5-2014

Physical Ability Testing: A Review of Court Cases 1992-2014

Joseph Westlin
Western Kentucky University, joseph.westlin991@topper.wku.edu

PHYSICAL ABILITY TESTING: A REVIEW OF COURT CASES 1992-2014

A Thesis
Presented To
The Faculty of the Department of Psychology
Western Kentucky University
Bowling Green, KY

In Partial Fulfillment
Of the Requirements for the Degree
Master of Arts

By
Joseph Westlin

May 2014

PHYSICAL ABILITY TESTING: A REVIEW OF COURT CASES 1992-2014

Date Recommended May 5, 2014

Elizabeth L. Shoenfelt, Director of Thesis

Reagan Brown

Amber Schroeder

5-14-14

Dean, Graduate Studies and Research Date

I would like to dedicate this to my future wife, Gabby Manny. Her love and support throughout the seemingly never ending process of writing a thesis was helpful beyond measure.

ACKNOWLEDGEMENTS

I would like to acknowledge and thank the member of my thesis committee, Betsy Shoenfelt, Reagan Brown, and Amber Schroeder, for their consistent willingness to help me and teach me throughout the course of the thesis process. I would especially like to thank Betsy Shoenfelt for her tremendous attention to detail and unending patience as I struggled through the process of writing my first document of this magnitude. I would also like to thank my parents, Dave and Lory Westlin, for providing me the ability and opportunity to pursue the highest levels of education I seek.

TABLE OF CONTENTS

PHYSICAL ABILITY TESTING: A REVIEW OF COURT CASES 1992-2014

Joseph Westlin May 2014 41 Pages

Directed by: Dr. Elizabeth L. Shoenfelt, Dr. Reagan D. Brown, and Dr. Amber N. Schroeder

Department of Psychology Western Kentucky University

Selecting employees for hire and promotion is one of the most essential functions of an organization. Many companies that have positions which contain a physical component rely on physical ability testing as part of their selection procedure. The establishment of both the Civil Rights Act and the Americans with Disabilities Act (ADA) had a profound impact on the manner in which selection testing may legally be conducted (Gutman, Koppes, & Vodanovich, 2011). The current study sought to analyze court cases involving physical ability testing. Results revealed that pure ability tests did not significantly differ from work sample tests with regard to whether court cases found for the plaintiff or defendant. Additionally, rulings did not significantly differ in ruling in favor of the plaintiff or defendant with regard to whether the position in question involved public safety. Finally, the ADA related cases did not significantly differ in their rulings in favor of the plaintiff or defendant after the 2011 modifications to the interpretation of disabled, as compared to before 2011. Future research should focus on the difference between court rulings involving physical ability tests in comparison to other forms of testing such as cognitive tests, and further investigate the role of the ADA in physical ability testing.

Introduction

Physical ability testing has been used in the selection of many positions ranging from police officers (Hoover, 1992), to NFL players (Lyons, Hoffman, Michel, & Williams, 2011), to Special Forces operatives personnel in the military (Picano, Williams, & Roland, 2006). Physical ability tests are part of a broader category of procedures designed to assess the ability of job applicants in order to select the most proficient applicants to be hired or promoted. These selection tests carry with them a legal necessity on the part of the employer to ensure that the test is both accurate in predicting job performance and, if it disparately impacts a legally protected group, that the test is job related and a business necessity (EEOC Uniform Guidelines, 1978).

The discussion of the legal nature of physical ability testing must begin by examining the impact that Title VII of the 1964, 1972, and 1991 Civil Rights Acts have had on selection, as well as how the 1990 Americans with Disabilities Act (ADA) impacted selection procedures. From here, an introduction to the various types of physical ability testing is provided, followed by an investigation of the various forms of validity and how they impact court rulings. Finally, a brief overview of the study at hand is provided. This study seeks to provide clarity on how court cases have ruled on physical ability testing since the Civil Rights Act of 1991 (CRA 1991). The study further seeks to analyze the court cases that have surrounded physical ability testing, by expanding on previous analysis performed in 2006 (Starling). Court cases involving physical ability testing will be coded and analyzed to provide a broader understanding of the patterns and practices of judicial rulings.

