RICCI V. DESTEFANO: RADICAL CHANGE IN DISPARATE IMPACT THEORY OR MUCH ADO ABOUT NOTHING?

ROBERT K. ROBINSON*  
DAVE L. NICHOLS**  
SAM COUSLEY***

I. INTRODUCTION

Ricci v. DeStefano,¹ popularly known as the “New Haven fire fighter’s case,” was decided by the Supreme Court of the United States on June 29, 2009. Many in the media contended that this decision marked a major setback for affirmative action.² Despite alarm in the headlines, the case actually had nothing to do with affirmative action; rather, it addressed very technical legal issues dealing with disparate impact theory and the employer’s burden of persuasion involving validated employment requirements, especially written examinations.

This paper serves three purposes. First, it examines how the Ricci decision affects current employment practices, especially validated selection metrics. To accomplish this, the authors provide a cursory explanation of the two unlawful forms of employment discrimination codified under Title VII of the Civil Rights Act of 1964, with particular emphasis being given to disparate impact. Second, it discusses the legal standards for demonstrating that a given selection device is job-related, and how this, a process known as validation, is generally accomplished. Finally, it analyzes the impact the Ricci decision has had on employer defenses predicated on validated test results.

* Ph.D., SPHR, Michael S. Starnes Professor, School of Business Administration, The University of Mississippi.  
** Ph.D., Associate Professor, School of Accountancy, The University of Mississippi.  
*** D.B.A., Clinical Professor, School of Business Administration, The University of Mississippi.  
¹ 129 S. Ct. 2658 (2009).  
II. UNLAWFUL EMPLOYMENT PRACTICES UNDER CIVIL RIGHTS ACT OF 1964

Title VII of the Civil Rights Act of 1964 governs employment practices in the workplace. Specifically, Title VII makes it unlawful for any employer, or his or her representatives, to:

(1) 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) 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 employee, because of such individual’s race, color, religion, sex, or national origin.3

Under Title VII, there are two major forms of unlawful discrimination which are actionable: Disparate treatment and disparate impact (adverse impact).4

A. Disparate Treatment

Disparate treatment is sometimes referred to as intentional discrimination. Under this form of discrimination, the employer intentionally withholds some job outcome because of an individual’s membership in a particular protected class. For example, an employer states to a witness that they do not wish to hire a female supervisor. As a consequence, no female applicant for that position will be considered, regardless of her qualifications. It is solely because the applicant is female, the wrong sex, that she is eliminated from consideration.

There are two ways of establishing actionable disparate treatment. First, it can be established by direct evidence. In this instance, evidence is produced that the employer in question openly intended to discriminate in his or her employment decision. As stated in our previous example, the employer making a comment in the presence of credible witnesses that he

would not hire a female for the job in question demonstrated disparate treatment. Rarely do such circumstances occur.

The other means of establishing disparate treatment is through the use of circumstantial evidence. The preferred method for establishing the requisite level of circumstantial evidence comes from the 1973 Supreme Court decision, *McDonnell Douglas v. Green.*\(^5\) Under the *McDonnell Douglas* burden shifting method, it must be demonstrated that:

1. The complaining party was a member of a protected class (i.e., race, color, religion, sex, or national origin).
2. The complaining party was qualified and applied (in the case of applications for employment or promotion), or held (in the case of demotions, layoffs, terminations, etc.).
3. The complaining party suffered an adverse employment consequence (i.e., was not hired, not promoted, bid off, terminated, etc.).
4. Someone who was similarly situated, but was not from the complaining party’s protected class, was treated differently (did not suffer the adverse employment consequence).\(^6\)

