2009

Fourth and Short on Equality: The Disparate Impact of the NFL's Use of the Wonderlic Intelligence Test and the Case for a Football-Specific Test Note

Christopher Hatch

CONNECTICUT
LAW REVIEW

VOLUME 41 JULY 2009 NUMBER 5

Note

FOURTH AND SHORT ON EQUALITY: THE DISPARATE IMPACT OF THE NFL'S USE OF THE WONDERLIC INTELLIGENCE TEST AND THE CASE FOR A FOOTBALL-SPECIFIC TEST

CHRISTOPHER HATCH

Prior to being selected in the NFL draft, a player must undergo a series of physical and mental evaluations, including the Wonderlic Intelligence Test. The twelve-minute test, which measures "cognitive ability," has been shown to have a disparate impact on minorities in various employment situations. This Note contends that the NFL's use of the Wonderlic also has a disparate impact because of its effect on a player's draft status and ultimately his salary. The test cannot be justified by business necessity because there is no correlation between a player's Wonderlic score and their on-field performance. As such, this Note calls for the creation of a football-specific intelligence test that would be less likely to have a disparate impact than the Wonderlic, while also being sufficiently job-related and more reliable in predicting a player's success.

NOTE CONTENTS

FOURTH AND SHORT ON EQUALITY: THE DISPARATE IMPACT OF THE NFL'S USE OF THE WONDERLIC INTELLIGENCE TEST AND THE CASE FOR A FOOTBALL-SPECIFIC TEST

CHRISTOPHER HATCH*

I. INTRODUCTION

Every February, hundreds of college football players are invited to Indianapolis to participate in the NFL combine, the league's evaluation program for potential NFL players. This "cattlecall"[1] for NFL hopefuls serves as an opportunity for personnel from every team to assess players' speed, strength, durability and acumen before determining whom to select in the NFL draft. Invitees to the combine are evaluated in many activities including a 40-yard dash, bench press, vertical jump, injury examination, personal interviews and an intelligence test.[2]

Every aspect of a player's makeup is analyzed over the course of two days—nothing is immaterial.[3] There can be tremendous payoffs for those who do well; successful performance can lead to a relatively obscure prospect being drafted, or an already established player can increase his value and ultimately his paycheck with an impressive showing.[4] Tenths of a second can literally equate to hundreds of thousands of dollars in gains or losses.[5]

Testing a potential NFL player's speed, strength and overall health is

---

* Brigham Young University, B.S. 2006; University of Connecticut School of Law, J.D. Candidate, 2009. Many thanks to Professor Peter Siegelman for his comments and guidance throughout the writing of this Note. All errors contained herein are mine and mine alone.

[1] Pete Prisco, *Cattle Call Under Way for Hopefuls at NFL Combine,* FLORIDA TIMES-UNION, Feb. 21, 1999, at C9, *available at* LEXIS, News Library, FLATUN File.

[2] Ryan Christopher DeVault, Wonderlic Test Begins at the 2009 NFL Combine; Every Prospect is Required to Take It (Feb. 19, 2009), http://www.associatedcontent.com/article/1494254/wonderlic_test_begins_at_the_2009_nfl.html?cat=14; Nathan Rush*, NFL Combine-Offense*, ATHLON SPORTS, Feb. 23, 2009, *available at* http://www.athlonsports.com/pro-football/16371/nfl-combine-offense; NFL Scouting Combine, http://www.nfl.com/combine/workouts (last visited Mar. 18, 2009).

[3] Ryan Rigmaiden, Michael Allan Post-Combine Interview!, (Feb. 27, 2007), http://sea.scout.com/2/622322.html (noting that potential NFL prospect Michael Allan was asked by the Philadelphia Eagles staff in an interview "who [he] would call if [he] was thrown in jail and was given . . . one phone call.").

[4] Michael Allan played college football in near-obscurity at Division III Whitworth University, but due to his exceptional performance at the combine he was drafted by the Kansas City Chiefs in the seventh round of the 2007 NFL draft. *See id.*; ESPN NFL DraftTracker, Round 7, http://insider.espn.go.com/nfldraft/draft/tracker/round?round=7&draftyear=2007 (last visited Mar. 18, 2009).

[5] Jim Corbett, *At the NFL Combine: Officials Reassess Wonderlic's Value*, USA TODAY, Feb. 23, 2007, *available at* http://www.usatoday.com/sports/football/draft/2007-02-22-combine-notebook_x.htm (noting that running back Maurice Jones-Drew ran a 4.39 second 40-yard dash at the 2006 combine, which the Jaguars personnel suggested was a significant reason he was drafted 60th overall).

undoubtedly essential in evaluating potential; however, the combine also tests for "intelligence," via the Wonderlic Personnel Test (Wonderlic), a controversial IQ test that has been shown to have a disparate impact on minority groups in a variety of employment situations.[6] The theory of disparate impact holds that it is unlawful under Title VII of the 1964 Civil Rights Act for an employer to use a facially neutral practice, such as a test or educational requirement, which has an adversely disproportional impact on a protected group, even without any discriminatory intent on the employer's part.[7] If the test has an adverse impact, the employer can avoid liability only if the test is shown to be "job related for the position in question and consistent with business necessity . . . ."[8]

In the framework of the NFL Draft, the Wonderlic almost certainly has a disparate impact because of its effect in determining draft ranking and ultimately a player's salary. Moreover, the test also does not appear to accurately forecast future performance in the NFL. As such, this Note concludes that the NFL's use of the Wonderlic cannot be justified by business necessity and would likely fail a Title VII disparate impact challenge. This Note calls for the creation of a football-specific "intelligence" examination that would be less likely to have a disparate impact on a player's draft status and salary. Such a test would also be sufficiently job-related and more reliable in predicting a player's success.

This Note will give a brief explanation of the Wonderlic test and its relation to disparate impact law. Then, in order to show that the test has a disparate impact on NFL draft pick selection, the Note will focus on the relationship between quarterbacks' Wonderlic scores, their draft positions, and their performance on the field. Quarterbacks' Wonderlic scores are the easiest to procure and the most useful to examine because their on-field performance is the most straightforward to evaluate and the position is generally regarded as the most cerebral.[9] If the test does not predict quarterbacks' future performance, *a fortiori*, it will be even less likely to predict the performance of players in less intellectually-demanding positions. In the interest of full disclosure, the data used in my analysis is not publicly available, but has been leaked through many different sources.

---

[6] Michael A. Reiter, *Compensating for Race or National Origin in Employment Testing*, 8 LOY. U. CHI. L. J. 687, 699–702 (1977).

[7] *See* BARBARA T. LINDEMANN & PAUL GROSSMAN, EMPLOYMENT DISCRIMINATION LAW 109–10 (4th ed. 2007) (explaining that "disparate treatment focuses on discriminatory *intent* while adverse impact focuses on discriminatory *consequences*").

[8] 42 U.S.C. § 2000e-2(k)(1)(A)(i) (2006).

[9] Michael Callans, President of Wonderlic Consulting stated that "[s]electing a new quarterback is like hiring a president for a company . . . They need the intelligence to think on their feet, evaluate all of their options and understand the impact their actions will have on the outcome of the game." Intelligence Testing in the National Football League, NFL Testing Provides Valuable Lesson for All Employees (Mar. 1, 2005), http://www.assessmentpsychology.com/nfl.htm [hereinafter Intelligence Testing].

Because the data is not completely reliable, I will only be offering a preliminary evaluation of the test's link to performance on the field.

II. BRIEF HISTORY OF THE WONDERLIC

The Wonderlic Personnel Test was developed by industrial psychologist Eldon F. Wonderlic and first distributed in 1937.[10] The test, which measures "cognitive ability," contains fifty questions that become progressively harder and must be completed in twelve minutes.[11] Wonderlic, Inc.—a for-profit corporation that markets the test—claims that the test is long enough to ensure validity while short enough not to intimidate.[12] Wonderlic representatives say that the test is administered 3,000,000 times a year, for different levels of employment, and by about 7,000 of their clients.[13] The score is calculated by simply totaling the number of correct answers in the allotted time;[14] a score of twenty-one is equivalent to the intelligence of a high school graduate and constitutes an approximate average score in the United States.[15] The NFL is the only sports league that uses the test.[16]

The history of the Wonderlic's use in the NFL is unclear, but one of its most ardent proponents was Tom Landry, former head coach of the Dallas Cowboys in the early 1970's.[17] Landry, who wanted to evaluate more than just on-field performance, contended that players who used their minds would have a strategic advantage over the other teams in the league.[18] After the Cowboys began using the test, other teams followed suit, leading to today's NFL-wide administration of the Wonderlic at the combine.[19]

III. THE NFL'S USE OF THE WONDERLIC AND DISPARATE IMPACT

The claim of disparate impact in employment discrimination law arose under Title VII of the Civil Rights Act of 1964, which made any forms of discrimination in the workplace illegal.[20] Title VII forbids discrimination in all aspects of employment on the basis of race, color, national origin,

---

[10] Wonderlic: Our History, http://www.wonderlic.com/about-us.aspx (last visited Mar. 19, 2009).