History of Title VII

Before 1964, many police departments based their physical requirements solely on the height and weight of applicants (Maher, 1984). Today it is apparent that this practice would be both discriminatory and subject to legal action. The Civil Rights Act of 1964 (CRA 64) drastically impacted hiring procedures for many businesses. The primary implication of this act from a selection standpoint was that it clarified which groups were protected from discriminatory practices. These groups are sex, race, color, religion, and national origin. Intentional discrimination was made illegal by this act. Title VII prevented discriminatory intent of an employer but did not clarify whether unintentional discrimination could be legally defensible (Hollar, 2000).

The *Griggs v. Duke Power Company* (1971) decision was a landmark ruling which clarified the ever-growing role companies had in preventing discriminatory practices. In this suit, the company was using high school diplomas as a minimum requirement for consideration for hire. Because this standard was found to have a disparate impact on black applicants, it was deemed illegal (Gutman, Koppes, & Vadonovich, 2011). Thus, the court found that if a hiring practice leads to a disproportionate failure to hire members of any protected group, an equally valid test with less adverse impact may be required. However, the seemingly more important finding from this case was what the Supreme Court stated about future cases. The judges noted that disparate impact should be deemed illegal only when there is no evidence that the hiring procedure is linked to job performance (Hollar, 2000). Therefore, disparate impact is legally defensible if test performance is statistically related to job performance (i.e., it has criterion-related validity).

Several similar cases appeared from 1971 to 1991; their results seemed to blur the burden of proof between the company and the employee. Once disparate impact had been demonstrated, the rulings differed with regard to which party was responsible for the burden of proof in relating the hiring practice to job performance. Hollar (2000) stated that, in general, the rulings began to favor the employer, placing the burden on the plaintiff to prove that there was no substantial tie between the hiring practice and job performance. *Wards Cove Packing Co. v. Atonio* (1989) in particular muddied the manner in which these cases were handled by courts, in that this case seemingly shifted the burden of proof onto the plaintiff to demonstrate a lack of necessity for the hiring procedure.

The Civil Rights Act of 1991 (CRA-91) provided distinct clarification as to how these disparate impact claims would be dealt with in the future. The CRA-91 placed the burden of demonstrating business necessity and job relevance on the employer in instances where disparate impact has been proven. The legality of a test is tried in three different phases (Gutman et al., 2011). First, the plaintiff must demonstrate that the test results in adverse impact for a protected group. Second, the defendant then must demonstrate the job relatedness and business necessity of the test. Finally, the plaintiff must show that there is the potential for an alternate test to have equivalent validity while reducing adverse impact. Title VII has an immense impact on how physical ability testing is treated in a legal setting. The Americans with Disabilities Act further added to our understanding of how physical ability testing may be treated in court.

The Americans with Disabilities Act

In addition to the protected groups mentioned in the CRA, the Americans with Disabilities Act created a distinct and different protection for those who have either physical or mental disabilities that interfere with a major life activity (Gutman et al., 2011). The 2008 Americans with Disabilities Amendment Act (ADAAA) broadened the definition of disability to include a lengthier list of major life activities which may be impaired due to disability (Gutman et al., 2011). In a 2011 article, Gutman stated that the definition of disability continues to broaden, and that the number of lawsuits won through ADA claims may grow in the future. Additionally, it should be noted that the ADA requires that testing that may be considered "medical" testing is only to be performed after a conditional job offer is made to an applicant (Gutman et al., 2011). Thus, any test that is medical in nature may not be legally defensible in court if it is completed before a conditional offer. The legal precedents surrounding physical ability testing provide a basis for how this form of testing may be legally handled as a whole; however, the specific types of physical ability testing have different pros and cons in a legal context.

Types of Physical Ability Tests

There are a large number of methods by which physical ability can be tested. However, these methods of testing can be grouped into three distinct categories. These categories have been adopted from Hoover (1992), who broke down all types of physical ability testing into what he called job simulations, physical agility tests, and physical norm testing.