If the complaining party can satisfy all four of the above criteria, a *prima facie* case of disparate treatment has been established and the burden of persuasion shifts to the employer. In his or her rebuttal, the employer can defend the employment decision on the grounds that it was necessitated as a *bona fide* occupational qualification (BFOQ)\(^7\) or it was predicated on a legitimate nondiscriminatory reason.\(^8\) The BFOQ allows the employer to legally discriminate based upon religion, sex and national origin; it is extremely difficult to prove and is rarely available to most employers. In the vast majority of cases, employers turn to legitimate nondiscriminatory reasons under which the employer has demonstrated that the decision was based on a reason other than the complaining party’s protected class status.\(^9\) Usually this is accomplished by offering evidence that the other candidate was more qualified than the complaining party.\(^10\)

---

\(^{5}\) 411 U.S. 792 (1973).
\(^{6}\) *Id.* at 802.
\(^{9}\) ROBERT K. ROBINSON ET AL., EMPLOYMENT REGULATION IN THE WORKPLACE: BASIC COMPLIANCE FOR MANAGERS (2010).
\(^{10}\) Reeves, 530 U.S. at 148; Bell v. E.P.A., 232 F.3d 546, 553-54 (7th Cir. 2000).
nondiscriminatory reasons include, but are not limited to: seniority, business-related work experience, job-related education, or productivity.\textsuperscript{11}

If the employer can carry his or her burden of persuasion, the complaining party is afforded a final rebuttal. The complaining party must then show that the employer’s legitimate nondiscriminatory reason is a pretext in order to prevail.\textsuperscript{12}

\section*{B. Disparate Impact}

Disparate impact, or unintentional discrimination, occurs when an employer’s specific facially neutral employment practice has the effect of disqualifying a disproportionate number of applicants from a protected class (for the purpose of our discussion, our examples focus on how this affects qualifying examinations for promotion). This is usually accomplished by the complaining party offering reliable statistical proof that a specific employment practice excludes more members of a given protected class than members of another class.\textsuperscript{13} The most common method authorized under the Uniform Guidelines on Employee Selection Procedures is to use the so-called four-fifths rule (eighty percent rule).\textsuperscript{14} If the pass rate of the complaining party’s protected class is less than four-fifths (eighty percent) of that of the group with the highest pass rate, disparate impact is said to have occurred and the \textit{prima facie} case has been established.\textsuperscript{15}

\subsection*{1. Imbalances and the Four-Fifths Rule}

As the Uniform Guidelines clearly explain, violating the four-fifths rule has serious potential consequences because Federal agencies use it to evaluate disparate impact. The Guidelines state, “A selection rate for any race, sex, or ethnic group which is less than four-fifths 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.”\textsuperscript{16}

The four-fifths rule has an interesting and curious history. Apparently, the rule was first used by the Technical Advisory Committee on Testing (TACT) of the California Fair Employment Practices Commission. TACT

\begin{thebibliography}{9}
\bibitem{11} Robinson et al., \textit{supra} note 9.
\bibitem{12} Tex. Dep’t of Cmty. Affairs v. Burdine, 450 U.S. 248, 258 (1981).
\bibitem{14} 29 C.F.R. § 1607.4(D) (2009).
\bibitem{15} \textit{Id.}
\bibitem{16} \textit{Id.}
\end{thebibliography}
was looking for a rule of thumb that could be used to determine when validity information was required. The rule was developed, at least in part, because officials thought that statistical hypothesis testing could be beyond the ability of enforcement personnel. Nothing suggests that TACT expected the Federal government to include the four-fifths rule in guidelines or for courts to use it when detecting adverse impact. However, the four-fifths rule has become so widely used, that it has almost taken on a life of its own, and it is considered to be the standard metric for detecting adverse impact.\textsuperscript{17} Given the widespread impact of the four-fifths rule, it is surprising how little planning was done before the rule was promulgated.

The rule was accepted by a wide variety of Federal enforcement agencies and the courts as a meaningful and legitimate indicator of problematic adverse impact without quantitative analysis of its validity and without giving much thought to its results under different scenarios. One study, that revealed the rule’s shortcomings, used a variety of computer simulations to measure the number of “false-positives” found with the four-fifths rule. In this context, a false-positive is finding adverse impact when none actually existed.\textsuperscript{18} This study examined the performance of the four-fifths rule when there were no differences in the majority and minority populations (equal means and standard deviations). When the selection ratios and the number of applicants were both low, the four-fifths rule implied adverse impact quite often.