[11] Darren Rovell, Grading Wonderlic and the Best Sports Video Game, (Feb. 21, 2007), http://www.cnbc.com/id/17258325; Wonderlic: Our History, *supra* note 10.

[12] Wonderlic: Our History, *supra* note 10.

[13] Rovell, *supra* note 11.

[14] McDonald P. Mirabile, *Intelligence and Football: Testing for Differentials in Collegiate Quarterback Passing Performance and NFL Compensation*, 8 SPORT J. 2005, *available at* http://www.thesportjournal.org/tags/2005?page=2.

[15] Reiter, *supra* note 6, at 702; Rovell, *supra* note 11.

[16] Rovell, *supra* note 11.

[17] Intelligence Testing, *supra* note 9.

[18] *Id.*

[19] *Id.*

[20] Michael S. Beer, *Title VII Today: The Shift Away From Equality*, 20 J. MARSHALL L. REV. 525, 525 (1987).

sex, and religion and has been developed through two separate legal theories—disparate treatment and disparate impact.[21] Disparate treatment, which encompasses intentional acts of employment discrimination and may be applicable in relation to the NFL in other respects, is not a focus of this Note.[22]

As stated above, the theory of disparate impact holds that it is unlawful for an employer to use a facially neutral practice, such as a test or educational requirement, which has a disproportionately adverse impact on a protected group, even without a showing of the employer's discriminatory intent.[23] The seminal disparate impact case is *Griggs v. Duke Power Co.*, in which the Supreme Court struck down an employer's hiring and promotion practices as being in violation of § 703 of Title VII.[24] In that case, the plaintiffs challenged an employer's requirement of a passing score on two intelligence tests (one of which was the Wonderlic) or a high school education as a condition for employment or promotion.[25]

The *Griggs* Court noted that Section 703(h) provides it shall not be an "unlawful employment practice for an employer . . . to give and to act upon the results of any professionally developed ability test provided that such test, its administration or action upon the results is not designed, intended or used to discriminate because of race, color, religion, sex or national origin . . . ."[26] But the Court also held that tests, even if impartial on their face and in terms of intent, may be illegal under Title VII if their results are disproportionately adverse to minority groups.[27] This adverse impact was apparent when the Court noted EEOC studies finding that the use of intelligence tests, including the Wonderlic, resulted in passage rates for whites of fifty-eight percent, compared to six percent for blacks.[28]

---

[21] 42 U.S.C. § 2000e-2(a) (2006); GEORGE RUTHERGLEN, MAJOR ISSUES IN THE FEDERAL LAW OF EMPLOYMENT DISCRIMINATION 7 (4th ed. 2004).

[22] *See* Tom Gage, *National Pastime Strikes Out with Black Athletes; Kids in Michigan, U.S. Choose the Flash of NFL and NBA Over Baseball's Slow Pace*, THE DETROIT NEWS, Apr. 10, 2005 at 1A, *available at* LEXIS, News Library, DETNWS File (noting that in 2004, sixty-nine percent of NFL players were black). For a cogent analysis of racial discrimination in the NFL, see Jason Chung, *Racial Discrimination and African-American Quarterbacks in the National Football League 1968–1999*, MCGILL U. 4–9 (NOV. 29, 2004), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=835204. The article points to two flawed arguments that have been used to show why African-American quarterbacks are not as successful in the NFL—the option argument and the intelligence argument. *Id.* The option argument notes that the NFL uses a pass-oriented offensive system which is more complex than option offenses that many black quarterbacks used in high school or college. *Id.* at 4. An option offense relies primarily on a quarterback's athletic ability to run the ball himself or choose to lateral the ball to a teammate. *Id.* at 5. The author notes that there was prejudice against option quarterbacks because of the simplicity of the offense. *Id.* at 6. The intelligence argument, driven by the use of the Wonderlic, will be examined in this Note.

[23] LINDEMANN & GROSSMAN, *supra* note 7, at 109–10.

[24] Griggs v. Duke Power Co., 401 U.S. 424, 436 (1971).

[25] *Id.* at 425–26, 428.

[26] *Id.* at 426 n.1 (citations omitted).

[27] *Id.* at 431.

[28] *Id.* at 430 n.6 (citations omitted).

The *Griggs* Court concluded,

> The Act proscribes not only overt discrimination but also practices that are fair in form, but discriminatory in operation. The touchstone is business necessity. If an employment practice which operates to exclude [members of a protected group] cannot be shown to be related to job performance, the practice is prohibited.[29]

To determine business necessity, the Court examined the evidence and noted that employees who had not graduated from high school or who had not take the intelligence tests had still been able to satisfactorily perform the duties of their job and that the company had not made a showing that the promotion requirements fulfilled a "genuine business need."[30]

The Supreme Court clarified the theory of disparate impact in *Albemarle Paper Co. v. Moody*, a case that established the framework with which to bring a claim.[31] The facts of the case were similar to those in *Griggs*; in order to receive a promotion or seniority advancement an employee had to have a high school diploma and pass two intelligence tests, one of which was the Wonderlic.[32] Intending to show that their tests met the business necessity standard, the defendants hired an industrial psychologist to validate the job relatedness of its testing program.[33] The study, validated at the trial level, showed a statistically significant correlation between test scores and ratings from supervisors,[34] but the Supreme Court rejected the legitimacy of the study based on, among other factors, the subjectivity of these ratings.[35]

The Court proceeded to elaborate on the allocation of the burden of proof in a disparate impact claim by holding that the complaining party must "ma[k]e out a prima facie case of discrimination, *i.e.*, has shown that the tests in question select applicants for hire or promotion in a racial pattern significantly different from that of the pool of applicants."[36] From there, the employer must show that their employment tests which have a discriminatory effect have a "manifest relationship" to the job in question.[37] If the employer meets their burden of showing that their test is

---

[29] *Id*. at 431.
[30] *Id.* at 432.
[31] Albemarle Paper Co. v. Moody, 422 U.S. 405, 425 (1975).
[32] *Id.* at 410–11.
[33] *Id.* at 411.
[34] *Id.*
[35] *See id.* at 431–36 (noting that the validation study was also defective because it produced an "odd patchwork of results" in that the test only showed significant correlation in two of the eight skilled lines of progression within the company, it mainly focused on the top job groups, and only dealt with job-experienced white workers and not job applicants, who were often non-white).
[36] *Id.* at 425.
[37] *Id.* (quoting Griggs v. Duke Power Co., 401 U.S. 424, 432 (1971)).

job related, "it remains open to the complaining party to show that other tests or selection devices, without a similarly undesirable racial effect, would also serve the employer's legitimate interest in 'efficient and trustworthy workmanship.'"[38]

This framework for disparate impact was ultimately codified through the Civil Rights Act of 1991 as Section 703(k) of Title VII.[39] First, the plaintiff must demonstrate a prima facie case that the defendant "uses a particular employment practice that causes a disparate impact on the basis of race, color, religion, sex, or national origin . . . ."[40] At that point, the burden of production and persuasion is shifted to the employer to show that their practice or selection device is "job related for the position in question and consistent with business necessity . . . ."[41] If the employer is successful at this stage, the plaintiff then has the burden to show that the employer refused to implement an alternative hiring practice that could achieve the same legitimate business goal with less adverse impact.[42]

The difficulty of making a prima facie case in an NFL draft context is that unlike *Griggs* and other cases in which the Wonderlic has been administered to potential employees, NFL teams apparently do not use a cutoff score; instead they employ the test as one factor among many in determining draft position (and ultimately, salary). When sufficiently high cutoff scores are used, the test can indisputably have a disparate impact on minorities.[43] For example, if an employer's minimum passing score is 21, the level of a high school graduate, 75.1% of all black applicants would be excluded compared to 34.9% of all white applicants[44]—certainly enough to show a significant disparity.

This analysis cannot be directly applied to the NFL's use of the Wonderlic, however, because the test does not form a rigid barrier to employment; if an individual's athletic ability is high enough, he will be

---

[38] *Id.* (quoting McDonnell Douglas Corp. v. Green, 411 U.S. 792, 801 (1973)).

[39] LINDEMANN & GROSSMAN, *supra* note 7, at 110.