Job simulations mimic the task requirements of the position into which applicants intend to be hired. It should be noted that job simulations may also be called work sample

tests. In the coding for this current study, this type of test is referred to as work sample tests. To accurately represent the job, these simulations should contain two essential forms of fidelity. Physical fidelity addresses how similarly the physical act performed in the test replicates the physical act performed on the job (e.g., scaling a wall to mimic scaling a fence; Goldstein, & Ford, 2002). Additionally, physical ability tests should have psychological fidelity. That is, the test should mimic the psychological demands of the conditions under which one would be asked to perform the job (Goldstein, & Ford, 2002). The primary advantage of a job simulation is that it often provides the best indication of how applicants will physically perform the duties on the actual job (Hoover, 1992) because applicants are performing the actions that will be required of them on the job after being hired. The second advantage of job simulations is that they are more likely to withstand a lawsuit on the basis of content validity (i.e., the test accurately represents facets of job) because the test simulates the actual job. Likewise, because the job simulation attempts to replicate the activities of the job, applicants are less likely to view the simulation as discriminatory because the test *looks* like the job (i.e., face validity; Ryan, Greguras, & Ployhart, 1996). However, there are also disadvantages to this type of testing. It can be unwieldy, expensive, and sometimes impractical (Hoover, 1992). For example, in assessing the potential of a job applicant to extinguish a fire, the creation of an entire scenario that simulates a fire would be beneficial to employers who are hiring firefighters; but at the same time, it would be remarkably expensive, difficult, and potentially hazardous.

The second form of physical ability testing is physical agility testing. Hoover (1992) identified agility tests as containing sets of specific exercises designed to

determine the physical abilities of performers. It should be noted that in common vernacular in Industrial Organizational Psychology, physical ability testing and physical agility testing are synonymous. Thus, it is important to indicate that physical agility testing, as used by Hoover, refers to tests that involve assessing pure ability rather than simply any form of physical ability testing. Therefore, in this paper, this type of test will be referred to as pure ability tests. Hough, Oswald, and Ployhart (2001) identified several examples of pure ability tests that include muscular strength, cardiovascular endurance, and flexibility. The primary advantages to this style of physical ability testing are that they are relatively inexpensive, convenient, and, more importantly, they do not focus on specific techniques which may be taught within a course of training (Hoover 1992). However, Hough et al. noted that in every single pure ability test performed in their study, women scored (often significantly) lower than their male counterparts. Due to the biological differences of men and women, men often score higher than females on pure ability tests, thus leading to adverse impact in hiring decisions. This difference between sexes inherently lends itself to a higher frequency of litigation.

One way to reduce this adverse impact is to use norm testing. In Industrial Organizational Psychology, what Hoover (1992) referred to as norm testing is called subgroup norming. Accordingly, the term subgroup norming will be used in this paper. Subgroup norming adjusts the standard by which individual scores are interpreted. Subgroup norming compares individual applicant performance against the norm for their respective subgroup (Hoover, 1992). Individual scores are compared to the norms within this subgroup and each applicant receives a percentile rank within their subgroup which is used as their test score. This serves to reduce the impact that gender (or another

protected group characteristic) has in the selection process, but also raises questions as to what cutoff scores should be used to determine who gets hired. If a police officer applicant needs to score in the 80th percentile to be considered for hire, a female applicant whose *actual* performance falls far below a male at the 80th percentile may not have the genuine ability to successfully perform the job. However, this subgroup norming would lead to fewer instances of adverse impact because applicants are compared only to the norms of their subgroup rather than to one another. Subgroup norming was rendered illegal by CRA-91. Subgroup norming may only be used in instances when there is empirical support for moderation by group membership (Gutman et al., 2011) Thus, each of the types of physical ability tests has merits and pitfalls. One of the primary ways in which we interpret the results of varying physical ability tests is through the validity evidence for that test and how the evidence would impact the test's legal defensibility.

Validity

The concept of validity has been described as several types of unique validities, or as one uniform concept. Recently, it has been suggested that validity should be defined as one overarching concept including all information regarding a given test that contributes to the evidence of validity for that test (McDaniel, Kepes, & Banks, 2011). However, in the instances of legal action, courts still lean heavily on the three types of validity presented in the Uniform Guidelines: content, criterion, and construct. According to the EEOC Uniform Guidelines (1978), content validity refers to the extent to which the test reflects all facets of a given job. Criterion validity refers to the how well the test predicts job performance. Construct validity equates to how well a test measures a characteristic of applicants which should lead to higher levels of performance on the job. These types

of validity constitute avenues to legal defensibility for a company whose test is under scrutiny.