For example, when twelve percent of 200 applicants were minorities and the selection rate was 30 percent, the four-fifths rule signified adverse impact nineteen percent of the time. In another simulation, when twenty percent of 200 applicants were minorities and the selection rate was three percent, the false-positive rate was between forty-seven and seventy-eight percent. When the computer simulation used a very large number (2,000) of applicants, the false-positive rate of the four-fifths rule was much lower. One conclusion was that organizations that are very selective (with extremely low selection rates) are more likely to have false-positive results based on the four-fifths rule.\textsuperscript{19} In all cases, the number of false-positives was dramatically reduced if the four-fifths rule was used as the first step, followed by statistical testing in the presence of a four-fifths rule violation. Since the four-fifths rule was developed to avoid using statistical testing, it is somewhat ironic that these tests may be needed to reduce false-positives following violations of the four-fifths rule. Some people support use of the four-fifths rule by arguing that violations of the rule are only \textit{prima facie}

\textsuperscript{17}Philip L. Roth, et al., \textit{Modeling the Behavior of the 4/5ths Rule for Determining Adverse Impact: Reason for Caution}, 91 J. APPLIED PSYCHOL. 508 (2006).
\textsuperscript{18}Id. at 508-19.
\textsuperscript{19}Id. at 508-12.
evidence of adverse impact. Organizations can still assert validity and business necessity as defenses; however, violations of the four-fifths rule can cause serious repercussions including EEOC investigations, law suits, and damaged reputations.20

Another problem with the four-fifths rule is its sensitivity to small sample sizes. In fact, if an organization hires a small number of people, a shift in just one or two hires can have a huge impact on the conclusion when the four-fifths rule is used to signal adverse impact. This deficiency has also been noted by the EEOC and the Office of Federal Contract Compliance Programs (OFCCP), as both have acknowledged that adverse impact results are questionable when they are derived from “numbers which are too small to be reliable . . . .”21

To demonstrate the effect of small sample sizes, assume there are twenty majority and ten minority applicants for six job vacancies. If five majority applicants and one minority applicant are hired, the majority proportion is 5/20 = 0.25; the minority proportion is 1/10 = 0.10; and the Adverse Impact Ratio (AIR) is 0.10/0.25 = 0.40. The AIR is less than four-fifths, so Federal enforcement agencies will generally regard the results as evidence of adverse impact. On the other hand, if four majority and two minority applicants are hired (a shift in just one person), the conclusion is reversed. The majority proportion is 4/20 = 0.20; the minority proportion is 2/10 = 0.20; and the AIR = 0.20/0.20 = 1.00. Since the AIR is greater than eighty percent, this would not indicate evidence of adverse impact.22

2. Other Means of Evaluating Imbalances

Although the EEOC and the OFCCP still recognize the four-fifths rule, courts now prefer more precise tests based on statistical theory.23 Statistical tests in adverse impact cases are based on the assumption that the minority and majority applicant pools have equivalent population characteristics. The actual applicants are viewed as random samples from the populations, and the null hypothesis is that the two population selection rates are equal (minority = majority). If the difference in sample selection rates is statistically significant, the null hypothesis is rejected, and adverse impact is suspected.

It is important to note that statistical and legal significance are two different concepts, and there is no straightforward association between the

20 Id. at 512.
21 29 C.F.R. § 16074(D); 41 C.F.R. § 60-3.4D (2009).
two. This statistical imbalance can be established using other statistical means such as showing that the pass rate falls outside of two to three standard deviations in an expected distribution, or falls outside of a ninety-five percent confidence interval. It is also important to remember that “statistical significance” does not mean “practical significance” or even “important.” Some authorities contend that there is no compelling reason to use statistical significance as the one and only, or even the main, measure for judging importance.

