

HARVARD MODEL CONGRESS 2010

During the hearings, the scores for individual applicants were not released, so no firefighter knew his personal score. Only the race and gender breakdown statistics were made available. The Civil Service Board split evenly, 2-2, denying certification of the test results and noting that even if the exam is job-related, as required by law, it could still have a disparate-impact on a minority group if there are equally valid, but less-discriminatory exams available.

The Plaintiffs in this case are 20 policemen – 18 white and one Hispanic – who scored well on the exam. The lead plaintiff is Sam B. Andrews. The Defendant is the City of Boston, represented by Mayor Thomas Meninger.

Complaint by the Plaintiff

The plaintiff brings this civil action in tort alleging a disparate intent under Title VII of the Civil Rights Act by the City of Boston in Fire Department promotion examinations administered in the spring of 2008.

Plaintiff in its complaint against defendant alleges: The Defendant acted with discriminatory intent when it threw out the results of a race-neutral examination that yielded unintended racial disparity. The Defendant's actions were based on the fact that it did not approve of the race of candidates who qualified for promotions.

The Defendant, as an employer, violated Title VII's prohibition "to adjust the scores of, use different cutoff scores for, or otherwise alter the results of, employment related tests on the basis of race" when it threw out the results 42 U.S.C. §2000e-2(1).

Plaintiff invokes the right to trial by jury in this case.

Answer of the Defendant

The Defendant denies the allegations and asserts that it did not act with discriminatory intent against the white candidates. Instead, it sought a race-neutral decision to deal with the obvious disparate impact on underrepresented candidates. There was no discriminatory intent because the decision treated every candidate the same: no one's test scores were considered for promotion, and everyone will have to have to take another exam. The City of Boston did not violate Title VII, but sought to comply with its requirements when it made the decision to not certify the results.

Relevant Sections of Title VII

Title VII of the Civil Rights Act of 1964, 42 U.S.C. §2000e-2

SEC. 2000e-2. [Section 703]

(a) Employer practices

It shall be an unlawful employment practice for an employer:

(1) to fail or refuse to hire or to discharge any individual, or otherwise to discriminate against any individual, with respect to his compensation, terms, conditions, or privileges of employment, because of such individual's race, color, religion, sex, or national origin; or

(2) to limit, segregate, or classify his employees or applicants for employment in any way which would deprive or tend to deprive any individual of employment opportunities, or otherwise adversely affect his status as an

HARVARD MODEL CONGRESS 2010

employee, because of such individual's race, color, religion, sex, or national origin.

(j) Preferential treatment not to be granted on account of existing number or percentage imbalance. Nothing contained in this subchapter shall be interpreted to require any employer, employment agency, labor organization, or joint labor management committee subject to this subchapter to grant preferential treatment to any individual or to any group because of the race, color, religion, sex, or national origin of such individual or group on account of an imbalance which may exist with respect to the total number or percentage of persons of any race, color, religion, sex, or national origin employed by any employer, referred or classified for employment by any employment agency or labor organization, admitted to membership or classified by any labor organization, or admitted to, or employed in, any apprenticeship or other training program, in comparison with the total number or percentage of persons of such race, color, religion, sex, or national origin in any community, State, section, or other area, or in the available work force in any community, State, section, or other area.

(k) Burden of proof in disparate impact cases

(1) (A) An unlawful employment practice based on disparate impact is established under this subchapter only if:

(i) a complaining party demonstrates that a respondent uses a particular employment practice that causes a disparate impact on the basis of race, color, religion, sex, or national origin, and the respondent fails to demonstrate that the challenged practice is job related for the position in question and consistent with business necessity; or

(ii) the complaining party makes the demonstration described in subparagraph (C) with respect to an alternative employment practice, and the respondent refuses to adopt such alternative employment practice.

(B) (i) With respect to demonstrating that a particular employment practice causes a disparate impact as described in subparagraph (A)(i), the complaining party shall demonstrate that each particular challenged employment practice causes a disparate impact, except that if the complaining party can demonstrate to the court that the elements of a respondent's decision-making process are not capable of separation for analysis, the decision-making process may be analyzed as one employment practice; (ii) If the respondent demonstrates that a specific employment practice does not cause the disparate impact, the respondent shall not be required to demonstrate that such practice is required by business necessity.

(C) The demonstration referred to by subparagraph (A)(ii) shall be in accordance with the law as it existed on June 4, 1989, with respect to the concept of "alternative employment practice".

(D) A demonstration that an employment practice is required by business necessity may not be used as a defense against a claim of intentional discrimination under this subchapter."

Applicable Federal Case Law

Duke Power vs. Griggs (1971) is a foundational case for defining disparate impact and disparate treatment under Title VII. Prior to the Civil Rights Act of 1964, Duke Power Co. segregated its employees by ethnicity. Following the Act, it changed its policies to require a high school diploma

HARVARD MODEL CONGRESS 2010

and IQ test for positions outside of the labor department. This eliminated a large number of African American candidates from such positions.