In addition to the various types of validity, court cases are sometimes determined by how appropriately a cutoff score has been set. Cutoff scores set a minimum score an applicant must achieve to be considered for hire. For example, if a cutoff score of a seven minute mile was set for fire-fighters, those who ran a mile in longer than seven minutes would not meet the cutoff and, therefore, would not be hired.

Many court cases surrounding physical ability testing focus primarily on the content validity of the selection procedure. Content validity refers to the representativeness of test content relative to the job itself. That is, how accurately the test reflects the job for which it is being used to make selection decisions (Gutman et al., 2011). Williams, Schaffer and Ellis (2013) stated that, on average, employers won only 10 % of the court cases in which they were defending their selection tests. Because of this infrequency of successful defense, coupled with the importance of content valid tests in physical ability testing, the litigation surrounding physical ability testing is substantial and relevant to selection as a whole.

The importance of content validity cannot be overstated. Sackett and Lieven (2008) noted that physical ability tests fare well in terms of job relevance, yet generally produce adverse impact towards female applicants. Physical ability tests are used extensively for positions of public safety (e.g. police officers, fire fighters, military service, etc.). However, because physical ability tests are often used in positions that require a level of physicality to protect public safety, there is an important balance between validity and the avoidance of disparate impact. It is important to place adequate

weight on both the validity of the test itself to ensure that qualified applicants are hired, and to reduce the adverse impact created by the use of the physical ability test without endangering the general public. Employers must determine an adequate validity-adverse impact trade-off.

The trade-off between utility and adverse impact is often a focal point of physical ability test court cases. Sady, and Dunleavy (2013) addressed the importance of content validity in selection testing. They stated that there are instances in which a test with high adverse impact will have the same criterion validity as another test; yet the former test should be used because it relates more closely to the demands or requirements of the work itself. That is, the tests are predicting performance (as a whole) equally well, but the former test is better for the specific critical components of the job that the organization needs to assess. Content validity is the most commonly addressed issue once disparate impact has been statistically demonstrated. However, organizations that use physical ability tests may struggle to defend their test unless they are able to demonstrate both its content and criterion validity.

There are instances in which the criterion validity of a test may not co-exist with content validity. Criterion validity at its base is simply how well a test predicts job performance (EEOC Uniform Guidelines, 1978). Sothman, Gebhardt, Baker, Kastello, and Sheppard (2004) identified a police officer hiring scenario in which males passed a physical ability test at a rate of 93.2% while females passed at a rate of 16.2%, clearly demonstrating adverse impact. However, Maher (1984) noted that females have been found to perform equally to men in instances of violence which would seemingly require physical abilities. In this instance, effective performance in the violent scenario is the

criterion that the test is supposed to predict. Therefore, although the test was content valid, it appeared to lack criterion related validity. Wax (2011) stated that demonstration of criterion validity is the most rigorous and demanding of the validities, and therefore it is the "gold standard" for Industrial Organizational Psychologists.

Additionally, there is concern regarding the construct validity of any test. In a selection context, construct validity reflects how well a test measures a specific characteristic in applicants which should, in turn, lead to stronger performance on the job (EEOC Uniform Guidelines, 1978). Essentially, construct validity refers to how well a given test measures underlying attributes that lead to successful job performance. However, the Uniform Guidelines also identify the difficulty in using construct validity as a primary measure of a test's functionality as it has low generalizability and its uses in employee selection were relatively unstudied. The wide range of validity interpretations can make ruling on court cases difficult, but there are some standards that have been utilized by the courts.

Court Cases

Historically, there have been four primary factors that influence court rulings with regard to physical ability testing: minimum qualifications, close approximation, manifest relationship, and the Spurlock Doctrine.