[40] 42 U.S.C. § 2000e-2(k)(1)(A)(i) (2006). Alternatively, if the complaining party cannot separate the respondent's employment practices in order to show a specific practice that has a disparate impact, "the decisionmaking process may be analyzed as one employment practice." *Id.* §2000e-2(k)(1)(B)(i).

[41] *Id.* § 2000e-2(k)(1)(A)(i).

[42] *Id.* § 2000e-2(k)(1)(A)(ii).

[43] *See* Reiter, *supra* note 6, at 701 (noting that a "very stable differential in raw scores achieved by Negro Applicant Populations exists . . . These mean score differentials are, as other researchers have noted in the study of mental ability, about one standard deviation apart when comparisons of Caucasians and Negroes are studied.") (citation omitted). Unfortunately, my research could not find any more recent data on white or black Wonderlic performance. However, significant research on test score disparities between races is readily available. *See* Christopher Jencks & Meredith Phillips, *The Black-White Test Score Gap*, *in* THE AFRICAN AMERICAN PREDICAMENT 63, 63 (Christopher H. Foreman, Jr. ed., 1999) (noting that "African Americans score lower than European Americans on vocabulary, reading, and math tests as well as on tests that claim to measure scholastic aptitude and intelligence" and that "the typical American black still scores below 75 percent of American whites on almost every standardized test.").

[44] Reiter, *supra* note 6, at 702.

drafted regardless of his score on the Wonderlic. Nevertheless, it is indisputable that the Wonderlic often plays a factor in a team's decision of whom to draft,[45] and therefore, how much money to offer a player after he has been drafted. The fact that there is not a cutoff score should not preclude a plaintiff's prima facie case. If the test is used to determine draft position, it will invariably have a disparate impact on a player's salary in the NFL; on average minorities will score lower on the test and thus be drafted later leading to a decrease in salary, or a player's poor performance on the test could lead to them not being drafted at all.[46] Proving wage-based disparate impact in this context presents unique challenges that will be discussed below.

Although disparate impact claims are generally developed through examining impacts on a group, the Supreme Court has stressed that the essential protection provided by Title VII is the opportunity for the *individual* to be treated fairly.[47] The notion of protecting the individual's opportunity is crucial in evaluating the NFL's use of the Wonderlic. The NFL would likely argue that many black athletes have scored poorly on the Wonderlic and still been drafted in the first round, including quarterbacks Vince Young,[48] Jason Campbell,[49] and Akili Smith,[50] thus suggesting the absence of any disparate impact in the selection process. However, there are players whose draft status possibly tumbled as a result of poor

---

[45] Michael Callans, President of Wonderlic Consulting, said he thinks "all teams look at [Wonderlic scores] to some degree. I haven't had a team tell me that they didn't think it was of any value at all." *See* Rovell, *supra* note 11.

[46] The prima facie case of disparate impact in this context cannot be rebutted by the fact that a strong majority of the players in the NFL are black. This argument, known as a "bottom line" defense, was rejected by the Supreme Court in *Connecticut v. Teal*, 457 U.S. 440, 445, 453 (1982). In that case, an employer used a written examination, shown to have a disparate impact, in its promotion decisions. *Id.* at 443–44, n.3. The employer raised a bottom-line defense, asserting that despite the written test, 22.9% of blacks who passed the examination were promoted, compared to 13.5% of whites, so that in the end, blacks were overrepresented among those who actually got jobs. *Id.* at 444. The Court turned to the language of Title VII which states that it is unlawful to "'limit . . . or classify . . . applicants for employment . . . in any way which would deprive or tend to deprive any individual of employment *opportunities.*'" *Id.* at 448 (emphasis original) (quoting 42 U.S.C. § 2000e-2(a)(2)(2006)).

[47] *Teal*, 457 U.S. at 448.

[48] Perhaps the most notorious example of a Wonderlic score being leaked occurred when University of Texas quarterback Vince Young's score at the 2006 NFL combine surfaced. Young, whose draft stock rose one month earlier after almost single-handedly beating the University of Southern California in the Rose Bowl, entered the combine as a certain top four pick, but when a score of 6 was reported for Young, his status as a prototype quarterback began to be questioned. *See* Pete Doherty & Jim Wyatt, *Will Wonderlic Cause Teams to Wonder about Young?* (Mar. 1, 2006), http://www.usatoday.com/sports/football/draft/2006-03-01-young-wonderlic_x.htm. According to Jeep Chryst, a former NFL assistant attending the combine as an at-large scout, Young's extremely low score "raise[d] a huge red flag." *Id.* Different sources began to report that the test was administered or graded incorrectly. *Id.* Young took the test again approximately two weeks later and scored a 16, still below average for non-college graduates and lower than the mid-20 range that NFL teams are looking for in a quarterback. *Id.*

[49] Michael A. McCann, The Wonderlic Test for the NFL Draft: Linking Stereotype Threat and the Law 26 (Oct. 1, 2006), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=934307).

[50] *Id.*

Wonderlic scores.[51]

Such a fall in draft standing is more likely to affect African-American football players, especially African-American quarterbacks who on average score under twenty on the Wonderlic more often than their white counterparts.[52] To succeed in establishing a prima facie case of disparate impact in this context a plaintiff would have to show, through statistical analysis, that minority players as a group score lower than whites on the Wonderlic (which is undisputed),[53] that they are drafted later than they would have been had the test not been administered, and that as a result the test has a disparate impact on minority players' salaries.[54]

Through empirical research conducted for this Note, I compiled Wonderlic scores for 104 quarterbacks from 1999 to 2006 (Figure 1). The Wonderlic scores used in my analysis come from Mac Mirabile's website.[55] Mirabile states that his results come from personal research and generally reliable sources, including notes from scouts and newspaper articles.[56] Nevertheless, because the scores are not made public knowledge they cannot be completely verified.[57] The data confirms that there is a marked disparity between black and white quarterbacks. Figure 2 shows the average Wonderlic scores and standard deviation by race, with number of observations. This data shows that black quarterbacks score, on average, a full 7.75 points lower than their white counterparts on the Wonderlic. In addition, there is good evidence that the higher a quarterback scores on the Wonderlic the earlier he is drafted. In other words, NFL teams do use the Wonderlic, among other things, in determining draft order, and hence, in setting salary. Figure 3 plots the Wonderlic score and the draft position of each of these quarterbacks. The downward sloping regression line implies that as a player's test score has a

---

[51] *See id.* at 26–27 (noting quarterbacks Jeff Blake, Kordell Stewart, Anthony Wright, and Randall Cunningham's draft status may have suffered as a result of the Wonderlic). Blake, Stewart, White, and Cunningham all scored under twenty on the Wonderlic. NFL Quarterback Wonderlic Scores, http://www.unc.edu/~mirabile/Wonderlic.htm (last visited Apr. 6, 2009).

[52] Chung, *supra* note 22, at 7.

[53] *See supra* text accompanying note 39.

[54] *See* GEORGE RUTHERGLEN, EMPLOYMENT DISCRIMINATION LAW, VISIONS OF EQUALITY IN THEORY AND DOCTRINE 80 (2nd ed. 2007) ("The disparity revealed by statistical evidence should be both statistically and practically significant . . . ."). A successful plaintiff would also have to convince a court that the test was not job related and consistent with business necessity. 42 U.S.C. § 2000e-2(k)(1)(A)(i) (2006).

[55] NFL Quarterback Wonderlic Scores, http://www.unc.edu/~mirabile/Wonderlic.htm (last visited Apr. 6, 2009).

[56] *Id.* Although the NFL takes precautions to ensure that Wonderlic scores are not released to the public, as fans' interest in the combine has increased so have leaks of individual players scores. Sam Walker, *The NFL's Smartest Team*, WALL ST. J., Sept. 30, 2005 at W1, *available at* LEXIS, News Library, WSJNL File (stating that National Football Scouting Inc., which runs the NFL combine, has closed the test to team personnel and upon the end of the combine burns the test scores on thirty-two DVDs and sends it by Federal Express to each team to prevent leaks).