Under Title VII, the *Duke Power* court found that if such tests disparately impact an ethnic minority, it must be demonstrated that the tests are reasonably job-related and using the test was a business necessity. This court also established that employer practices need not be overtly discriminatory, but could be facially neutral and applied in a discriminatory manner. The court thus reasoned that Congress intended Title VII to be used not only to deal with employer motivation, but consequences of employer practice.

In *Washington v. Davis* 426 US 229 (1976), two African Americans sued Washington, D.C. after being turned down for the DC police department, claiming that the use of verbal test, on which African Americans scored significantly lower, had a racially disparate impact on them. The case did not involve Title VII, but rather the Equal Protection Clause of the Fourteenth Amendment and established that, "[an] official action will not be held unconstitutional solely because it results in a racially disproportionate impact." Instead, a plaintiff must prove discriminatory motive on the state actor's part. The court noted that "disproportionate impact is not irrelevant, but it is not the sole touchstone of an invidious racial discrimination forbidden by the Constitution."

In *McDonald v. Santa Fe Transportation Co.* 427 US 273 (1976), the Supreme Court held that Title VII prohibits racial discrimination against whites as well as blacks.

In *United Steel Workers of America v. Weber* 443 US 193 (1979), the Supreme Court held that private sector employers and unions may lawfully implement voluntary affirmative action plans to remedy past discrimination. The Court held that an employer and union do not violate a collectively bargained plan by reserving 50 percent of the slots in a training program in a traditionally segregated industry for black employees. The program is lawful because it does not "unnecessarily trammel the interests of white employees," does not "create an absolute bar to the advancement of white employees," and is "a temporary measure . . . not intended to maintain racial balance, but simply to eliminate a manifest racial imbalance."

In *Teal v. Connecticut* 1978 US 159 (1982), the Supreme Court held that an employer is liable for racial discrimination when any part of its selection process, such as an invalidated examination or test, has a disparate impact, even if the final result of the hiring process is racially balanced. The *Teal* decision means that fair treatment of a group is not a defense to an individual claim of discrimination.

In *City of Richmond v. Croson* 488 US 469 (1989), the city maintained a set-aside program for awarding municipal contracts to minority businesses. The court held "that the city has failed to demonstrate a compelling interest in apportioning public contracting opportunities on the basis of race. To accept Richmond's claim that past societal discrimination alone can serve as the basis for rigid racial preferences would be to open the door to competing claims for "remedial relief" for every disadvantaged group. The dream of a Nation of equal citizens in a society where race is irrelevant to personal opportunity and achievement would be lost in a mosaic of shifting preferences based on inherently immeasurable claims of past wrongs. [Citing *Regents of the University of California v. Bakke*]. Courts would be asked to evaluate the extent of the prejudice and consequent harm suffered by various minority groups. Those whose societal injury is thought to exceed some arbitrary level of tolerability then would be entitled to preferential classification. We think such a result would be contrary to both the letter and the spirit of a constitutional provision whose central command is equality."

In *Grutter v. Bollinger* 539 US 306 (2003), Barbara Grutter, a white ap-

HARVARD MODEL CONGRESS 2010

plicant to the University of Michigan Law School, filed suit under Title VII and the Equal Protection Clause, claiming that she was discriminated against for being white when the school rejected her. The Supreme Court found that "Equal Protection Clause does not prohibit the law school's narrowly tailored use of race in admissions decisions to further a compelling interest in obtaining the educational benefits that flow from a diverse student body," upholding the race-conscious admissions policies of the University of Michigan Law School. The quota system, however, was deemed unconstitutional in *Gratz v. Bollinger*, 539 US 244 (2003).

Post Summation Instructions to the Jury

This case deals with Title VII. Title VII, codified as 42 USC § 2000e, is part of the landmark Civil Rights Act of 1964 that has been the cornerstone of equal employment opportunity and has had far-reaching impact in forbidding discrimination in public facilities, the government; and employment since it was passed by Congress. It was originally phrased to protect historically disadvantaged groups, and was later amended to include women and whites. It created the Equal Employment Opportunity Commission (EEOC), whose guidelines are used in this case to help evaluate the adverse impact on African American candidates.

This civil case is unique in that the facts of the case are not the focus of the dispute. The question is whether or not the city acted in a discriminatory manner when it threw out the exam results. In considering the facts and arguments of the case, you may consider the distinction between disparate impact and disparate treatment. Both are requirements under Title VII – employers are allowed neither to engage in activities of disparate treatment, nor in activities that result in disparate impact. In this case, the city believed that the promotional examination had a *disparate impact* on African American applicants, as the racial disparity in the test results demonstrated. Therefore, the city took action to throw out the test results, eliminating the opportunity for the white applicants and a Hispanic applicant that scored high to be promoted. This, the plaintiffs argues, is *disparate treatment* based on race; the argument essentially claims that because a disproportionate number of white candidates were eligible, the city chose to throw out the results, which amounted to discriminatory treatment on the basis of race.

The defendant admits, as the plaintiff argues, that the action taken to refuse certification of the exam was motivated by concerns about the racial distribution of the test results. This is undisputed. The defendant admits that the City acted to avoid possible litigation by minority policemen under Title VII. If the exams had not yielded racially disparate results, the results would have been certified, and the high-scoring plaintiffs would have had the opportunity to be promoted. The court can infer that the action was motivated by a concern for the lack of minority applicants eligible for promotion, and therefore by a concern that there were too many white candidates.

