

Defendants’ expert admitted that the written, multiple-choice component of the sergeant exam was not, standing alone, sufficiently representative of the job. RA1892, 2111-2112, 2185. Indeed, the court accepted that the exam was “minimally valid” only because it understood Dr. Outtz to testify that essential skills and abilities not assessed by the written test were “attested to by the E&E component.” RA123-124.⁷ But where a “selection procedure purports to measure a knowledge, skill, or ability,” it is “essential” that the employer produce “evidence that the selection procedure *measures* and is a representative sample of the knowledge, skill, or ability.” 29 C.F.R. 1607.15(C)(4) (emphasis added), 1607.14(C)(4). Thus, an employer must demonstrate that its exam *actually tests* the skills and abilities required for successful job performance, see, *e.g.*, *United States v. City of New York*, 731 F. Supp. 2d 291, 308 (E.D.N.Y. 2010); *Legault*,

⁶ In addition, according to the 1991 testability analysis, the “[a]bility to communicate orally and in writing,” among other important skills and abilities, was not tested. Attachment A (#107).

⁷ When asked directly whether the exams “test enough of the supervisory skills to meet validity requirements,” however, Dr. Outtz did not cite the E&E but answered only in numerical terms: “The examinations here test enough of the KSAs that were identified in the job analysis to satisfy the requirements for content validity.” RA1971; see also RA2056-2057.

842 F. Supp. at 1488, and “utilizes appropriate metrics to measure those skills,” *United States v. Massachusetts*, 781 F. Supp. 2d 1, 18 (D. Mass. 2011).

Yet the court required no such showing of Boston. Indeed, it is difficult to see how the E&E component, which is just a rating sheet that indicates work experience, academic degrees and certificates obtained, and courses taught (see RA2611-2619), could measure the 24 KSAs attributed to it. Such KSAs included, *inter alia*: “Knowledge of local ordinances” (#4); “Knowledge of proper behavior in the courtroom” (#29); “Knowledge of the limits of one’s own authority” (#56); “Skill in interviewing and interrogation” (#80); “Skill in perceiving and reacting to the needs of others” (#85); “Ability to write (e.g., to prepare reports, etc.)” (#87); and “Ability to interpret policy” (#116). Attachment A. Dr. Outtz contended the E&E could measure these things by virtue of the three years’ experience a police officer needs to take the sergeant exam. Thus, if “you’ve been a police officer for three years,” it was no “stretch to assume you have knowledge of local ordinances.” RA2151 (discussing #4). A similar conclusory assertion was offered for knowledge of courtroom behavior (#29): “police officers and police sergeants are going to be called, sooner or later, to come into court and work with a prosecutor to prosecute a case.” RA2155. He likewise explained (citing #56): “I think if you have been a police officer for three years, you are going to get an

understanding of the limits of your authority, * * * what you can and can't do, because your authority will be tested every day.” RA1903-1904.

With this kind of logic, Boston might as well have abandoned altogether the effort to test a representative sample of KSAs because one could just as easily assume that officers with 3+ years' experience have many or all KSAs marked as testable by the written test without testing for any of them. The court's failure to assess whether the E&E component actually measured the KSAs attributed to it—much less *essential* skills and abilities such as leadership, decision-making, and interpersonal relations—casts doubt on its conclusion that the exam as a whole was representative and therefore valid. RA120.⁸

⁸ A key factual issue the court did not evaluate was whether the E&E component was actually linked to the KSAs to be tested by it. See 29 C.F.R. 1607.14(C)(6). Dr. Outtz attempted to supply a linkage by citing a 1991 Validation Report attachment that asked for ratings of the usefulness of specified education degrees, certificates/licenses/trainings, and work experience. RA1890-1891 (citing RA3726-3730); see also RA3739-3746. But the attachment does not purport to link any of these achievements with the important KSAs supposedly testable by the E&E. RA3726-3730; see also RA3317-3322 (linking exam subject areas to 59 KSAs deemed testable by the written test without addressing the KSAs attributed to the E&E).

Likewise, the consulting firm's 2000 job analysis generated a list of 155 KSAs for police sergeants (RA4043, 4058-4065), but the court did not assess how the 2005 and 2008 test plans purported to use that job analysis to allocate KSAs between the multiple-choice test and the E&E, or evaluate the number or criticality of KSAs to be tested by either component. Instead, the court determined that the test plans could be traced to the 2000 job analysis, which could be traced to the

(continued...)