[57] *Id.*

positive effect on his expected draft position. The estimated equation for this line is

Draft # = 165 – 1.95 x Wonderlic Score

(t-stat) (t-stat)

What this means is that a player with a Wonderlic score of 0 would be predicted to be drafted #165 and each additional point on the Wonderlic score raises a quarterback's draft number by just under two positions. The *p-value* (a statistic that measures the probability of obtaining a result at least as extreme as the one observed) for this analysis is 0.113, which is close to statistical significance at the ten percent level.[58]

Looking at the data another way, I examined drafted quarterbacks who scored greater than or equal to one standard deviation above and below the mean score in my sample (26.25) to show the impact of the test on a player's draft status and future salary (Figure 4). This analysis did not show statistical significance, largely because of the small sample size and the substantial deviations in draft numbers within both the high-scoring and low-scoring groups. The average draft position of those who scored greater than or equal to 33 (one standard deviation above) on the Wonderlic was 107.8. Of those who scored less than or equal to 19 (one standard deviation below) the average draft position was 136.4. While a difference of over twenty-eight draft slots is not shocking, it is also not insignificant—with thirty-two teams in the league, it equates to being drafted almost one full round later and likely having a smaller rookie contract. For example, the difference in average annual salary between the 108th pick and 137th in the 2008 NFL Draft was $69,500.[59] Because black quarterbacks score lower on the test than whites,[60] the Wonderlic would have a disparate impact on their salary.

The inherent problem in making this prima facie case is that the Wonderlic test is not used as the sole factor to set a player's salary. There are many other factors that are used in assessing a player including speed,

---

[58] The p-value measures the probability of getting a difference between the sample mean and the null hypothesis which is numerically greater or equal to that actually observed. Generally, in order to show statistical significance the p-value must be smaller than or equal to the significance level. *See* JOHN E. FREUND & GARY A. SIMON, MODERN ELEMENTARY STATISTICS 314 (8th ed. 1992).

[59] After being drafted 108th by the Denver Broncos, Kory Lichtensteiger signed a four-year $2.189 million dollar contract. Kory Lichtensteiger, http://www.rotoworld.com/Content/player pages/player_main.aspx?sport=nfl&id=4816 (last visited Apr. 6, 2009); 2008 NFL Draft Round 4 http://sportsillustrated.cnn.com/football/2008/draft/breakdowns/by_rounds/4.html (last visited Apr. 6, 2009). John David Booty, drafted 137th by the Minnesota Vikings, signed a four-year $1.911 million dollar contract. John David Booty, http://www.rotoworld.com/Content/playerpages/player_ main.aspx?sport=NFl&id=4732 (last visited Apr. 6, 2009); 2008 NFL Draft, Round 4 http://sportsillustrated.cnn.com/football/2008/draft/breakdowns/by_rounds/4.html (last visited Apr. 6, 2009).

[60] Of the 103 quarterbacks sampled in the above analysis, seventy-eight white quarterbacks had an average score of 28.15 and twenty-five black quarterbacks had an average of 20.4. *See infra*, Figure 2.

strength, interview skills, and college performance.[61] It is concededly difficult for any individual player to show how heavily his Wonderlic score weighed in a team's decision to not draft him. A team can point to many other factors besides Wonderlic scores as reasons for not drafting a particular player. However, a compelling argument has been made by Ian Ayres that these other factors must be ignored in a disparate impact case in order to determine if they cause a racial disparity.[62] Ayres notes that controlling for non-race factors is inappropriate because the purpose of the analysis is to see if these non-race factors cause racial disparities.[63] Thus, attempting to control for these other factors (college performance, physical tests, etc.) can bias the determination as to whether the employer's policies causes an unjustified disparate impact.

Secondly, because the NFL is more popular than ever; the demand for combine and draft information is insatiable, including twenty-six hours of live coverage of the combine on the NFL Network.[64] There are many sources that leak Wonderlic scores to the public and the media often belittles players with low test results.[65] It is entirely possible that as players' scores become public knowledge, it could more readily affect teams' decisions regarding which prospects to select, thus worsening the already disparate effect of the Wonderlic on future NFL players.

IV. LEGAL ANALYSIS OF DISPARATE IMPACT IN SALARY SETTING

Although this Note does not cover a classic disparate impact claim based on a facially neutral test affecting the *selection* rates of a protected class, the making of a prima facie case by a plaintiff should not be prohibited if he can show that the NFL's use of the Wonderlic resulted in black players being drafted later and thus receiving a smaller initial contract. Courts have addressed disparate impact in relation to wage in many circumstances, including the Age Discrimination in Employment Act

---

[61] *See* Dougherty & Wyatt, *supra* note 48 (noting that "[i]n addition to the Wonderlic, teams also determine a player's intelligence through interviews and their success in school.").

[62] Ian Ayres, Three Tests for Measuring Unjustified Disparate Impacts in Organ Transplantation: The Problem of Included Variable Bias 2 (Dec. 28, 2003) (unpublished Yale Law and Economics Research Paper No. 290, *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=483242) (stating that it is "*necessary* to intentionally omit non-race based variables from a regression to test whether those variables produced a racially disparate impact.").

[63] *Id.*

[64] Jon Ourand, Reebok to Sponsor NFL Network's Scouting Combine Coverage (Feb. 15, 2008), http://www.sportsbusinessdaily.com/article/118558.

[65] McCann, *supra* note 50 at 22 (noting that there are "myriad sources of leaks" including team officials, players' agents, rival agents and players themselves). After Vince Young scored a 6 on his first Wonderlic attempt, one author wrote that "Young is perceived as laying an IQ egg that might kill the draft goose . . . ." Jon Saraceno, *Who Knows if this Longhorn is Short on IQ,* USA TODAY, Feb. 28, 2006, *available at* http://www.usatoday.com/sports/columnist/saraceno/2006-02-28-saraceno-young_x.htm.

(ADEA) in *Smith v. City of Jackson, Mississippi*.[66] That case involved a group of plaintiff police officers that challenged the City's salary increases because they were more generous to officers under the age of forty.[67] The Supreme Court held that the plaintiffs' could not make a disparate impact claim because they did not identify a specific test, requirement, or practice that had an adverse impact on older workers.[68] Because no prima facie case could be made, the Court found the City's intention to raise salaries in order to become competitive with regional averages sufficiently related to the goal of retaining the police force and a reasonable factor other than age.[69]

The Ninth Circuit Court of Appeals addressed wage-based disparate impact in the context of sex discrimination in *American Federation of State, County, and Municipal Employees (AFSCME) v. State of Washington*.[70] In that case, state female employees in job categories of at least seventy percent female brought suit alleging Title VII sex discrimination in compensation.[71] The plaintiffs claimed that they were compensated at lower rates than employees in jobs where males predominate, although evidence existed that the jobs were of comparable worth.[72] The court held that the decision to base compensation on market forces "involves the assessment of a number of complex factors not easily ascertainable, an assessment too multifaceted to be appropriate for disparate impact analysis."[73] It further reasoned that a compensation system based on supply and demand is not the type of employment practice contemplated by *Griggs* because it does not constitute a single, specific employment practice that has a disparate impact.[74]

However, the notion of disparate impact in salary setting has not been fully explored by the courts—in fact, after thorough research I could only find one case directly on point. In *Donnelly v. Rhode Island Board of Governors for Higher Education*, certain female faculty members at the University of Rhode Island contended that the university's salary plan had a disparate impact on women's pay.[75] The plan had progressively higher minimum salaries for the humanities tier, natural sciences tier, and business tier.[76] The plaintiffs contended that because a greater percentage

---

[66] Smith v. City of Jackson, Miss., 544 U.S. 228, 233 (2005).
[67] *Id.* at 230.
[68] *Id.* at 241.
[69] *Id.* at 242.
[70] Am. Fed. of State, County, and Mun. Employees (AFSCME) v. State of Wash., 770 F.2d 1401, 1403 (9th Cir. 1985).
[71] *Id.*
[72] *Id.* at 1403–04.
[73] *Id.* at 1406.
[74] *Id.*
[75] Donnelly v. R.I. Bd. of Governors for Higher Educ., 110 F.3d 2, 4 (1st Cir. 1997).
[76] *Id.* at 2.

of women were in the lower paying tiers, this pay schedule disparately impacted their salaries.[77] The First Circuit Court of Appeals rejected their claim, finding that the faculty members' choice of academic field combined with the national market, rather that the plan itself, was responsible for the differences in salary between tiers.[78]

These cases, although relevant, are distinguishable from the NFL Draft context. Unlike *City of Jackson* and *AFSCME,* a specific employment practice, namely the use of the Wonderlic, is identifiable as creating a disparate impact on the salaries of minority football players. The test is administered to all potential draft picks and those who score lower on the test are, on average, drafted later and receive smaller initial contracts.[79] Also, the market forces defense is not appropriate because the NFL is a monopsonistic employer; there is no competitive market wage for football players because the NFL comprises *the* entire market. Hence, the ordinary reply of "we're just paying market wages" does not apply here. Because potential NFL players have no equally lucrative employment opportunities, the league cannot justify a competitive market defense.[80]

The NFL's use of the Wonderlic is unusual because it is a factor in determining an employee's salary. An employment test is typically used, as mentioned in *Griggs* above, to see if an individual is qualified to do the job[81]—salary is then based on the job itself. Because the NFL is such a unique employer it is not surprising that few cases are analogous. Intelligence test scores rarely determine wages; instead they are used to make hiring decisions. The evidence presented in this Note shows that test scores are used to determine wages in the NFL. A player that scores higher on the Wonderlic is, on average, drafted earlier and therefore receives a higher salary. Consequently, because blacks score lower than whites on the Wonderlic, the test has a disparate impact on minority players' salaries.