The plaintiff's case rests on proving that there was discriminatory intent. They claim that the defendant used their desire to comply with Title VII's non-disparate-impact statute as a pretext to advance the interests of the African American candidates. Thus, the plaintiff claims, the defendant was actually violating Title VII's anti-disparate-treatment statute. The defendant claims that the test had a *disparate impact* on minority candidates, and refused certification. The plaintiff argues that throwing out the exam results amounted to *disparate treatment* of the white candidates. The court notes that this is the opposite scenario of usual Title VII violations. The law was phrased for the classic scenario where the defendant argues that test

HARVARD MODEL CONGRESS 2010

results are valid in order to justify their use. Ordinarily, as contemplated by the statute, the "complaining party" bears the burden of proving a disparate impact, and the "respondent," defense, bears the burden of "demonstrat[ing] that the challenged practice is job related for the position in question and consistent with business necessity," or, alternatively, the "complaining party" may prevail by showing that an alternative employment practice with less disparate impact existed and that the respondent failed to utilize it. See 42 USC § 2000e-2(k). In this case, the defendant has the burden of showing that there was disparate impact, which is usually the petitioner's burden.

As the jury, you must consider whether or not each side has met its burden. The plaintiff must prove that there was discriminatory intent against white candidates and that there was no disparate impact on minority candidates. The defense must, for their part, prove that there indeed was disparate impact; they do not have the burden to prove that they acted discriminately (that is the burden of the plaintiff).

Witness List

Plaintiff:

Sam B. Andrews
Alexander/a Stevens
Mario/a Lopez

Defense:

Mayor Thomas Meninger
Martin/a Robespierre
John/Johanna Malkin

Note on the Order of the Witnesses

The order of witnesses specified above is random and not binding. Teams should feel free to present witnesses in whatever order they deem most rewarding.

HARVARD MODEL CONGRESS 2010

Affidavit of Sam B. Andrews

Witness for the Plaintiff

My name is Sam B. Andrews, and I'm 33 years old. I've been a fire fighter in the Boston Fire Department for the last three years. Currently, I live in Dorchester with my wife and two kids. I became a firefighter fairly late, although both my father and grandfather were firefighters. I was determined to make my own way, and I didn't know whether choosing to follow my father's footsteps was just a default for me or whether I would be happy with it. After college, I experimented with a series of jobs that had me bouncing across the country; I was a bartender, bouncer, border patrol agent, paralegal, a miner in Alaska...you name it. Eventually, I had exhausted myself and realized that I just wanted to go home. It was such a strong urge that I knew my calling was unquestionable. I made my way back to Boston and went into training in late 2005.

At the end of 2007, there was all this hype around the station that they were going to be promoting soon, trying to fill the recently opened lieutenant and captain positions at the department. Sure enough, in November of 2007, they announced that promotional examinations would be administered on February 23rd, 2008. I was excited, but nervous. Imagine the SATs all over again, but even more stressful, because this exam literally establishes your career path. Luckily, I qualified. We needed a minimum of 30 months at the department, a high school diploma, and training courses that I had already taken. I studied like I had never studied before. Apparently, the examinations were going to be extraordinarily difficult. I bought all the test material I could afford and even hired a tutor. I have Attention Deficit Disorder (ADD), which I take medication for, but it makes it harder for me to focus and study. I wanted this so badly and I knew that it would take my life and the lives for my wife and kids in the right direction. I was willing to invest in order to make sure I scored well.

There were a few firefighters like me, who saw the exam as something that could move them forward. We arranged study groups that met twice a week in the months leading up to the exam and even met over Christmas. There's no such thing as vacation when your future is on the line. We posted signs to get as many people to join our groups as possible - the more brains we could pool the better. Apparently, the exams were being made professionally, and the company was supposed to survey firefighters from the New England area to see what sort of knowledge was crucial to know as a captain or as a lieutenant. Of course, this included method: how a firefighter was supposed to perform his job. In December, January, and February, I studied for five to eight hours a day for the exam. My buddy even quit his part-time job so that he could have more time to study. This was serious, and we all needed to get our heads in the game. I did notice that there were a few guys in the department who were planning on taking the exam that missed most of our study sessions, and I never saw them crack a book. That's why I thought it was so unfair when Mayor Meninger announced that he was throwing out the exam results. I couldn't understand. I studied so hard and other people didn't - it wasn't because the test was unfair, it was because they didn't study!

Let me lay out the timeline, though. We took the exam on February 23rd, 2008. The exam can't straightaway determine who gets promoted; the process is more complicated to make sure the best-qualified firefighters are chosen. The Boston City Fire Department is governed by a city charter that requires that the department follow the "rule of three." Basically, if you did the best, it doesn't mean that you automatically get the position. The administration is allowed to choose from the top three scorers for each position. Theoretically, then, even the person that scored best on the exam

HARVARD MODEL CONGRESS 2010

wouldn't necessarily be promoted (although I can't see this happening). The charter also requires that 60% of the score be based on the written examination, and 40% of the score be based on the oral examination. The results have to be certified by the Boston Civil Service Board (BCS) to make sure all guidelines were followed and that the exam had identified the most qualified candidates. The results went to the BCS in March of 2008. We never got our scores back before we were informed that, apparently, there was evidence of a racial disparate impact that was going to lead them to throw out the results. This was ridiculous - I didn't know everything that was going on at the administrative level, but I learned a lot over the course of April and May because I made sure to be at the BCS hearings.