Because the court did not analyze whether HRD's examination actually tested enough critical skills and abilities to render it representative, the court also could not undertake the required analysis regarding whether the exam's components were correctly weighted. A proper validation analysis requires an employer to demonstrate that the relative weight of exam components correlates with the importance of the information being tested. *Legault*, 842 F. Supp. at 1488 (testing procedures must "accurately and reasonably test the skills identified in the job analysis in accordance with their relative importance"). Yet nothing in the court's opinion indicates that Boston made any such showing. To the contrary, the 20% weight accorded the E&E rating was apparently based on longstanding "standard practice" (RA3316), rather than any calibrated formula. Without a better understanding of which KSAs were actually tested by the E&E and their relative importance, the court had no way to assess whether the 20% weight was too much or too little.⁹

(...continued)

1991 job analysis, on which the court relied in concluding that a representative sample of KSAs was tested. RA120.

⁹ In fact, the 20% weight significantly overstates how much the E&E could separate candidates because all eligible candidates automatically received a minimum of 14 of 20 possible points on the E&E. RA304-307, 3297. Moreover, the evidence suggests that, as a practical matter, the E&E actually had a maximum impact of 2 or 3 points. RA1157-1158, 2179-2180. Given the court's understanding that the E&E actually tested for critical skills and abilities (RA123-

(continued...)

Because the court misunderstood content-validity requirements and left key gaps in its analysis of the E&E component, its validity assessment is fatally flawed and must be redone.

B. The District Court Failed To Consider Whether The Selection Of A Cutoff Score And Rank-Order Use Of Exam Results Were Valid

As this Court has recognized, even assuming the exams were “minimally valid” in content, “a test should be valid for the purpose for which it is used,” *Beecher*, 504 F.2d at 1025 n.14, and “should receive no more weight in the selection process than its validity warrants,” *id.* at 1026. The Guidelines provide that evidence of both the validity and utility of a selection procedure should support the “operational use” to which the procedure is put. 29 C.F.R. 1607.5(G). Moreover, the exam must be used with “a scoring system that usefully selects from among the applicants those who can better perform the job.” *Guardians*, 630 F.2d at 95; accord *City of New York*, 731 F. Supp. 2d at 313-314.

1. As the Guidelines specify, “[w]here cutoff scores are used, they should normally be set so as to be reasonable and consistent with normal expectations of proficiency within the work force.” 29 C.F.R. 1607.5(H). A discriminatory cutoff

(...continued)

124), however, the E&E’s weighting and minute impact on scores are difficult to justify. The problem is exacerbated by the promotion of candidates in rank order. See 29 C.F.R. 1607.14(C)(9) (where selection procedure is used to rank candidates, it “should measure those aspects of performance which *differentiate* among levels of job performance”) (emphasis added).

score is impermissible unless the employer shows that it “measures the minimum qualifications necessary for successful performance of the job in question.”

Lanning v. Southeastern Pa. Transp. Auth., 181 F.3d 478, 489 (3d Cir. 1999); accord *Isabel v. City of Memphis*, 404 F.3d 404, 413-414 (6th Cir. 2005) (exam not valid where “cutoff score was nothing more than an arbitrary decision and did not measure minimal qualifications”); *Green v. Town of Hamden*, 73 F. Supp. 2d 192, 199 (D. Conn. 1999). Moreover, there “should generally be some independent basis for choosing the cutoff.” *Guardians*, 630 F.2d at 105.

Here, the district court never found that the passing point measured the minimum qualifications necessary to perform the job or was anything other than an arbitrary cutoff. Cf. *Beecher*, 504 F.2d at 1023 (passing score of 70 “selected arbitrarily”); *Bradley v. City of Lynn*, 443 F. Supp. 2d 145, 152 (D. Mass. 2006) (HRD used “passing point of seventy, an arbitrary number that has been used since at least 1971”). Boston argued below that the passing point was irrelevant because it seldom, if ever, reached candidates near the cutoff given its low selection rates. That argument necessarily assumes, however, that Boston justified its use of the exam as a rank-ordering device. See *Isabel*, 404 F.3d at 415 (rejecting city’s argument that none of plaintiffs would have ranked high enough even without a cutoff, because rank orderings were not valid). Yet, as discussed next, the court completely overlooked this aspect of the validity analysis as well.

2. To select among candidates based on rank-ordering, an employer must establish “*such substantial test validity* that it is reasonable to expect one- or two-point differences in scores to reflect differences in job performance.” *Guardians*, 630 F.2d at 100-101 (emphasis added); *Johnson v. City of Memphis*, 770 F.3d 464, 482 (6th Cir. 2014) (“[R]anking is a valid, job-related selection technique only where the test scores vary directly with job performance.”) (citation omitted); see also 29 C.F.R. 1607.14(C)(9). An employer cannot justify selecting candidates on the basis of ranked scores merely by showing that the examination has *some* level of validity. *Guardians*, 630 F.2d at 101. Instead, it faces the “substantial task” of (1) making “a substantial demonstration of job relatedness and representativeness to show a sound basis for making rank-ordering hiring decisions”; and (2) demonstrating “an adequate degree of reliability.” *Id.* at 103-104; accord *Johnson*, 770 F.3d at 482. Ultimately, a “high degree of correlation between examination score and job performance” is “needed to justify rank-ordering.” *City of New York*, 637 F. Supp. 2d at 131.