V. BUSINESS NECESSITY

Once a prima facie case is established, the NFL would have the burden of production and persuasion to show that the use of the Wonderlic is "job related for the position in question and consistent with business necessity. .

---

[77] *Id.* at 3–4.

[78] *Id.* at 5.

[79] *See supra* text accompanying notes 43–46, 51, 58–60.

[80] Ian Ayres contends that the market power defense, rooted in profitability, should be rejected by courts when used to extract super-competitive profits. Paying employees less than their marginal product because they have few alternatives should not be a business justification. Ian Ayres, *Market Power and Inequality: A Competitive Conduct Standard For Assessing When Disparate Impacts are Unjustified*, 95 CAL. L. REV. 669, 673–74 (2007).

[81] Griggs v. Duke Power Co., 401 U.S. 424, 431 (1971) (noting that an employment test must be shown to bear a demonstrable relationship to job-performance ability).

. . ."[82] Arguably, the interpretation of the phrase "business necessity" is still unclear, however, because in *Griggs*, the terms "job-related" and "business necessity" were used synonymously.[83] These terms can have entirely different levels of application, leading to different conclusions.[84] The Supreme Court attempted to clarify the standard in *Dothard v. Rawlinson* it stated that a challenged employment practice must be "necessary to safe and efficient job performance,"[85] but later relaxed the employer's required showing in *Wards Cove Packing Co. v. Atonio*, by calling for the challenged practice to "serve[], in a significant way, the legitimate employment goals of the employer."[86] The 1991 Civil Rights Act overruled that portion of *Wards Cove*, and returned to the "job related for the position in question and consistent with business necessity" standard,[87] but despite this, no universal definition exists.[88] For example, the Third Circuit Court of Appeals ruled that in order for a cutoff score on a test to be job-related it must equal the minimum qualifications necessary to do the job.[89]

It is undisputed that in some vocations a job applicant's Wonderlic score meets the business necessity standard because it is a valid predictor of job performance. For example, it has been used and upheld for jobs in a chemical manufacturing plant,[90] and is routinely used to select dental hygienists and some entry-level workers.[91] However, courts have found it not to be "job related" in the context of promotion within a paper mill,[92] or in hiring employees at a power plant.[93] What becomes apparent is that the Wonderlic's justification under business necessity is almost never certain, regardless of the employment position. Equal Employment Opportunity Commission (EEOC) personnel selection guidelines require statistical evidence that the "selection procedure is predictive of or significantly correlated with important elements of job performance."[94] This section presents evidence that these EEOC standards are not met because Wonderlic scores do not predict performance in the NFL.

---

[82] 42 U.S.C. § 2000e-2(k)(1)(A)(i) (2006).

[83] LINDEMANN & GROSSMAN, *supra* note 7, at 148.

[84] *Id.* at 148–52 (noting the federal appellate courts' different interpretations of the two terms).

[85] Dothard v. Rawlinson, 433 U.S. 321, 332 n.14 (1977).

[86] Wards Cove Packing Co. v. Atonio, 490 U.S. 642, 659 (1989).

[87] 42 U.S.C. § 2000e-2(k)(1)(A)(i) (2006).

[88] *See* Ayres, *supra* note 80, at 670 (stating that "a persuasive answer [to the business justification defense] has eluded both scholars and judges.").

[89] *See* Lanning v. S. Pa. Transp. Auth. (SEPTA), 181 F.3d 478, 489 (3d Cir. 1999) (noting that a "discriminatory cutoff score is impermissible unless shown to measure the minimum qualifications necessary for successful performance of the job in question.").

[90] Cormier v. P.P.G. Indus., 519 F. Supp. 211, 214, 255 (W.D. La. 1981).

[91] Rovell, *supra* note 11.

[92] Albemarle Paper Co. v. Moody, 422 U.S. 405, 435–36 (1975).

[93] Griggs v. Duke Power Co., 401 U.S. 424, 431–32 (1971).

[94] 29 C.F.R § 1607.5(B) (2008).

Analysis of the "business necessity" issue in the NFL context is challenging because there is no uniform way that all thirty-two teams "use" the test. Some teams use the information to guide selections in later rounds when less information is available on specific players.[95] Other teams apparently weigh the test more heavily the closer the player is to the ball, under the theory that those who regularly touch the ball have to make more strategic decisions.[96] While seemingly all coaches are skeptical of the test, almost all of them use it in some form while evaluating players.[97] Once the test is shown to have a disparate impact, if it is used, the scores must sufficiently relate to job performance.[98]

Not surprisingly, Wonderlic Inc., the company that administers the test, is convinced that its test meets "business necessity" standards primarily because of the amount of money involved in selecting a draft pick.[99] However, such reasoning is unlikely to satisfy any definition of the admittedly vague phrase "business necessity." Others claim that due to the complexities of the play calling in the modern NFL, a Wonderlic score can point to a player's aptitude to memorize an intricate playbook,[100] but this assertion would have to be proven by showing that a player's memorizing ability was significantly correlated with Wonderlic score.[101] Either way, at

---

[95] Rovell, *supra* note 11.

[96] *See* McCann, *supra* note 50, at 21 (noting that players who touch the ball more frequently in a game are required to use a higher level of thinking than those who play more physical, rudimentary, positions); *see also* Dougherty & Wyatt, *supra* note 48 (noting that former Baltimore Ravens coach Brian Billick expects quarterbacks to score higher on the Wonderlic than other positions based on their responsibilities on the field).

[97] Super Bowl champion coach Tony Dungy said, "The ability to win, delivering in the clutch cannot be measured with pen and paper . . . . We have had some really 'Wonderlic Smart' guys who turned out to be 'football dumb.'" *See* Jon Entine, Dark Thoughts, http://www.jonentine.com/articles/dark_thoughts_recon.htm (last visited Mar. 20, 2009). Nevertheless, Dungy still reviews every score. *Id.* Pittsburgh Steelers coach Mike Tomlin stated that "[p]ersonally, I've never been a Wonderlic guy. . . . It doesn't measure football intelligence. You don't know the background, the way guys have prepared for the test. You've got to go based on your interactions with people and what you see on tape." Corbett, *supra* note 5. *But see* Entine, *supra* (noting that former Minnesota Vikings coach Dennis Green said regarding the Wonderlic scores, "I never pay any attention to them. . . . I don't even look at the score. The only thing I'm concerned about is how the guy has performed of [sic] the field.").

[98] 42 U.S.C. § 2000e-2(k)(1)(A)(i) (2006).

[99] *See* Intelligence Testing, *supra* note 9 (noting that given the investment involved a draft pick, teams need all the information they can get). Michael Callans, President of Wonderlic Consulting stated that"[w]hether you are hiring a mailroom clerk or a CEO, a defensive lineman or a quarterback, intelligence is an accurate determiner of success . . . Smart people achieve more, they are better leaders, and they add greater value to the company." *Id.*

[100] Walker, *supra* note 56 (noting that "[i]f the coach calls 'zero type wing ride,' for instance, each player has to know instantly what to do, where his teammates will be going and how to adjust to the other team's behavior"). Tampa Bay Buccaneers general manager and Wonderlic proponent Bruce Allen, whose 2005 offensive line averaged Wonderlic scores above 30 (higher than the average attorney's score) said, "You need to 'get it' quick . . . We don't have a lot of patience in the NFL right now." *Id.*

[101] *See* 29 C.F.R. § 1607.5(B) (2008) (stating that a valid test should predict an important element of the job).

this stage the NFL would have the burden of showing a business justification for its use of the Wonderlic.[102] When most employers' selection practices are challenged, they present their own validation studies to the court.[103] So-called "criterion validation," which is most often used in evaluating tests, involves comparing test performance with the outcome the test is designed to predict.[104]

There are two types of criterion validation: predictive and concurrent.[105] Predictive validation would require the NFL to administer the Wonderlic at the combine, but not release the scores to the teams. The teams would then select their draft picks and after a certain interval of time, job performance would be evaluated. A comparison of predicted performance (based on the test score) with actual performance would be used to assess the test's validity.[106] If the test was not sufficiently correlated with actual performance it would be invalid.[107] Concurrent validation would require that the Wonderlic be administered to players already in the NFL who are representative of candidates "normally available in the relevant labor market for the job . . . in question."[108] These test scores would be compared with their current level of ability. This could be unnecessary because almost all players in the NFL have already taken the Wonderlic in preparation for the draft and a finding of job-relatedness could be based on those earlier scores. Moreover, scores on the test would not reflect work experience gained by the player already on the job, obfuscating the correlation between test scores and criterion.[109] In other words, a player's Wonderlic score might not increase corollary to his knowledge and expertise gained while playing in the league.