At the hearings, they had testimony brought forth from so many different people, including citizens and ourselves. I was at least pleased that I had the opportunity to throw my two cents in. I didn't pay attention to all the statistics all that closely, but the most important result was that there would be no African Americans eligible for either promotion, and there would be one Hispanic eligible for lieutenantcy (but apparently, he scored third, which meant he was a shoe-in). The president of the company (PTI), which created and administered the exams, also came in to speak, which I believe made everything clear: the alleged disparity wasn't a result of a few kinks in a supposedly unfair exam, it was simply how the numbers played out. In all honesty, there were something like nine African Americans and seven Hispanics taking either exam *total*. Think about all the kids, any race, that didn't study in high school or college for exams - seven is nothing. Plus, there were way more white candidates that scored poorly and wouldn't have qualified either - it seemed proportional to me to the composition of the testing pool. Maybe the department should look at getting a more diverse group of firefighters in general instead of making us suffer like this.

There was all this talk about alternative testing. I completely agree that it's not desirable for there to be no Hispanics or African Americans promoted. I think it's incredibly unfortunate, but don't discriminate against the people that did study hard and are qualified because the city wants to fudge the numbers. There is no reasonable alternative testing. According to the chief and a number of city officials at the BCS hearings, preparing another exam would take years. Also, there were members of the department who couldn't afford the testing materials that, for example, I myself purchased. That, again, is unfortunate; however, this talk about developing a testing center to provide materials is ludicrous as a solution for our *current* situation. I think that's a brilliant idea to work on, but it's also going to take a few years and a lot of money to make it happen. Until then, who's going to fill those vacant spots?

Take it from a real firefighter (and some other firefighters also testified this at the BCS hearings), that test was fair. I felt that everything on there, I should have known. Of course, I could have messed up and probably did, but I prepared myself for those questions. A lot of them were straight out of the official guidebook that sits in our department's reading room. A lot of them were based on information and experience we would gain from being in the field. I don't think that the results had anything to do with race - I think it had to do with experience and time investment. Unfortunately, some of those kids just haven't been in the department long enough to have learned this information, and that means they're just not qualified. I attended every single study session, and I never noticed a disparity in the race of the people that attended - it was a really mixed bunch. Like I said, it's unfortunate that the numbers came out the way they did, and I support a test center for the future, but don't punish those of us who poured ourselves into this exam because you don't like the numbers. There will be opportuni-

HARVARD MODEL CONGRESS 2010

ties for promotions in the future, and hopefully by then, there will be more resources available, but that's just not possible now.

Sam B. Andrews

Sam B. Andrews

Subscribed and sworn to me on this, the 5th day of June, 2009.

Alexandra Makhholm, Notary Public

HARVARD MODEL CONGRESS 2010

Affidavit of Alexander/Alexandra Stevens

Witness for the Plaintiff

My name is Alexander/a Stevens, and I'm 52 years old. I live in Cincinnati, Ohio and currently am a director at Promotional Testing Industries (PTI). A director is responsible for an assigned project and leads a team in providing whatever consulting services are necessary. I was the director of the Boston Firefighter project that we were commissioned to do in 2007. I've been a director for the last five years, but have been part of PTI for the last 20 years. I probably have the most in-depth understanding of the examination at issue in this case.

PTI was chosen out of a large number of security sector consulting companies through the common bid process. Although we deal with the entire security sector in some capacity or another, we specialize in promotional testing for fire and police departments. The process of developing an exam is long and involved: how is a test supposed to measure how qualified a candidate is to keep you and your family safe from harm? PTI recognizes that the responsibility is enormous. We normally take about two months to process a standard job, during which time one team focuses on one project. We develop questionnaires based on information we gather from personal interviews with previous lieutenants and captains in the area and from observations gather from riding along with them into the field. These questionnaires about the job are administered to most of the incumbent lieutenants and captains in the department.

Any expert will tell you that there are racial disparities in testing. This has historically been the case, and no one has ever completely been able to get results with absolutely no apparent racial disparity; however, the data supports a generally accepted racial disparity. We follow the guidelines of the Equal Employment Opportunity Commission (EEOC), which does not allow more than a four-fifths difference. The EEOC "four-fifths rule" provides that a selection tool that yields,

"[a] selection rate for any race, sex, or ethnic group 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 by the Federal enforcement agencies as evidence of adverse impact, while a greater than four-fifths rate will generally not be regarded by Federal enforcement agencies as evidence of adverse impact." 29 CFR § 1607.4(D).

Because racial disparity is always an issue in testing, we took precautionary measures. We actually over-sampled minority candidates during the job analysis. It is well-established that potential disparity arises from the fact that different groups (whether they be different age groups, racial groups, or even income groups) can perform and explain their job differently. We wanted to make sure that no one was disadvantaged by having information gathered from predominantly one group.