That “substantial demonstration” was never made here. Moreover, in 2008, HRD proposed to implement test bands as wide as seven points. See RA133-134; *Pratt v. Dietl*, No. 09-1254, at 3-4 (Mass. Super. Ct. Apr. 15, 2009) (Appellants’ Addendum 126-127). That effort represents a concession that one- or two-point differences in scores did not reflect differences in job performance.

The court's failure to address the exam's validity as a rank-ordering device is particularly significant given its finding that the exam was only "minimally" valid. RA124. Boston's attempt to justify rank-ordering by citing the exam's alleged "reliability" is beside the point. Reliability asks only whether an exam would produce consistent results if taken repeatedly. *Guardians*, 630 F.2d at 101. A test for how well sergeant candidates memorize a page of the phonebook might yield consistent scores and candidate rankings but would not be job-related.

While acknowledging that "[t]he use of a ranking device requires a separate demonstration that there is a relationship between higher scores and better job performance" (RA95-96), the court's only analysis relevant to rank-ordering was its discussion of whether one particular alternative—"banding"—was an available option that would reduce adverse impact. See RA132-134. But the court discussed banding only *after* finding the exam to be job-related and consistent with business necessity. At that point, the burden had shifted to plaintiffs to prove the availability of less discriminatory alternatives. RA124-125. Before it ever reached that point, however, the court was required to examine whether Boston had met *its* burden to establish a valid basis for rank-ordering. By failing to impose the appropriate burden on Boston, the court's analysis, again, was faulty.

Nor is it relevant, as the court believed, whether banding was permissible under Massachusetts law. RA133-134. The court was required to examine

whether the rank-ordering of candidates was valid under Title VII, which supersedes state law. 42 U.S.C. 2000e-7; see, e.g., *Bridgeport Guardians, Inc. v. City of Bridgeport*, 933 F.2d 1140, 1148 (2d Cir. 1991); *Guardians*, 630 F.2d at 104.

For the above reasons, this Court should vacate and remand to permit the district court to reevaluate the validity of the Boston exams under correct legal principles.

II

THE COURT DID NOT CORRECTLY ANALYZE WHETHER PLAINTIFFS ESTABLISHED A PRIMA FACIE CASE OF DISPARATE IMPACT AGAINST THE NON-BOSTON DEFENDANTS

The district court ruled that plaintiffs failed to establish a prima facie case of disparate impact against the non-Boston jurisdictions. In reaching that conclusion, the court held it inappropriate for plaintiffs to aggregate data either across jurisdictions or over time, even though the challenged exams in these cities and MBTA were identical. It focused solely on promotion-rate data while ignoring other data plaintiffs had presented suggesting the exams had an adverse impact on minorities. The court's analysis of these issues was deficient, necessitating a remand.

1. In evaluating whether plaintiffs established a prima facie case against the smaller jurisdictions, it is important not to lose sight of the forest for the trees.

This Court has recognized that “[t]he utility of statistical evidence ‘depends on *all* of the surrounding facts and circumstances.’” *EEOC v. Steamship Clerks Union*, 48 F.3d 594, 604-605 (1st Cir. 1995) (*Steamship*) (emphasis added) (quoting *International Bhd. of Teamsters v. United States*, 431 U.S. 324, 340 (1977)). Here, one thing we know about these tests is that they had an overwhelming and statistically significant disparate impact in Boston (to be sure, with the inclusion of Boston-specific questions) where sample sizes were large. See RA4307-4313. In the face of that undeniable impact, the court’s evaluation of whether plaintiffs established a prima facie case was highly artificial, treating in a vacuum the question whether the exams had a disparate impact in particular jurisdictions.

In addressing whether plaintiffs could rely on aggregated data for the smaller jurisdictions, the court focused solely on promotion data. But plaintiffs presented *several* kinds of aggregated statistics on adverse impact: across jurisdictions, plaintiffs presented statewide differences in mean test-scores, passing rates, and promotion rates by exam year, with promotion-rate data also aggregated statewide over time. RA4307-4309, 4315, 4319. Over time, using data from exams given in 2003 to 2008, plaintiffs presented differences in mean scores, passing rates, and promotion rates by jurisdiction. RA4310. Taken together, the aggregated evidence suggests that the challenged exams had a statistically significant disparate

impact on minority test-takers in the smaller jurisdictions, just as they undeniably did in Boston.¹⁰