This Note contends that neither concurrent nor predictive validation would show that a football player's Wonderlic score predicts his future success in the NFL. McDonald P. Mirable examined the relationship between "intelligence" based on the Wonderlic and football ability in NFL rookie quarterbacks between 1989 and 2004.[110] Through empirical analysis he concluded that "there exists no statistically significant relationship between intelligence and quarterback performance at either the collegiate or professional level."[111] Mirabile concluded that if the Wonderlic had no relation to ability, the NFL would be better suited to

---

[102] 42 U.S.C. § 2000e-2(k)(1)(A)(i) (2006).
[103] LEX K. LARSON, EMPLOYMENT DISCRIMINATION § 27.01 (2d ed. 2006).
[104] *Id.* at § 27.01(1).
[105] *Id.*
[106] *See* 29 C.F.R. § 1607.14(C)(4).
[107] *See id.* § 1607.14(B)(5) (noting that a selection procedure is generally valid when performance on the procedure and performance on the measured criterion is statistically significant at the 0.05 level).
[108] *Id.* § 1607.14(B)(4).
[109] LARSON, *supra* note 103, at § 27.01(1)(b).
[110] Mirabile, *supra* note 14.
[111] *Id.*

spend their time and resources in some other way.[112]

In order to evaluate the Wonderlic's ability to predict a quarterback's success over their career, I compiled the Wonderlic scores and corresponding NFL passer rating of 116 quarterbacks over the past four decades (Figure 5).[113] The passer rating is a statistic based on four categories—percentage of completions per attempt, average yards gained per attempt, percentage of touchdown passes per attempt, and percentage of interceptions per attempt.[114] A passer rating can range from 0.0 to 158.3.[115] The league regular season average in 2008 for qualified quarterbacks (defined as quarterbacks who threw at least fourteen passes per game played) was 84.2.[116] The statistic has its weaknesses; it does not measure intangibles such as leadership, play calling, or meaningful touchdowns, but it has been consistently used since 1973.[117]

Although a passer rating can be determined after a quarterback has thrown one attempt,[118] only quarterbacks who had attempted at least twenty-five passes were included in my analysis to give a reasonable sample size. If "intelligence" as measured by the Wonderlic was necessary to be a successful NFL quarterback, one would expect a positive correlation to exist between Wonderlic score and passer rating. In other words, using Figure 6 with the X-axis representing Wonderlic score and the Y-axis representing NFL passer rating, there should be a positive relationship between the two. It is obvious by looking at Figure 6 that such a relationship is not present. Instead of a positive relationship, the regression line has a slightly downward (negative) trend, although the coefficient is not significantly different from zero. The linear equation for

---

[112] *Id.*

[113] *See supra* notes 55–59 and accompanying text.

[114] NFL Quarterback Rating Formula, http://www.nfl.com/help/quarterbackratingformula (last visited Apr. 7, 2009).

[115] Richard Sandomir, *The N.F.L.'s Passer Rating, Arcane and Misunderstood*, N.Y. TIMES, Jan. 14, 2004, at D1, *available at* LEXIS, News Library, NYT File.

[116] Draft Order, http://www.nfl.com/stats/player (follow "complete list" hyperlink; then follow "regular season" hyperlink) (last visited Mar. 20, 2009). To compile the above statistic, I averaged the passer rating for every qualified quarterback.

[117] NFL Quarterback Rating Formula, *supra* note 114; *see* Sandomir, *supra* note 115 (noting also that the statistic is weighted against freewheeling quarterbacks as it rewards a high percentage of completions and yards per attempt while punishing interceptions).

[118] NFL Quarterback Rating Formula, *supra* note 114.

Figure 6 is

QB Rating = 75.65 - .12 x Wonderlic Score

(t-stat) (t-stat)

This means that with each point scored on the Wonderlic, a quarterback's rating actually *decreases* by .12. The p-value for this analysis was 0.491, leading to the conclusion that no statistically significant relationship between Wonderlic score and NFL passer rating exists; therefore its use would likely not meet the "business necessity" standard. This evidence, while certainly not dispositive, appears to show that a player's Wonderlic score has nothing to do with their ability to play quarterback. And, if the Wonderlic cannot predict success for a quarterback then, *a fortiori*, it should not predict success for other, less intellectually demanding football positions.

Others have also examined the relationship between Wonderlic score and quarterback performance, reaching different conclusions. The employee testing firm Criteria Corp. found a correlation between test score and passing yards once a quarterback had thrown for over 1000 yards.[119] The authors of this study originally noted that no association between Wonderlic and passing yards could be found for the sixty-one quarterbacks they evaluated that were drafted from 2000–2004.[120] However, once they included only quarterbacks who had passed for over 1000 yards, the authors found a correlation coefficient (otherwise known as *r*)[121] of .51—showing strong positive correlation between aptitude and performance.[122]

While there is perhaps some validity to this study, using passing yards with a 1000 yard cutoff may be misleading. In the 2008 regular season, the average NFL team threw for 211.6 yards per game;[123] throwing over 1000 yards in your career means that you have played in multiple games. Using passing yards as a metric is an accurate way to measure long term skill and durability, but conversely it brings in variables including resistance to injury and a team's propensity to throw the ball, something that varies throughout the league and something that the Wonderlic is not designed to predict.

Admittedly, passer rating may be an inaccurate measure in the case of a quarterback who has one good game only to be injured and never play

---

[119] Criteria Employees Testing Blog, *The Wonderlic as a Predictor of Performance in the NFL*, http://blog.criteriacorp.com/blog/bid/4920/The-Wonderlic-as-a-Predictor-of-Performance-in-the-NFL

[120] *Id.*

[121] The correlation coefficient, or *r*, is a common measure of the correlation between two variables. Ranging from -1 to +1, it measures the strength of linear dependence. *See* FREUND & SIMON, *supra* note 58, at 470–71.

[122] Criteria Employees Testing Blog, *supra* note 119.

[123] To reach this number I averaged the passing yards per game average of all thirty-two NFL teams. The statistics I used are available at http://www.teamrankings.com/nfl/stats/?cat=team&pan=7&conf=0.

again, or for a quarterback who is benched upon the return of the team's regular starter. Yet in today's NFL the demand for decent quarterback play is so high that even one well played game may be enough to lead to future opportunity. Because quarterback rating is normalized per passing attempt and is less affected by a team's offensive scheme, it is a better gauge of a quarterback's ability for this study than passing yards. It provides a standardized method of comparing Wonderlic performance to in-game performance. As the above analysis shows, no statistically significant relationship between Wonderlic score and quarterback performance can be found, thus the test would be unlikely to meet "business necessity."

VI. THE CASE FOR AN ALTERNATIVE TEST

In the unlikely event that NFL's use of the Wonderlic test were found to be justified under the "business necessity" standard, plaintiffs would still be entitled to offer an alternative selection practice that meets the employer's legitimate interests while having a less discriminatory effect.[124] It must be noted that courts are generally apprehensive to mandate alternative hiring practices because they are "less competent than employers to restructure business practices . . . ."[125] Also, the burden rests squarely on the plaintiffs to demonstrate that this alternative practice exists and that it would have less of a disparate impact on the protected class.[126]

Despite the plaintiff carrying this burden, there is some evidence of a responsibility on the part of an employee who uses a selection procedure to investigate alternatives and choose the one with the least discriminatory impact.[127] The Uniform Guidelines on Employee Selection Procedures state that when a validity study is called for, the employer should include "an investigation of suitable alternative selection procedures and suitable alternative methods of using the selection procedure which has as little adverse impact as possible . . . ."[128] Once the employer has made a "reasonable effort to become aware of such alternative procedures and validity has been demonstrated in accord with these guidelines, the use of the test or other selection procedure may continue until such time as it should reasonably be reviewed for currency."[129] Although the Guidelines themselves cannot impose an obligation on the employer to investigate alternatives, they may not be ignored by the employer either.[130] Courts

---

[124] *See* 42 U.S.C. § 2000e-2(k)(1)(A)(ii)(2006).
[125] Furnco Constr. Corp. v. Waters, 438 U.S. 567, 578 (1978).
[126] LARSON, *supra* note 103, at § 24.01.
[127] *Id.* at § 24.02.
[128] 29 C.F.R. § 1607.3(B).
[129] *Id.*
[130] LARSON, supra note 103, at § 24.02.