Based on the information we gathered, the team developed a written examination to cover the material. For each test, we compiled a list of sources (training manuals, department procedures, and other materials). The sources were approved with minor changes by the chief and his assistant. We then used these sources to create the multiple-choice test. Pursuant to BCS guidelines, each test had 150 questions and was written at an 11th grade reading level. The candidates were then given until February, when the test was administered, to study. The department released the list of sources that the questions were drawn from and even noted the chapters upon which the test was heavily focused. The composite cutoff score was chosen to match up with

HARVARD MODEL CONGRESS 2010

minimal competence.

PTI also developed the oral examination, which composed 40% of the score. We developed case scenarios to test incident-command skills, fire-fighting tactics, interpersonal skills, leadership, and management ability. The oral exams were administered and scored by 40 assessors. The assessors at the time all held positions higher than lieutenant or captain, but none of them were from the department itself. The department encountered problems in previous examinations that made it too much of a liability to use people from within the department to assess candidates. Their credentials were submitted to the chief for approval. 66% of the assessors were minorities and composed two-thirds of the three-person panel. The day of the oral exam, the assessors were trained by PTI for 3 hours on how to score candidates. Candidates were allowed to sketch out a response on paper before articulating it so that they could collect their thoughts. It was a test of their practical abilities as firefighters and how they would perform in emergency situations. According to the contract we had with the Boston Fire Department, our team was prepared to issue a technical report that described our methodology in detail and analyzed the testing results. The city never requested this report. The mayor, however, did request to meet with me personally to address his concerns about the racial disparity he thought he saw. I reassured him that the numbers were consistent with previous statistics and were within a professionally acceptable range.

The BCS then held hearings to decide whether or not to certify the exams. I testified at these hearings and answered more questions, e-mails, and phone call interviews than I have done thus far in my career as a director. The BCS seemed to conduct a very extensive investigation. I disagreed completely with one piece of evidence that was frequently tossed around: that a simple numerical disparity could constitute racial disparity and justify throwing out the results. When the pool is so small, statistics are insignificant - you can't prove anything conclusively with only 80 candidates being analyzed; other factors must be considered, most importantly, whether or not there is an alternative examination that is proven to minimize racial disparity. There wasn't any evidence to show this. Unfortunately, years down the road, we may learn that there are better tests, but that's the way of progress. We're always getting better, and we can never be perfect now - we can only be the best that we can be, given the situation. The BCS brought in a competing company president to speak. It was a load of bull. The guy has had it in for my company for years now and is clearly just insinuating how he should do the job in the future. I had wished that the BCS would see through that, but apparently they didn't. It seemed like they *wanted* to throw out the exams and were spending so much time and energy to justify doing so.

The lead plaintiff was a top scorer and was nearly guaranteed a promotion. Candidates of all races passed the exam. The apparent racial disparity was comparable to that of other promotional exams. Also, there were only a few vacancies in the first place (six spots for lieutenant and nine spots for captaincy). The racial make-up of the candidates broke down typically (in parentheses, I calculated the pass rate for each group. The first number refers to how many candidates within the group achieved the minimal score to pass):

For Captaincy (4 spots; 7 eligible):
19 (64%) of 30 whites
3 (39%) of 7 African Americans*
3 (38%) of 8 Hispanics*

HARVARD MODEL CONGRESS 2010

**NB – Those who passed did not score high enough to be eligible for a promotion.*

For Lieutenant (6 spots; 9 eligible):
14 (58%) of 25 whites
2 (32%) of 5 African Americans
1 (20%) of 5 Hispanics*

**NB – Only member of underrepresented group that scored high enough to be eligible for a promotion.*

These numbers can obviously be studied in great detail. Take, for example, the captaincy numbers. Roughly 16% of the applicants were African American. Therefore, one would expect no more than one or two of the captaincy openings to be filled by an African American under the best conditions if everyone studied equally hard. The Boston Fire Department is composed of 25% African Americans, 20% Hispanics, and the remainder are non-minorities. None qualified for either position in this case. Only one Hispanic was eligible for a captaincy promotion, and he did so well that it was basically guaranteed. Many departments across the country have their tests checked by people within the department in order to make sure that the material is a fair and an accurate test of a candidate's abilities. However, Boston specifically requested that PTI not do this because there had been allegations of cheating on previous examinations. There is a higher likelihood of cheating if people within the department have already seen the exam. PTI did not believe it was necessary to have the examinations reviewed by outside experts; it's not standard procedure.

Alexander/Alexandra Stevens

Alexander/Alexandra Stevens

Subscribed and sworn to me on this, the 5th day of June, 2009.

Alexandra Makhholm, Notary Public

HARVARD MODEL CONGRESS 2010

Affidavit of Mario/a Lopez

Witness for the Plaintiff

My name is Mario/a Lopez, and I'm 65 years old. I currently live in Milton, Massachusetts with my spouse. I've been a psychology professor at Boston University for the last 5 years, conducting research on the influence of race and culture on performance tests. I trained as a psychologist at the University of Michigan, but applied my degree and training for the federal government. I worked in the Department of Defense as an adviser on employee assessments and as a psychologist from 1995 to 2004. As you can imagine, our government is very particular when it comes to its employees and wants to make sure that the most qualified candidates are where they are most needed. My area of expertise has never focused on firefighters, per se.