Perhaps most striking were the statewide results for mean-score differences. Plaintiffs' expert, Dr. Wiesen, found "highly statistically significant" score disparities in every year between minority and non-minority test-takers taking the same exams (RA338), with a probability that the differences occurred by chance less than .001, or 1 in 1000, in nearly every year. RA4307, 4315, 4319; see also RA338-341, 358, 1534-1536. These results strongly supported plaintiffs' prima facie case. See, e.g., *Isabel v. City of Memphis*, 404 F.3d 404, 409-413 (6th Cir. 2005) (relying on mean-score differences to find an adverse impact); *Bridgeport Guardians, Inc. v. City of Bridgeport*, 933 F.2d 1140, 1147 (2d Cir. 1991) (comparing white and minority test scores); cf. *Fudge v. City of Providence Fire Dep't*, 766 F.2d 650, 659 (1st Cir. 1985) (Breyer, J., concurring) (noting absence of evidence that blacks had "lower *mean* score" than whites).

Plaintiffs also examined each jurisdiction separately. RA4311-4327. There were virtually no minority promotions to sergeant in the non-Boston jurisdictions.

¹⁰ For example, the 2005 statewide exam revealed a statistically significant disparity between minority and non-minority test-takers in passing rates. RA4308. Statewide exams in every year from 2005 to 2008 demonstrated statistically significant mean-score differences. RA4307, 4315, 4319. Jurisdiction-specific data, aggregated over years, yielded statistically significant disparities in certain cities for promotion rates, passing rates, or mean scores. RA4310.

Ibid. Even though sample sizes were too small to reveal a statistically significant adverse impact for promotions, there were several statistically significant disparities with respect to passing rates, effective passing rates, or mean scores. *E.g.*, RA4311. The court provided no reason for disregarding this data, which it had agreed was relevant. RA103.

Add the Boston adverse-impact evidence to the other data the court overlooked, and it becomes reasonable to infer that differences in promotion rates were in fact the product of an exam with a well-known risk of disparate impact (RA102), rather than “attribut[able] to random chance.” *Jones v. City of Boston*, 752 F.3d 38, 43 (1st Cir. 2014).

2. Nevertheless, focusing on plaintiffs’ promotion data, the court decided that aggregation across jurisdictions would be inappropriate because of possible differences in the original selection and subsequent training of each city’s candidates. RA106. This reasoning, however, runs counter to the Guidelines and would foreclose aggregation in many instances involving the same selection procedure, making it difficult to hold smaller municipalities to Title VII’s standards.

To be sure, the Guidelines caution that “greater differences in selection rates may not constitute adverse impact where the differences are based on small numbers and are not statistically significant.” 29 C.F.R. 1607.4(D). But where

evidence of adverse impact is “based upon numbers which are too small to be reliable,” the Guidelines provide that “evidence concerning the impact of the procedure over a longer period of time and/or evidence concerning the impact which the selection procedure had when used in the same manner in similar circumstances elsewhere may be considered in determining adverse impact.” *Ibid.*; see also Questions and Answers on the Federal Executive Agency Guidelines on Employee Selection Procedures, 44 Fed. Reg. 11,996, 12,000 (1979) (Question 27) (if test is “administered and used in the same fashion for a variety of jobs, the impact of that test can be assessed in the aggregate”).

Here, the *same* tests, based on the *same* job analyses, were used in the *same manner* in each jurisdiction, all of which are located in and subject to the laws of the same State. Courts have upheld aggregation of data even where exams were not identical. See, e.g., *Paige v. California*, 291 F.3d 1141, 1148 (9th Cir. 2002) (upholding aggregation of different supervisory exams where positions were sufficiently similar). Nor is it unusual to aggregate data across cities making their own hiring decisions using their State’s civil service exam. See, e.g., *Vulcan Pioneers, Inc. v. New Jersey Dep’t of Civil Serv.*, 625 F. Supp. 527, 534-535, 544 (D.N.J. 1985) (“extreme similarity” of tests rendered aggregation across municipalities “entirely appropriate”), *aff’d*, 832 F.2d 811 (3d Cir. 1987).

The court's response that there could be differences in original selection procedures or in-service training is a *non sequitur*. Any such differences should accrue equally to all candidates in the jurisdiction—minority or non-minority. Such hypothetical variations cannot be a reason to reject aggregation.

The court's reliance on Simpson's Paradox, where combining different sample sizes and selection rates can lead to false-positive results (RA107), was equally inapposite. While more conservative methods are available to analyze confounding effects in aggregated data,¹¹ Dr. Wiesen testified that he found no indication that any municipalities were skewing the results; such distortion was improbable, he explained, where the adverse-impact ratio for promotions in most cities was zero. RA740-747; see RA4311, 4858-4859. Defendants introduced an exhibit showing the operation of Simpson's Paradox using two hypothetical jurisdictions (RA4860) but presented no proof that the paradox had actually affected plaintiffs' statewide promotion data. RA2493. Indeed, defendants' expert Dr. Silva conceded it was appropriate to combine data across jurisdictions. RA2495. This speculative flaw in one subset of plaintiffs' data supplied no basis for the court categorically to reject aggregation of data across jurisdictions. See *Guardians Ass'n of New York City Police Dep't v. Civil Serv. Comm'n of City of*

¹¹ For example, the Mantel-Haenszel test allows statisticians "to investigate the consistency of data trends over time while avoiding errors due to aggregation." *Bazile v. City of Houston*, 858 F. Supp. 2d 718, 741 (S.D. Tex. 2012).