Regardless of the statistical method used, once the imbalance has been demonstrated and the prima facie case has been established, the burden of persuasion shifts to the employer. It is absolutely essential to note that the presence of this disparate impact (the statistical imbalance) does not establish unlawful discrimination – it is inferred, but it is not yet proven.

3. Validation

If the employer can prove that the selection practice which causes the imbalance is a business necessity/job-related (both terms tend to be used interchangeably), then no violation of Title VII is said to have occurred. This is a far more onerous burden for an employer than the legitimate nondiscriminatory reason defense required in cases of intentional discrimination. Under a business necessity defense, the employer must be able to demonstrate, through a process known as validation, that there is a relationship between the neutral employment practice (a promotion examination score, e.g.) and the applicant’s or job incumbent’s ability to perform some aspect of the job in question. The premise is fairly simple, in the case of testing, passing scores predict the candidate’s ability to satisfactorily perform the job in question. Non passing scores would indicate the opposite.

27 James G. Combs, Big Samples and Small Effects: Let’s Not Trade Relevance and Rigor for Power, 53 Acad. Mgmt. J. 1, 10-12 (2010).
30 The authors do not wish to mislead the reader by implying that the business necessity defense is the only defense available to the employer; merely that it is the most common and the best defense. See Julia Lamber, Alternatives to Challenged Employee Selection Criteria: The Significance of Nonstatistical Evidence in Disparate Impact Cases under Title VII, 1985 Wis. L. Rev. 1 (1985). Other defenses include demonstrating that the complaining party’s
If the employer has shown that his or her selection practice is indeed validated, the complaining party has, as in disparate treatment, an opportunity to rebut the employer’s defense. In the case of disparate impact, the complaining party will prevail only if he or she can demonstrate that there is another selection practice which effectively predicts future job performance, but has less disparate impact on the underrepresented protected class. To illustrate, assume that the employer’s Test A has correlation between its test scores and job performance of \( r = 0.793 \).\(^{31}\) This would be a strong positive relationship, hence it would, by itself, be a good predictor. Unfortunately, assume that only fifty percent of African Americans pass the test compared to seventy percent of white applicants. Since the pass rate for African Americans compared to whites is 71.4% which is less than eighty percent, disparate impact has occurred.\(^{32}\) Fortunately, since the employer has already validated the test, he or she is able to defend the statistical imbalance created by the test because it is shown to be business related. However, assume that the complaining party in this scenario is able to provide the court with another test (Test B) which is also a good predictor, say \( r = 0.791 \). But this test has a fifty-eight percent pass rate among African Americans and a seventy-five percent pass rate among whites. The African American to white ratio is still below eighty percent but at 77.3%,\(^{33}\) it is higher than the 71.4% produced by Test A. The complaining party wins. If, on the other hand, the complaining party cannot produce another test, the employer prevails.

Now that the basics of disparate impact theory have been presented, the time has come to look at the findings of fact, and the technical issues contained in *Ricci*.

---

\( r = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n} (X_i - \bar{X})^2 \sum_{i=1}^{n} (Y - \bar{Y})^2}} \)

\(^{31}\) 0.50/0.70 = 0.7143 which is less than 80%, hence, under the *Uniform Guidelines on Employee Selection Procedures*, disparate impact would be concluded (see, 29 C.F.R. § 1607.4D)).

\(^{32}\) 0.58/0.75 = 0.7733 which is still less than 80%, but greater than the 0.7143 produced by Test A, hence it causes less disparate impact.
III. THE NEW HAVEN FIRE FIGHTER’S CASE

The central issue in *Ricci* was whether the City of New Haven could refuse to certify test results of a previously validated promotion examination purely on the grounds that these results created disparate impact. The promotion evaluation for lieutenant and captain consisted of two components: a written examination which counted for sixty percent of the candidate’s score, and an oral component which accounted for the remaining forty percent. A combined score of seventy percent would be a passing score.\(^{34}\) In November and December 2003, forty-one firefighters took the examination for captain and seventy-seven for lieutenant.\(^{35}\) The outcome of the examinations was as follows:

<table>
<thead>
<tr>
<th>EEO Classification of the Candidate</th>
<th>Candidates Taking Captain’s Examination</th>
<th>Candidates Passing Captain’s Examination</th>
<th>Pass Rate</th>
<th>Candidates Taking Lieutenant’s Examination</th>
<th>Candidates Passing Lieutenant’s Examination</th>
<th>Pass Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>25</td>
<td>16</td>
<td>.64</td>
<td>43</td>
<td>25</td>
<td>.581</td>
</tr>
<tr>
<td>Black/African American</td>
<td>8</td>
<td>3</td>
<td>.375</td>
<td>19</td>
<td>6</td>
<td>.316</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>8</td>
<td>3</td>
<td>.375</td>
<td>15</td>
<td>3</td>
<td>.2</td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td>22</td>
<td></td>
<td>77</td>
<td>34</td>
<td></td>
</tr>
</tbody>
</table>

The statistical outcomes based on the pass rates by each of the three protected classes were sufficient to establish a *prima facie* case under disparate impact analysis.\(^{36}\) The selection rate of Black/African American applicants for Captain was 0.586\(^{37}\) and 0.544 for lieutenant,\(^{38}\) both of which were well below the .8 specified in the *Uniform Guidelines*.\(^{39}\)

Although the four-fifths rule was violated in *Ricci*, statistical analysis of the data produces non-significant results for the Captain’s exam using Fisher’s Exact Test (for the white versus black comparison) and the Freeman-Halton extension of the Fisher test (for the white versus black versus Latino comparison). For the Lieutenant’s exam, the results were not statistically significant for the white versus black comparison, but they were significant for the white versus black versus Latino comparison.

However, the City was even more concerned by the bottom-line statistics. Bottom line statistics are those which consider the total number of applicants who begin a selection process and are eventually offered employment or promotion. In *Ricci*, promotions were determined by the City

---


\(^{35}\) Id.

\(^{36}\) 29 C.F.R. at § 1607 (2009).

\(^{37}\) \(0.375/0.64 = 0.586\).

\(^{38}\) \(0.316/0.581 = 0.544\).

\(^{39}\) 29 C.F.R. at § 1607.4(D) (2009).
Charter which required that any civil service vacancy (firefighters are such positions) had to be filled from among “the top three individuals with the highest scores on the exam.” None of the Black/African American applicants would be eligible for promotion under this scheme. This concern was the motivation factor behind the City’s refusal to certify the examination results. The end result, 17 white and one Latino firefighter were denied promotion.

It is interesting that in this case the Supreme Court never addressed the employer’s use of bottom-line statistics, even though employers are denied their use by the Civil Rights Act of 1991 and Connecticut v. Teal. When the affected firefighters filed a complaint with the EEOC, the City offered its justification for the refusal to certify the examination results was based on the grounds that to do otherwise would expose the City to litigation for disparate impact (unintentional discrimination). In communicating its rejection of the examination results, the City enunciated the action was motivated by “too many whites and not enough minorities would be promoted if the lists were certified.” The district court agreed with this line of reasoning and ruled that because the results would undermine the City’s diversity goals, the City was justified in its refusal to certify the examination.

At this juncture, the City in its attempt to avoid a disparate impact claim had stumbled into a disparate treatment (intentional discrimination) claim. Because none of the successful applicants were Black/African American, it would now be argued that the decision not to certify the examination results was a race conscious decision, since it was based solely on the race of the successful candidates.

In a single paragraph opinion, the United States Court of Appeals for the Second Circuit affirmed the decision “for the reasons stated in the thorough, thoughtful, and well-reasoned opinion of the [district court].” There was no further explanation. Just three days later, on June 12, 2008, the Second Circuit denied the firefighters’ request for a rehearing en banc in a complicated 12-page decision in which the page numbers of the decision were presented in a manner which appeared to be out of sequence.