I was asked by the BCS to shed some light on what may be going on in this issue. I didn't review the examinations in length, although BCS offered me that option. The percentages presented to me are not that inconsistent with those on promotional exams nationwide. Looking at particular questions would not be an efficient use of my time, nor, given the numbers, is there any useful information I could gather from examining a technical test I wouldn't know the first thing about. It is sufficient for me to understand the methodology and manner of administration, both of which check out. PTI not only followed standard procedure, but also took precautionary measures to minimize racial disparity. I have to say that their services went above and beyond what was required of them. Their job analysis was incredibly extensive, and the assessors they gathered for the oral examination were very qualified. In every step, the Fire Chief was involved and gave his approval the entire way. PTI is really at the forefront of promotional testing.

Nonetheless, there are several factors that could account for statistical disparity. Although PTI aimed to have a proportionally high number of minority firefighters answering their questionnaires, 70% that were returned were completed by white male firefighters. Most of the literature on firefighters is very clear on the fact that different groups perform their job differently. Therefore, there is a possibility that if the questions were very specific to certain areas and based on techniques or knowledge that is particular to one group, the underrepresented groups could have suffered disparately. However, to address the question of alternative testing, no test that could be administered would completely get rid of racial disparity in the results.

I spoke with a former fire chief, Vincent Simmons, who did look into the exams in great detail and asked him for his opinion. According to him, the knowledge required for the exams was extensive; in reality, candidates had to know the "entire book." Simmons took the same view I did. Statistical disparity will exist, just as a margin of error will always exist, so we employ confidence intervals to analyze data. Although it may not have been desirable for the city, the statistics were well within an expected standard deviation from what would presumably be the true value. Qualifications are hard to measure, and humans are the trickiest of all things to measure. To say that race is the only factor causing the disparity is a conclusion I cannot reach with any degree of certainty.

However, a representative racial distribution in the leadership is crucial to the proper functioning of a firefighting force. There is a balance between keeping the integrity of the test and getting the results you want. Sometimes numbers just play out the way they do. The city could have decided to change the number of openings, or not to follow the "rule of three" and move to a "rule of five," but they chose not to. They also had the option of weighing the two parts of the test differently. A 30/70 breakdown rather

HARVARD MODEL CONGRESS 2010

than a 40/60 one would have made more underrepresented candidates eligible for promotions. This would have allowed the most qualified candidates to pass on without throwing out the results. It would have also made the promotion pool more diverse. It seems severely unfair to have thrown out a test that was prepared as properly as it could have been given the circumstances, but that produced results the city didn't want because it feared being sued under Title VII.

Mario/a Lopez

Mario/a Lopez

Subscribed and sworn to me on this, the 5th day of June, 2009.

Alexandra Makhholm, Notary Public

HARVARD MODEL CONGRESS 2010

Affidavit of Thomas Meninger

Witness for the Defense

My name is Thomas Meninger, and I am 65 years old. I am the current mayor of Boston and have been for the last 20 years. I live in Jamaica Plain with my wife and son. I'm representing the actions of the City of Boston in this case, although I didn't make the decision to not certify the 2008 examinations of the Boston Fire Department by myself. In fact, it was completely up to the Boston Civil Service Board; however, I did support their decision 100%.

Before I jump into justifying the Board's decision, let me point out that we didn't discriminate against the white candidates when we threw out the results. That's absurd! *Everyone* was treated exactly the same - we didn't raise the scores of underrepresented groups or fudge the percentages to get numbers we liked. In the face of strong evidence, we believed that the city could do better without harming those who had already done well. If the candidates who brought this issue to court are truly qualified, then they will do well on the next exam. I understand how frustrating it must be to have studied and then to realize you'll have to take the exam at some future point, and I am truly sorry for this, but they'll be prepared that much more for the next exam. The aim behind our decision was to disadvantage *no one*.

Over the course of six BCS hearings, it became clear to us that there were significant flaws in the exams, that the most qualified candidates had not been identified, and that it may have been possible to create a less discriminatory alternative test. One thing you'll notice is that no one is denying the fact that the results show a severe racial disparity. People involved in the testing business all testified to this at the hearings.

We ultimately decided that we would administer a new, less discriminatory test in the future. We considered other options, of course. We decided that we would not undertake affirmative action policies, proportional promotions, or adjust test scores to benefit underrepresented candidates. These were all methods employed in the past and still used in various other cities. The courts, however, have ruled out quota systems, and we wanted to make sure that the most qualified candidates were promoted. Our concern was not with qualifying underrepresented candidates, but rather making sure that there were underrepresented firefighters who were certainly just as qualified as white candidates, but we were simply not able to demonstrate this on paper given PTI's exams. The number one problem we identified was that candidates who could not afford access to materials were at a severe disadvantage because the test was so heavily source-based. It is true that prior promotional results revealed disparities that were even worse than those evidenced here. However, in those exams the underrepresented candidates had scored high enough to be promoted. In this case, only one underrepresented firefighter was eligible for promotion. It wasn't even guaranteed, although it was likely, that he would gain that promotion. It is my understanding that that particular firefighter supports Andrews in this case.