New York, 630 F.2d 79, 88 n.7 (2d Cir. 1980) (to accept “unsupported possibilities” of misleading statistics would “create an onerous burden of proof” in excess of Title VII standards).

The court’s rejection of plaintiffs’ data aggregated over time (see RA4310) is similarly problematic. The concern the court articulated about aggregating data over exam years—the possible effect of repeat test-takers (RA107-108)—is insufficient, standing alone, to disallow this data. As noted, the Guidelines specifically contemplate aggregation “over a longer period of time.” 29 C.F.R. 1607.4(D). And aggregation of exams over years is routinely accepted when the selection procedure has not changed. See, e.g., *Paige*, 291 F.3d at 1149 (pre-liability data may be included “if promotional practices remain similar over a longer period of time”); *Bouman v. Block*, 940 F.2d 1211, 1226 (9th Cir. 1991) (aggregating results of two years’ examinations to observe “trends from past examinations to see if the total pass rate showed evidence of discrimination”). “[W]hen a company’s hiring policies have remained unchanged over a period of time and there have been no substantial changes in the conditions determining their application, it would be unreasonable to require a plaintiff to break his or her data into year-by-year subgroups.” David C. Baldus & James W. L. Cole, *Statistical Proof of Discrimination* § 7.1, at 100-101 (1987 Supp.).

To be sure, more conservative techniques evaluating the effect of repeat test-takers were available had defendants made some showing that the results were distorted by their presence. Similarly, had defendants made some showing that participation by repeat test-takers had a “race” effect such that it altered disparities between minorities and non-minorities, plaintiffs would have been required to respond. But defendants made no such showings, and the court was not free to adopt unsupported speculation. This Court rejected the same argument in *Jones*, involving repeated drug testing of some individuals, in part because the defendant “offer[ed] no analysis of the actual magnitude and effect of the claimed lack of independence in year-to-year results.” 752 F.3d at 48. “[A] defendant who asserts that a plaintiff’s prima facie case is insufficient must point out real deficiencies, not simply hurl epithets from behind gauzy generalizations.” *Steamship*, 48 F.3d at 604.

For these reasons, this Court should vacate and remand for the district court to reconsider whether plaintiffs have established a prima facie case of disparate impact against the non-Boston defendants.

CONCLUSION

This Court should vacate the judgment and remand for further proceedings.

Respectfully submitted,

P. DAVID LOPEZ
General Counsel

VANITA GUPTA
Acting Assistant Attorney General

CAROLYN L. WHEELER
Acting Associate General Counsel
Appellate Services

Equal Employment Opportunity
Commission
131 M Street N.E.
Washington, DC 20507
(202) 663-4739

s/ Bonnie I. Robin-Vergeer
SHARON M. MCGOWAN
BONNIE I. ROBIN-VERGEER
Attorneys
Department of Justice
Civil Rights Division
Appellate Section
Ben Franklin Station
P.O. Box 14403
Washington, DC 20044-4403
(202) 353-2464

CERTIFICATE OF COMPLIANCE

1. This brief complies with the type-volume limitations of Federal Rule of Appellate Procedure 32(a)(7)(B), because it contains 6,997 words, excluding the parts of the brief exempted by Federal Rule of Appellate Procedure 32(a)(7)(B)(iii).

2. This brief complies with the typeface requirements of Federal Rule of Appellate Procedure 32(a)(5), and the type style requirements of Federal Rule of Appellate Procedure 32(a)(6), because it has been prepared in a proportionally spaced typeface using Word 2007 in 14-point Times New Roman font.

s/ Bonnie I. Robin-Vergeer
BONNIE I. ROBIN-VERGEER
Attorney

Date: March 6, 2015

CERTIFICATE OF SERVICE

I hereby certify that on March 6, 2015, I electronically filed the foregoing
BRIEF FOR THE UNITED STATES AS AMICUS CURIAE
SUPPORTING PLAINTIFFS-APPELLANTS AND URGING VACATUR with
the Clerk of the Court for the United States Court of Appeals for the First Circuit
by using the appellate CM/ECF system.