40 Ricci v. DeStefano, supra note 34, at 145.
44 Morris & Libsenz, supra note 24, at 152.
45 Id. at 162.
46 Id. at 159-60.
47 Ricci v. DeStefano, 530 F.3d 87 (2nd Cir. 2008).
48 Ricci, 530 F.3d at 89.
IV. AN ETHICAL QUANDARY

The Supreme Court took a different approach to the case and would eventually find for the complaining parties, but would also open a potential Pandora’s Box. The Court ruled that an employer can engage in intentional discrimination for the expressed purpose of avoiding or remedying disparate impact, but only if the employer can show that he or she has a strong basis in evidence that it will be liable for disparate impact if it fails to take race-conscious action. Specifically, the Court ruled:

Under Title VII, before an employer can engage in intentional discrimination for the asserted purpose of avoiding or remedying an unintentional, disparate impact, the employer must have a strong basis in evidence to believe it will be subject to disparate-impact liability if it fails to take the race-conscious, discriminatory action. This wording is bothersome because the Supreme Court is explicitly stating that there are instances in which Title VII’s prohibition on race-conscious, discriminatory action can be circumvented in order to avoid unintentional discrimination. This becomes one of those rare instances in which a semantic inconsistency occurs. An employer is, in effect, permitted to violate Title VII in order to comply with Title VII.

Previously this line of reasoning had been applied only to those circumstances involving government entities who were previously culpable for historical discrimination. For instance, Justice Powell’s opinion in *Wygant v. Jackson Board of Education* concluded there may be circumstances in which it may be necessary to consider race in employment decisions, but only when it is done to remedy the effects of past discrimination. The justification in these instances was remedial; to eliminate the present effects of the employer’s past discrimination practices, The Supreme Court has even previously held that such race-conscious remedial action could be extended even to those who were not victims of discrimination.

This intentional use of race-conscious decision-making for “remedial” purposes was also held to be constitutional under the equal protection clause

---

49 Ricci v. DeStefano, supra note 41, at 2681.
50 Id. at 2677.
51 Id. at 2661.
52 Id. at 2677
54 Id. at 279-80; Fullilove v Klutznick, 448 U.S. 448, 486 (1980).
of the Fourteenth Amendment in City of Richmond v. J. A. Croson Co. In short, an employer could, in certain instances, be permitted to engage in intentional discrimination contrary to Title VII’s ban against such race-conscious action. It should be pointed out for clarity that these previous decisions all involved government-initiated affirmative action plans, and that the City of New Haven did not invoke such plans in its defense.

The Ricci Court has borrowed from these cases to justify employers’ race preferences when they can produce a strong basis in evidence that remedial action is needed. This is essentially providing an affirmative action justification when no affirmative action plan is present.

What is learned from Ricci is that engaging in intentional discrimination is not permitted in order to avoid litigation for unintentional discrimination. However, it would be permitted in order to avoid liability for unintentional discrimination. The burden is clearly on the employer to provide evidence that there is a very real likelihood for liability under disparate impact theory. Because the City had a properly validated promotion examination, it could not produce tangible evidence that its selection process would result in liability for unintentional discrimination. The validated tests provided a job-related defense for the statistical disparities in the test results. As a consequence, the City was precluded from justifying its refusal to certify the examination results.

V. CONCLUSION

In the Ricci case, there was a little something old and a little something new. The “old” involved the Court’s upholding the axiom that valid tests are still the employer’s best defense against disparate impact claims. The City could weather the disparate impact litigation because the promotion examinations were job-related. An individual can challenge under disparate impact theory but a job-related/business necessity defense is adequate for explaining and justifying the disparity.

The “new” involved the expansion of disparate treatment to avoid disparate impact liability. Resolving competing expectations under the disparate-treatment and disparate-impact provisions: A respondent must demonstrate, based on the strong basis in evidence that, had it not certified the results, it would have been subject to disparate-impact liability.