The composite score was composed of 60% written and 40% oral. There have been arguments made that a 30/70 breakdown would have shifted the numbers to allow underrepresented candidates to have a shot at promotions. First of all, the 40/60 breakdown is not arbitrary. It is based on a union contract with Boston Firefighters that outlines that a 40/60 breakdown is the best balance of required job knowledge and situational skills. We can't arbitrarily change it to 30/70 just so the numbers are more appealing. Also, no one was able to offer a convincing argument of how 30/70 wasn't an arbitrary breakdown and was actually a more appropriate balance to measure the most qualified candidates.

HARVARD MODEL CONGRESS 2010

There are also arguments to forget the "rule of three" that has directed promotions for the last decade. It was initially written into the city charter in order to prevent "banding," which rounded scores up to the nearest whole number and treated all whole number candidates equally. I can see why this can appeal to some people: judging whether or not a person is eligible to save lives based on a few decimal points can seem absurd. However, given the state of the system at the time, it allowed the city to promote the people it wanted because it seemed numerically like they qualified. Also, the city would have violated Title VII if it had rounded up the numbers to make underrepresented candidate scores seem higher.

Everything considered, the city saw that there was only one alternative to deal with the obvious disparate impact: to decline to use the results and administer a better test. We had an expert from another company come to the BCS hearing, who told us how we could have made it better. I'm sorry to the candidates and families who are frustrated by this process, but we ask them to understand that we were between a rock and a hard place. This isn't a case of reverse discrimination; this is a case of making sure that the city did everything in its power to identify the highest quality candidates.

Thomas Meninger

Thomas Meninger

Subscribed and sworn to me on this, the 5th day of June, 2009.

Alexandra Makhholm, Notary Public

HARVARD MODEL CONGRESS 2010

Affidavit of Martin/a Robespierre

Witness for the Defense

My name is Martin/a Robespierre, and I am the president of Identification Industries (ID) located in Concord, New Hampshire. By training I'm an industrial/organizational psychologist. I created this consulting company in 2000, when I was incredibly dissatisfied with state of the promotional testing industry. Most companies were not keeping up with testing methodology.

I was asked to testify at the BCS hearings in 2008 regarding potential racial disparity in the testing results of the Boston Fire Department promotions. ID was second in line to create these tests so we had already drawn up proposals on how we would do so. I was able to testify at two of the six hearings, although I was available to answer questions at the other four. I looked extensively into the methodology and exam questions themselves, and I believe that they unduly favored the overrepresented candidates.

Although I wasn't given the job analysis data from PTI, I can assure you that there is a severely adverse impact on underrepresented groups evidenced by the results. These could not be ignored, especially when I could identify what PTI could have done differently. When I first looked at the numbers, I was fairly surprised by the extent of the disparity. Normally, some sort of disparity exists on standardized tests throughout the country, so I wasn't sure what to expect at first. The pass rates for white candidates were just so much higher (around 68% for captains, I believe) while it was somewhere in the 30s for minority candidates. EEOC requires a four-fifths standard: this means that underrepresented minorities needed to have a pass rate of around 54% in order to make the tests appear valid. Anything below that, and you'll earn the wrath of the EEOC (Equal Employment Opportunity Commission created by the Civil Rights Act) and have a Title VII lawsuit on your hands.

I identified two major factors that contributed to the disparity. First, the composite score of 40/60 meant that each multiple-choice question on the written part (that constituted 60% of the score) was worth more individually. If the weighting was 70% instead, each multiple-choice question would decrease in importance. The second factor that is directly related to the composite breakdown was that PTI failed to undertake a departmental review. Rather than explaining to the City of Boston why an interdepartmental review was necessary, they simply accepted it. They then took no measures to account for the disparity caused by this factor. When you don't undertake an interdepartmental review, you inevitably get questions that are based on the source material, but that are not relevant to the City of Boston. Therefore, unless you poured through the materials that you managed to buy, you wouldn't be able to answer those questions. Therefore, that one multiple choice question that came out of left field could add up to decrease your score by a fraction and because of the "rule of three" put you out of the running for a promotion. It's a very frustrating situation that is understandably difficult to solve; you change something somewhere, and it may have an impact somewhere you never expected.

ID also proposed an assessment center that simulated real-world situations. We firmly believe that not all candidates can articulate themselves as well on paper or orally as they could in the field. Ultimately, what are we measuring? We're measuring how a candidate would respond to emergency situations. Sure, you may be able to brainstorm for a month how you could handle one situation, but would you have the instinct and reflexes to stay cool under real-life pressure? The assessment center would be a much better indicator than some oral exam.

I can't say what the right decision for the City of Boston is. There

HARVARD MODEL CONGRESS 2010

is no question that there was an adverse impact demonstrated by the results, and those couldn't stand without some alteration. However, knowing whether the city is in a position to cough up the money to create assessment centers requires information to which I don't have access. A few years down the road, this will certainly be the case. There is no question in my mind that the tests could be better and that the most qualified candidates could not have been identified. Identification Industries, Inc. would do anything it could to improve examinations for the future.