I further certify that all participants in this case are registered CM/ECF users
and that service will be accomplished by the appellate CM/ECF system.

s/ Bonnie I. Robin-Vergeer
BONNIE I. ROBIN-VERGEER
Attorney

ADDENDUM

TABLE OF CONTENTS

	PAGE
Attachment A: Testability Analysis, Appendix to 1991 Validation Report, Attachment EE (Record Appendix 3927-3933) ¹	1

¹ Record Appendix pages 3929 and 3930 are out of order and have been switched in Attachment A.

ATTACHMENT A

TESTABILITY ANALYSIS

Page 1

	WRITTEN TEST	EDUCATION & EXPERIENCE	NOT TESTED
KNOWLEDGES, SKILLS, ABILITIES AND PERSONAL CHARACTERISTICS REQUIRED			
Required Upon Promotion To Sergeant and Higher Level Titles			
1. Knowledge of the Massachusetts General Laws pertaining to public safety.	X		
3. Knowledge of constitutional law as it relates to the police function (i.e., 1st, 4th, 5th amendments).	X		
4. Knowledge of local ordinances.		X	
5. Knowledge of what constitutes violations and crimes.	X		
6. Knowledge of what constitutes harassment by police.			X
7. Knowledge of state and local laws relating to youths and their offenses.	X		
8. Knowledge of the individual and constitutional rights of people.	X		
9. Knowledge of bail procedures.			X
10. Knowledge of how to look up sections of the law.		X	
11. Knowledge of penal law.			X
12. Knowledge of the Child Protective Act.			X
13. Knowledge of ABC law.			X
16. Knowledge of the federal, state and local regulations regarding the administration of intoxication tests.		X	
17. Knowledge of department procedures regarding the checking of prisoners.		X	
18. Knowledge of department rules, regulations, policies and procedures regarding the processing and supervision of prisoners.		X	
19. Knowledge of department procedures regarding receipt of bail money.			X
20. Knowledge of all departmental rules, regulations, standards, policies and procedures.		X	
21. Knowledge of procedures/techniques involved in a preliminary criminal investigation (e.g., searching the crime scene, determining the exact nature of the crime, interviewing witnesses and victims, protecting physical evidence).	X		
22. Knowledge of procedures/techniques involved in a continuing criminal investigation (e.g., reviewing reports from preliminary investigation, developing a theory, pursuing leads).	X		
23. Knowledge of recording procedures/techniques at the crime scene (e.g., field notes, photographs, sketches, etc.).	X		
24. Knowledge of procedures/techniques involved in handling, protecting and collecting physical evidence at a crime scene.	X		
25. Knowledge of the informant system and how informants can be used to obtain information.	X		

TESTABILITY ANALYSIS (CONT.)

	WRITTEN TEST	EDUCATION & EXPERIENCE	NOT TESTED
KNOWLEDGES, SKILLS, ABILITIES AND PERSONAL CHARACTERISTICS REQUIRED			
Required Upon Promotion to Sergeant and Higher Level Titles			
26. Knowledge of specialized investigative procedures (e.g., surveillance and undercover operations).	X		
27. Knowledge of the procedures involved in confessions, admissions, and written statements.	X		
28. Knowledge of search and seizure procedures	X		
29. Knowledge of proper behavior in the courtroom (e.g., case preparation, appearance).		X	
30. Knowledge of the rules of evidence (e.g., types of evidence, rules of exclusions, corpus delicti).	X		
31. Knowledge of the various social classes and their impact on police work (e.g., upper class, lower-middle class, etc.).			X
32. Knowledge of urban and suburban population factors and how they affect police work.			X
33. Knowledge of town and city growth and how it affects police response.			X
34. Knowledge of road transportation as it relates to police operations (e.g., traffic control).	X		
36. Knowledge of the different kinds of citizen alienation within a community and how it affects the police.			X
37. Knowledge of the police patrol function (e.g., public service, criminal investigation, etc.).	X		
38. Knowledge of the police patrol goals and objectives (e.g., the importance of goals and objectives, difficulties associated with goals and objectives).	X		
39. Knowledge of external and internal factors affecting patrol operations (e.g., community influences, political influences, police organization and structure.)		X	
40. Knowledge of occupational hazards of police patrol (e.g., stress, style of management, role conflict, fatigue).	X		
41. Knowledge of various police patrol methods (e.g., automobile vs. foot patrol, motorcycle, etc.).	X		
46. Knowledge of the communication function as it relates to patrol operations (e.g., dispatch procedures, recording complaints, computer-aided dispatch, etc.).	X		
47. Knowledge of the police records system (e.g., public information, criminal prosecution, etc.)	X		
50. Knowledge of the various communities within the Department's jurisdiction and the factors which make them unique.		X	

TESTABILITY ANALYSIS (CONT.)