have generally given little weight to the Guidelines' directive and alternative testing techniques proposed by plaintiffs rarely prevail,[131] because they are usually unable to show that their plan would have a less discriminatory effect or that it would serve the employer's purposes.[132]

In *Murphy v. Derwinski,*[133] the plaintiff was able to show an alternative selection practice that the court found to meet the employer's goals. In that case, the plaintiff applied for employment as a Roman Catholic chaplain at a VA hospital in Colorado but was rejected because she was not ordained and did not have an ecclesiastical endorsement—a Veteran Administration requirement.[134] The Tenth Circuit Court of Appeals held that the requirement had a disparate impact on women because the plaintiff, as a woman, was unable to be ordained a priest.[135]

Although the court found that the decision could be justified by business necessity, it ruled that requiring only the ecclesiastical endorsement standard would ensure that the hospital's patients receive effective spiritual service while simultaneously not having a disparate impact on women.[136] In support of their conclusion, the court noted that women serve as Roman Catholic chaplains in non-VA hospitals "without disruption of service.[137]

Building on the case law mentioned above, this Note contends the NFL should abandon its use of the Wonderlic test and replace it with a football-specific intelligence test that (like the ecclesiastical endorsement requirement) could equally serve the NFL's interest in evaluating a player's mental acumen. This Note does not dispute the NFL's need to measure intelligence in order to provide more information to teams before they make their draft choices. Players spend hours in meetings every day and must know their responsibilities inside out.[138] However, a player's answer on a question such as "[p]aper sells for 21 cents per pad. What will four pads cost?" is not optimally designed to determine how well a player will receive instruction from coaches, associate with teammates, or learn a playbook.[139] Legal arguments aside, a football-specific test would provide more useful information for teams in assessing players. There is evidence that some teams have already implemented such tests into their evaluation

---

[131] *Id.* at § 24.03.
[132] *Id.*
[133] Murphy v. Derwinski, 990 F.2d 540 (10th Cir. 1993).
[134] *Id.* at 542.
[135] *Id.* at 544–45.
[136] *Id.* at 545–46.
[137] *Id* at 545.
[138] Elizabeth Merill, *In NFL, the Playbook is Sacred*, ESPN, Aug. 29, 2007, http://sports.espn.go.com/nfl/preview07/news/story?id=2973338.
[139] *See* So, how do you score?, http://espn.go.com/page2/s/closer/020228test.html (last visited Mar. 20, 2009) (providing fifteen sample Wonderlic questions).

process, although few details are known as to what is tested.[140]

The football-specific intelligence test this Note recommends would require players at the combine to study, based on their position, plays in a playbook upon which they would later be tested. Each team places great emphasis on players understanding their respective playbooks—some can be 800 pages in length and require two years study for full assimilation.[141] Because agents, in preparation for the Wonderlic, frequently have their players take a practice version of the test before the combine,[142] in similar fashion, a prospect under this Note's recommendation would receive, before the combine, a basic playbook to study in order to prepare for the test.

The test could be administered in twelve minutes, identical to the Wonderlic, but would assess players' abilities to match specific names of plays with their corresponding diagram. It could also give a name of a play and ask a player to choose a multiple choice answer that best describes the basic purpose of the play—whether it be a screen pass, play action, draw, or any other formation. The test would evaluate a player's ability to think quickly, but in relation to information that they will actually have to learn in order to play in the NFL.

Unfortunately, no promise can be made that such a test would not continue to have a disparate impact on minority football players. However, the test would be more likely to meet the business necessity standard and would represent a good-faith attempt to put all potential draftees on equal footing. Despite the courts' unwillingness to make business decisions for employers, it is entirely possible that a plaintiff could show that a football-specific intelligence would be justified under "business necessity," while possibly not discriminating.

The NFL should not wait for legal challenge to begin investigating different testing methods, but should take the initiative themselves, under the Guidelines mentioned above, to develop a test that can accurately predict football success by measuring the skills necessary to play the game. The league should begin this process through an employment study known as job analysis. Job analysis is a survey that establishes whether a challenged employment practice sufficiently predicts success on the job by determining what actually is required by the job.[143] It is often conducted

---

[140] *See* Corbett, *supra* note 5 (noting that New York Jets general manager Mike Tannenbaum and former head coach Eric Mangini ask prospects game-related football questions in quick fashion, as well as fire questions at them while watching college game film); *see also* Dougherty & Wyatt, *supra* note 49 (stating that Houston Texans general manager Charley Casserly uses another test that "relates more to football.").

[141] Merill, *supra* note 138.

[142] *See* Dougherty & Wyatt, *supra* note 48 (stating that because of the emphasis placed on the Wonderlic, most agents ensure their clients take practice versions of the test).

[143] LARSON, *supra* note 103, at § 27.05[1].

through interviews with employers and employees and observing employees while on the job.[144] The league could conduct interviews with teams to determine how football "intelligence" is best measured and from that information develop a league-wide test administered at the combine. Such a test might be slightly different from the one proposed above, but having standardized results that are applicable to each prospect would serve the NFL's interest.

Finally, the question of why a potential NFL player has not yet challenged the use of the Wonderlic must be addressed. I contend that a player who is ultimately drafted in the first four rounds (out of a total of seven), although perhaps negatively affected by the Wonderlic, is not likely to contest its use even if they have a valid prima facie case. These players usually have some form of commitment from the team that drafted them and the negative impact a lawsuit would have on their reputation might not make it a prudent career move. However, a minority player who scored poorly on the Wonderlic and was drafted in the late rounds or not drafted at all might be more willing to challenge its use. The reasoning for this is simple: the employment substitutes available to a player if they do not make the NFL, especially if they did not complete their degree, have drastically different levels of compensation. Take an average player for example: Leroy Harris, selected in the middle of the fourth round by the Tennessee Titans in 2007, received a four-year $2.11 million dollar contract.[145] If the next best alternative to an NFL career is a salary of $15,000 a year, a potential draft pick who was denied a spot in the NFL because of his low Wonderlic score could find litigation appealing and may even be wise to pursue it.

VII. CONCLUSION

Because of the Wonderlic's racially-different results, the test almost certainly has a disparate impact on minority players' salary opportunities in the NFL. Potential draft picks with higher Wonderlic scores are drafted earlier, increasing their initial professional salaries. Black quarterbacks who, on average, score lower on the test may be financially punished by it. As shown above, the test does not meet the "business necessity" standard because it cannot reliably predict job performance. In fact, it appears that the higher a quarterback scores on the Wonderlic the worse they perform on the field, and therefore it cannot be justified as representing the minimum intelligence standards necessary for the job. This Note argues that the league should abandon its use of the Wonderlic and instead

---

[144] *Id.*

[145] Leroy Harris Player Page http://www.rotoworld.com/Content/playerpages/player_contract.aspx?sport=NFL&id=4311; SI.com 2007 NFL Draft, Round 4 http://sportsillustrated.cnn.com/football/2007/draft/breakdowns/by_round/4.html (last visited Apr. 16, 2009).

administer a football-related "intelligence" test to potential players. This test might not have a disparate impact on minorities and would more likely be justifiable under business necessity. Perhaps most importantly it could provide the most relevant information in evaluating a football player's ability.

APPENDIX

Figure 1: 1999–2006 Quarterback Score and Draft Position

Drew Henson	42	192
Alex Smith	40	1
Eli Manning	39	1
Todd Husak	39	202
Charlie Frye	38	67
Craig Krenzel	38	148
Ryan Fitzpatrick	38	250
Omar Jacobs	36	164
Aaron Rodgers	35	24
Drew Stanton	35	43
J.T. O'Sullivan	35	186
Kellen Clemens	35	49
Matt Leinart	35	10
Charlie Whitehurst	33	81
Tom Brady	33	199
Giovanni Carmazzi	32	65
Jesse Palmer	32	125
Joey Harrington	32	3
Patrick Ramsey	32	32
Sage Rosenfels	32	109
Andrew Walter	31	69
Craig Nall	31	164
Dave Ragone	31	88
J.P. Losman	31	22
Kliff Kingsbury	31	201
Kurt Kittner	31	158
Matt Schaub	31	90
Ingle Martin	30	148
John Beck	30	40
Josh Heupel	30	177
Josh McCown	30	81
Matt Mauck	30	225