Martin/a Robespierre

Martin/a Robespierre

Subscribed and sworn to me on this, the 5th day of June, 2009.

Alexandra Makhholm, Notary Public

HARVARD MODEL CONGRESS 2010

Affidavit of John/Johanna Malkin

Witness for the Defense

My name is John/Johanna Malkin, and I'm 47 years old. I was born and raised in Boston and have been a member of the Boston Civil Service Board (or BCS as we refer to it) for the last 6 years. The Civil Service Board is an independent five-member organ of the City that administers the City's civil service employment system. This includes supervising the process of competitive examinations and reviewing their results before certifying lists of eligible candidates.

Our relationship with the "City" is often misunderstood. The BCS made the decision to not certify the exams. The City and its mayor, although they contributed their viewpoints and elaborated on their interests at the public hearings, did not make this decision. The BCS eventually split 2-2, which, by default, meant we couldn't take action to certify the exams. Our fifth member could not be involved in the deliberation process because her brother was a firefighter impacted by the promotion exams. The board received the breakdown of the scores and race of each candidate from the Director of Human Resources. It was brought to our attention that there was a severe adverse impact. The Director of HR presented us with the results and outlined the reasons that the board could refuse to certify the results.

We heard evidence during the first hearing in April 2008 that, although most exams have some disparate impact, in this case it was more significant. The pass rates were more disparate than they generally are. The Corporate counselor of Boston informed us that these numbers set off a warning bell because they evidenced a severe disparate impact. This meant that the EEOC would soon be swarming the city. Disparate impact called for very close examination of the policies and procedures of the city. We were informed that the city could be sued under Title VII (I don't think the irony that the city is still being sued under Title VII escapes anyone).

We also heard from the firefighters who took the exam themselves. Their scores were not reported to them because potential bias is always an issue. A number of firefighters still spoke to certify the exams. Sam Andrews was a very vocal voice. I truly sympathize with the time, energy, and money he poured into his studying. Although it may feel unfair that we didn't certify the exams, the studying he put in will be worth it in the long-run, regardless of what happens with his promotion. Also, Andrews scored well and was eligible for promotion, but he did not receive a promotion automatically. Because of the "rule of three," a candidate with a top score could, theoretically, not be chosen. The BCS operated under the understanding that we were not snatching a promotion away from any candidate because none had received one; they were simply identified as being qualified and eligible for one. There were also a number of firefighters who spoke against certifying the exam results. They testified that numerous questions were outdated, and others were not at all relevant to the job in Boston. A full set of the source materials cost \$600, and many firefighters reported that they could not afford this cost. However, this is a problem dealt with by students across the country, as another firefighter pointed out. As a result, study groups were formed and costs were split among group members.

The corporate counselor in the second hearing, I believe, made it very clear to the board that the city was in the running for a high-profile Title VII lawsuit if we certified the results. We were also informed that PTI was supposed to draw up a technical validity report, which we never saw. When the PTI director was questioned about it, he informed us that the city never asked for it. At this point, with the director in the position to answer all of our questions, we didn't feel that it was necessary to ask for a report.

HARVARD MODEL CONGRESS 2010

PTI also never offered to prepare one, and BCS believed it would not support the legality of the exams.

It was also brought to our attention throughout the hearings that there were possible alternative testing mechanisms. The ID president testified that assessment centers would identify the most qualified candidates more accurately than the PTI exams. Also, there needed to be an interdepartmental review or some way to weed out the questions that were not relevant to Boston's Fire Department. It's understandable that different knowledge is necessary when dealing with fires in California, Vermont, and Boston. There is also a significant difference between rural areas and large cities. To expect the candidates to know everything that came in nationally-approved books was ridiculous, even in our minds. We then considered a hypothetical situation: say we *did* certify the results and had a lawsuit on our hands. In that Title VII lawsuit, we would definitely lose, as the plaintiff would easily be able to prove that there was a disparate impact. We had so much evidence before us that an adverse impact existed, and then we ignored it? No judge would buy it.

I'm not going to tread lightly around the issue and say that we completely ignored race in our decision. We were conscious of it, of course. But, as has been proven time and time again, ignoring race is not the solution; rather, acknowledging it certainly is. Our decision was race-neutral in the sense that *no group* was placed at an advantage or disadvantage. Everyone will have to retake the exam, and no one's test results will be considered. We certainly did not treat the white candidates differently, although they somehow feel as if their situation is unfair. We always think that our own individual person suffers the most, when in reality we should try to keep in mind the bigger picture. The BCS and city officials certainly have a tough position in society: we have to balance disparate interests. There will always be someone that is unhappy. What's important to remember is that we're moving in the direction of progress. The 1965 Civil Rights Act was passed just a few decades ago, and we're still struggling to deal with the inequalities in the US. I hope that everyone will come to understand the fine balancing act we have to maintain in considering everyone's interest and the general good of our whole society.

John/Johanna Malkin

John/Johanna Malkin

Subscribed and sworn to me on this, the 5th day of June, 2009.

Alexandra Makhholm, Notary Public