	WRITTEN TEST	EDUCATION & EXPERIENCE	NOT TESTED
KNOWLEDGES, SKILLS, ABILITIES AND PERSONAL CHARACTERISTICS REQUIRED			
Required Upon Promotion to Sergeant and Higher Level Titles			
52. Knowledge of criminal law.	X		
53. Knowledge of community emergency services.			X
54. Knowledge of the types and uses of various weapons.			X
56. Knowledge of the limits of one's own authority.		X	
57. Knowledge of court decisions affecting departmental operations (e.g., search and seizure, liability, detention and arrest, discipline, criminal law, etc., evidence, rules of exclusions, corpus delicti).	X		
58. Knowledge of administrative rules, regulations and procedures affecting the operations of the police department.		X	
59. Knowledge of procedures/techniques involved in homicide investigations.	X		
60. Knowledge of procedures/techniques involved in juvenile incidents.	X		
61. Knowledge of procedures/techniques involved in riot incidents.	X		
62. Knowledge of procedures/techniques involved in situations involving barricaded suspects and/or hostages.	X		
63. Knowledge of the procedures/techniques when a major disaster occurs (i.e., hurricane, floods, earthquake, tornado, etc.).		X	
64. Knowledge of the grievance process.	X		
65. Knowledge of personnel evaluation techniques.	X		
66. Knowledge of proper English grammar and spelling.			X
67. Knowledge of the principles and practices of management including planning, organizing, directing, motivating, controlling and decision making.	X		
70. Knowledge of the behavior and/or characteristics of individuals under the influence of narcotics.			X
71. Knowledge of the various types of narcotics and their effects on a user.	X		
72. Knowledge of gangs and gang violence.			X
77. Knowledge of the techniques of handling and transporting injured persons.		X	
78. Knowledge of the types and availability of public and private organizations for providing recreational services.		X	
79. Knowledge of the types and availability of public and private organizations for providing medical services.		X	
80. Skill in interviewing and interrogation.		X	

TESTABILITY ANALYSIS (CONT.)

	WRITTEN TEST	EDUCATION & EXPERIENCE	NOT TESTED
KNOWLEDGES, SKILLS, ABILITIES AND PERSONAL CHARACTERISTICS REQUIRED			
Required Upon Promotion To Sergeant and Higher Level Titles			
81. Skill in driving a patrol car.			X
83. Skill in use of firearms.			X
84. Skill in identifying problems, securing relevant information from both oral and written sources, identifying possible causes of problems, and analyzing and interpreting data and complex situations involving conflicting demands, needs, or priorities.	X		
85. Skill in perceiving and reacting to the needs of others.		X	
86. Ability to read (e.g., to review subordinates' written reports, department policies, etc.).			X
87. Ability to write (e.g., to prepare reports, etc.).		X	
88. Ability to make and carry out decisions quickly.			X
89. Ability to choose actions appropriate to the situation, not to over-react.	X		
90. Ability to confront problems, take charge, assume responsibility.	X		
91. Ability to apply appropriate sections of the law to each specific case.	X		
92. Ability to be confidential.		X	
93. Ability to be objective, impartial, avoid preconceived opinions.			X
94. Ability to follow policies and procedures.		X	
95. Ability to compute basic arithmetic.			X
96. Ability to motivate staff.	X		
97. Ability to supervise.	X		
98. Ability to give clear, concise verbal orders.			X
99. Ability to give constructive criticism.	X		
100. Ability to appropriately delegate assignments.	X		
101. Ability to define problems.			X
102. Ability to demonstrate administrative judgement.	X		
103. Ability to accept responsibility.		X	
104. Ability to be objective.			X
105. Ability to plan.	X		
106. Ability to coordinate.	X		
107. Ability to communicate orally and in writing.			X
108. Ability to develop alternative solutions to problems, to evaluate courses of action, and to reach logical decisions based on the information at hand.	X		
109. Ability to maximize human potential of subordinates through training and developmental activities.	X		
110. Ability to recognize and resolve performance deficiencies in subordinates.	X		

TESTABILITY ANALYSIS (CONT.)