Philip Rivers	30	4
Quincy Carter	30	8
Stefan LeFors	30	121
Brady Quinn	29	22
Chris Weinke	29	106
Mark Bulger	29	168
Rex Grossman	29	22
Brian St. Pierre	28	163
Brooks Bollinger	28	200
Cade McNown	28	12
Drew Brees	28	32
Kevin Kolb	28	36
Seth Burford	28	216
Josh Booty	27	162
Kyle Boller	27	19
Mike McMahon	27	149
Tim Rattay	27	212
Adrian McPherson	26	152
Akili Smith	26	3
Carson Palmer	26	1
Dan Orlovsky	26	145
Jay Cutler	26	11
Kyle Orton	26	106
Ben Roethlisberger	25	11
Bradlee Van Pelt	25	250
Brock Huard	25	77
Byron Leftwich	25	7
Casey Bramlett	25	218
Chad Pennington	25	18
Joe Germaine	25	101
Josh Harris	25	187
Ken Dorsey	25	241
Randy Fasani	25	137
Shaun King	25	50
Spergon Wynn	25	183

David Carr	24	1
Jamarcus Russell	24	1
John Navarre	24	202
Luke McCown	24	106
Marques Tuiasosopo	24	59
James Killian	23	229
Jason Campbell	23	25
Jeff Smoker	23	201
Jordan Palmer	23	205
B.J. Symons	22	248
Chris Simms	22	97
Tim Couch	22	1
Jeff Rowe	21	151
Scott Covington	21	245
Tyler Thigpen	21	217
Cody Pickett	20	217
Corey Jenkins	20	181
David Greene	20	85
Michael Vick	20	1
A.J. Feeley	19	155
Bruce Gradkowski	19	194
Derek Anderson	19	213
D.J. Shockley	19	223
Tavaris Jackson	19	33
Daunte Culpepper	18	11
Jarious Jackson	18	214
Joe Hamilton	18	234
Aaron Brooks	17	131
Rohan Davey	17	117
Chris Redman	16	75
Vince Young	16	3
Jim Sorgi	15	193
David Garrard	14	108
Donovan McNabb	14	2
Seneca Wallace	14	110

Tee Martin	11	163
Michael Bishop	10	277

Figure 2: Wonderlic Score by Race

	Average Score	Standard Deviation	# of Observations
White	28.15	5.70	78
Black	20.4	6.08	25

Figure 3: 1999-2006 Quarterback Wonderlic Score vs. NFL Draft Position

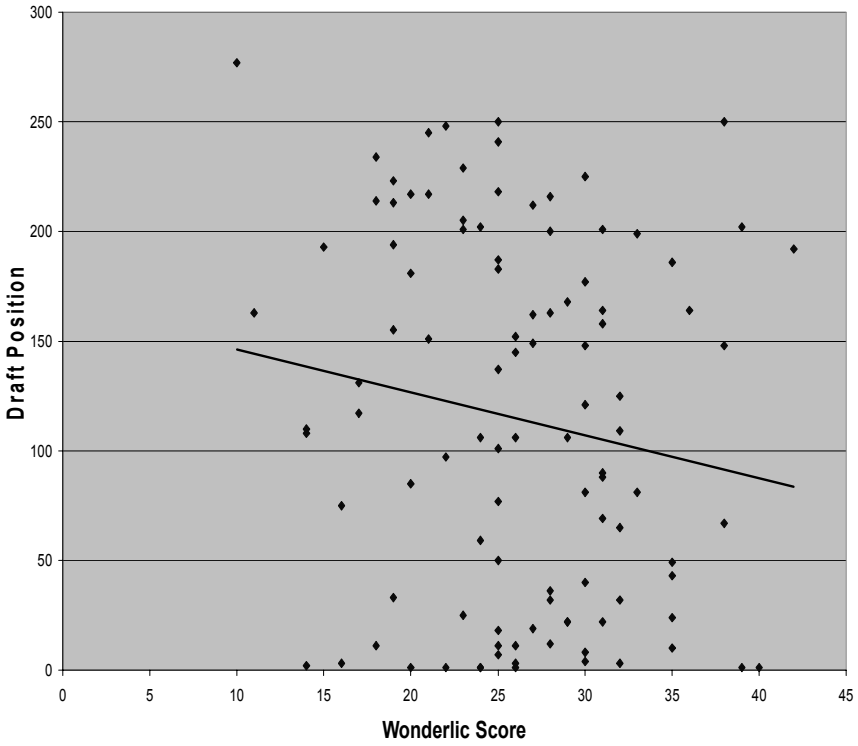

Figure 4: Average Draft Position When Wonderlic Score is One Standard Deviation Above and Below the Mean

	One Std. Deviation Above	One Std. Deviation Below
Wonderlic Score	≥33	≤19
Draft Postion	107.8	136.4

Figure 5: Quarterback Wonderlic Score and NFL Passer Rating

Hugh Millen	41	73.5
Alex Smith	40	63.5
Eli Manning	39	73.4
Brian Griese	39	83.6
Charlie Frye	38	70.1
Ryan Fitzpatrick	38	58.2
Craig Krenzel	38	52.5
Tony Romo	37	96.5
Drew Bledsoe	36	77.1
Anthony Dilweg	36	72.3
Jason Garrett	36	83.2
Matt Leinart	35	71.2
Kellen Clemens	35	60.7
Aaron Rodgers	35	73.3
Steve Stenstrom	35	62.5
Tom Brady	33	92.9
Steve Young	33	96.8
Joey Harrington	32	69.4
Patrick Ramsey	32	74.8
Sage Rosenfels	32	82
Jesse Palmer	32	59.8
Andrew Walter	31	56.3
J.P. Losman	31	77.3
Matt Schaub	31	80.7
Rick Mirer	31	63.5
John Beck	30	62
Philip Rivers	30	86.6
Josh McCown	30	71.6
Quincy Carter	30	71.7
Kerry Collins	30	73.3
David Klingler	30	65.1
Rex Grossman	29	70.9
Chris Weinke	29	62.2
Mark Bulger	29	88.1
Matt Hasselbeck	29	86.2
Bill Musgrave	29	71
Troy Aikman	29	81.6
John Elway	29	79.9
B. Bollinger	28	75.2
Drew Brees	28	87.9

Doug Johnson	28	69.4
Cade McNown	28	67.7
Peyton Manning	28	94.7
Kyle Boller	27	71.9
Mike McMahon	27	55.1
Tim Rattay	27	81.9
Ryan Leaf	27	50
Kelly Stouffer	27	54.5
Rich Gannon	27	84.7
Jay Cutler	26	88.2
Kyle Orton	26	62.2
Jared Lorenzen	26	58.3
Carson Palmer	26	90.1
Akili Smith	26	52.8
Tony Banks	26	72.4
Rob Johnson	26	83.6
Rodney Peete	26	73.3
Ben Roethlisberger	25	92.5
Bradlee Van Pelt	25	39.6
Byron Leftwich	25	79.7
Ken Dorsey	25	63.5
Randy Fasani	25	42.8
Shaun Hilll	25	101.3
Chad Pennington	25	88.9
Spergon Wynn	25	39.5
Shaun King	25	73.4
Brock Huard	25	80.3
Danny Wuerffel	25	56.4
Brodie Croyle	24	65.8
Luke McCown	24	75.5
John Navarre	24	43.9
David Carr	24	74.4
Marques Tuiasosopo	24	48
Billy J. Tolliver	24	67.7
Jamarcus Russell	24	55.9
Jason Campbell	23	77.3
Tim Hasselbeck	23	63.6
Chris Simms	22	71.2
Tim Couch	22	75.1
Trent Dilfer	22	70.2

Mark Brunell	22	84.2
Brett Favre	22	85.7
Chris Miller	22	74.9
Don Majkowski	21	72.9
Cody Pickett	20	16.4
Michael Vick	20	75.7
Tavaris Jackson	19	69
Bruce Gradkowski	19	65
Derek Anderson	19	78.9
AJ Feeley	19	69.6
Cleo Lemon	18	72.2
Daunte Culpepper	18	89.9
Ray Lucas	18	74.3
Aaron Brooks	17	78.5
Jeff Blake	17	78
Vinny Testaverde	17	75
Vince Young	16	69
Chris Redman	16	79.5
Anthony Wright	16	66.3
Heath Shuler	16	54.3
Elvis Grbac	16	79.6
Jim Sorgi	15	90.4
Steve McNair	15	82.8
Randall Cunningham	15	81.5
Jim Kelly	15	84.4
Dan Marino	15	86.4
Terry Bradshaw	15	70.9
Seneca Wallace	14	78.9
David Garrard	14	87.7
Donovan McNabb	14	85.8
Charlie Batch	14	77.9
Kordell Stewart	14	70.7
Neil O'Donnell	13	81.8
Jeff George	10	80.4
Vince Evans	8	63

Figure 6: Wonderlic Score vs. NFL Passer Rating