	WRITTEN TEST	EDUCATION & EXPERIENCE	NOT TESTED
KNOWLEDGES, SKILLS, ABILITIES AND PERSONAL CHARACTERISTICS REQUIRED			
Required Upon Promotion To Sergeant and Higher Level Titles			
111. Ability to efficiently establish an appropriate course of action for self/others.	X		
112. Ability to accomplish a specific goal.			X
113. Ability to make proper assignment of personnel.	X		
114. Ability to approach work in a logical and organized manner.	X		
115. Ability to organize and present information in a convincing manner or to modify the approach to gain agreement or acceptance.			X
116. Ability to interpret policy.		X	
117. Ability to present a positive and confident impression in dealing with others.			X
118. Ability to communicate thoughts and ideas clearly to others in a confident manner.			X
119. Ability to bring calm to control surroundings when in stress producing situations.			X
120. Ability to apply high standards of work performance, accuracy and completeness of detail.			X
121. Ability to influence others towards goal attainment.	X		
122. Ability to understand the needs and motivation of others.	X		
123. Ability to deal effectively with the public.	X		
124. Ability to discipline subordinates.	X		
125. Ability to understand complex oral and written instructions.			X
126. Ability to plan, organize and control the use of manpower, equipment and budgetary resources to achieve department goals and objectives.	X		
127. Ability to develop new policies and procedures to achieve department goals and objectives.	X		
128. Ability to coordinate the efforts of others in accomplishing assigned work objectives.	X		
129. Ability to establish rapport with persons from different ethnic, cultural and/or economic backgrounds.			X
130. Ability to detect and identify drugs and alcohol by smell.			X
131. Ability to identify abnormal behavior and/or characteristics of individuals under the influence of narcotics.	X		
132. Ability to organize work by establishing reporting relationships and by assigning work accordingly.	X		
133. Ability to work independently.			X
135. Ability to read and interpret such documents as maps, charts, etc.	X		

TESTABILITY ANALYSIS (CONT.)

	WRITTEN TEST	EDUCATION & EXPERIENCE	NOT TESTED
KNOWLEDGES, SKILLS, ABILITIES AND PERSONAL CHARACTERISTICS REQUIRED			
Required Upon Promotion To Sergeant and Higher Level Titles			
136. Ability to work accurately with names, numbers, codes and/or symbols.		X	
138. Ability to understand spoken English words and sentences.		X	
139. Ability to understand written sentences and paragraphs.			X
140. Ability to use English words or sentences in speaking so others will understand.			X
141. Ability to use English words or sentences in writing so others will understand.			X
142. Ability to produce a number of ideas about a given topic.			X
143. Ability to come up with creative solutions to problems or to develop new procedures for situations where standard operating procedures do not apply.			X
144. Ability to remember information, such as words, numbers, pictures, and procedures.			X
145. Ability to tell when something is wrong or is likely to go wrong.	X		
148. Ability to apply general rules to specific problems to come up with logical answers.			X
149. Ability to combine separate pieces of information, or specific answers to problems, to form general rules or conclusions.	X		
150. Ability to correctly follow a rule or set of rules to arrange things or actions in a certain order.			X
152. Ability to combine and organize different pieces of information into one meaningful pattern quickly.			X
153. Ability to detect a known pattern that is hidden in other material.			X
154. Ability to tell where you are in relation to the location of some object or to tell where the object is in relation to you.			X
157. Ability to move controls of a machine or vehicle.			X
158. Ability to coordinate movements to two or more limbs (i.e., two arms, two legs, or one leg and one arm) together, such as in moving equipment controls.			X
159. Ability to choose between two or more movements quickly and accurately when two or more different signals are given.			X
160. Ability to adjust an equipment control in response to changes in the speed and/or direction of a continuously moving object or scene.			X

	WRITTEN TEST	EDUCATION & EXPERIENCE	NOT TESTED
KNOWLEDGES, SKILLS, ABILITIES AND PERSONAL CHARACTERISTICS REQUIRED			
Required Upon Promotion To Sergeant and Higher Level Titles			
161. Ability to give one fast response to one signal when it appears.			X
162. Ability to keep the hand and arm steady.			X
163. Ability to make skillful, coordinated movements of one hand, a hand together with its arm or two hands to grasp, place, move or assemble objects.			X
166. Ability to make a quick, single movement of the arms or legs.			X
167. Ability to concentrate on a task and not be distracted.			X
168. Ability to shift back and forth between two sources of information.			X
169. Ability to use muscle force in order to lift, push, pull, or carry objects.			X
170. Ability to use short bursts of muscle force to propel oneself or an object.			X
173. Ability to bend, stretch, twist, or reach out with the body, arms, or legs.			X
174. Ability to bend, stretch, twist, or reach out with the body, arms, and/or legs, both quickly and repeatedly.			X
175. Ability to coordinate the movement of the arms, legs, and torso together in activities where the whole body is in motion.			X
176. Ability to regain one's body balance, or to stay upright when in an unstable position.			X
177. Ability to exert oneself physically without getting out of breath.			X
178. Ability or capacity to see close environmental surroundings.			X
179. Ability or capacity to see distant environmental surroundings.			X
180. Ability or capacity to match or discriminate between colors.			X
181. Ability to see under low light conditions.			X
182. Ability to perceive objects or movement towards the edges of the visual field.			X
183. Ability to distinguish which of several objects is more distant from or nearer to the observer or to judge the distance of an object from the observer.			X
184. Ability to see objects in the presence of glare or bright ambient lighting.			X
185. Ability to focus on a single source of auditory information in the presence of other distracting and irrelevant auditory stimuli.			X
187. Ability to identify the direction from which an auditory stimulus originated relative to the observer.			X