Minimum qualifications set the lowest possible standards that will lead to effective performance in a given position. Minimum qualifications represent a non-compensatory model because high scores on another facet of the application process cannot compensate for a score that falls below the minimum qualification (Sady, & Dunleavy, 2013). Hollar (2000) suggested that minimum qualifications are particularly

difficult for employers to define. Physical ability tests often inherently produce adverse impact, so determining the cutoff scores that balance quality of the hiring process with reducing adverse impact can become excessively difficult. For instance, in *Lanning v. Southeastern Pennsylvania Transportation Authority (SEPTA) (1999)*, the cutoff score of 12 minutes for a mile and a half run led to just 6.7% of females successfully meeting the standard. However, this hiring company (SEPTA) also had the added difficulty in that their position dealt with public safety.

After the initial trial and appeal, SEPTA was given the option to modify their selection procedure, and the case was remanded (Hollar, 2000). The view of the dissenting judge in this case was that rulings which favor the modification of a test that uses a cutoff score could lead to an outright abandoning of cutoff scores. This could prove problematic, particularly in instances where public safety is concerned. This concern stems from the possibility that an applicant who was not held to a cutoff score would be hired, and would be unable to properly protect the public (e.g., a fireman who cannot help someone out of a burning building). Ultimately, *Lanning v. SEPTA* (1999) was ruled in favor of the defendant. Minimum qualifications may be considered the most stringent requirement in the selection process of a physical ability test because they are cutoff scores; those scoring below the threshold are no longer considered for the position.

A less stringent standard that has been established in PAT court cases is *close approximation*. Close approximation and content validity are virtually synonymous; that is, both involve the demonstration that the physical ability test resembles what is actually performed on the job. In the landmark case of *Dothard v. Rawlinson* (1977), both gender and physical requirements (height and weight) were scrutinized as selection procedures

for prison guards. The court found that height and weight were not inherently job related. However, the court found that being male was a bona fide occupational qualification (BFOQ) because females in this job would be at high risk of sexual assault, which could prove unsafe for the prison as a whole. The primary implication of this case for physical ability cases is not the establishment of BFOQ, but rather the establishment that height and weight were not deemed appropriate measures for hiring. In contrast, in *Hardy v. Stumpf* (1978), Oakland police applicants were asked to scale a six-foot wall. Those who failed in this task were not considered for hire. The court upheld the selection procedure and ruled in favor of the defendant, presumably because police officers may need to scale something of this size in the line of duty (e.g., a fence). This case demonstrates that in instances where adverse impact exists, the defendant may be successful *if* there is a successful presentation of the job relevance of the selection procedure. The primary difference between these two cases is in the way the defendants measured the target behavior. In *Dothard v. Rawlinson* (1977), the defendants were using height and weight requirements as substitutes (proxy variables) for actual physical performance. In contrast, *Hardy v. Stumpf* (1978) measured the ability to climb a wall by directly measuring the target attribute.

Even less demanding than close approximation is the standard of a *manifest relationship*. The demonstration of a manifest relationship requires only that the employer prove that there is some relationship between the test and the job (Hollar, 2000). Close approximation necessitates that the test closely reflects the job. In contrast, demonstration of any link from the test to the job is all that is required to show a manifest relationship. In *Eison v. City of Knoxville* (1983), a large number of physical fitness tests

were used to assess potential police officers. The defendant needed to demonstrate only that the cutoff scores set were set for a reason and were not arbitrary. Proving a manifest relationship is one of the simplest defenses that a company can provide to demonstrate the necessity of their specific selection procedure. Wax (2011) stated that the demonstration of a manifest relationship requires only a "plausible match" between the test and the functions of the job.

The *Spurlock Doctrine*, in its most basic form, states that the employer's responsibility to prove business necessity is lessened for positions where public safety is a concern (Hollar, 2000). Therefore, if a physical ability test is legally challenged, the employer is still responsible for demonstrating the content validity of the test (given that adverse impact has already been demonstrated). However, if the test can be linked to an aspect of the position which involves public safety, the defendant has a greater likelihood of success in defending the use of its selection procedure. Wax (2011) stated that the Spurlock Doctrine has also been referred to as the "demonstrably necessary" test. Wax continued that in lower courts this doctrine has been used to require a relationship between the criterion and job, but does not in fact require statistical validation. Thus, interpretations of the Spurlock Doctrine are slightly varied, but both interpretations place a lighter burden on the employer to demonstrate business necessity than in instances where public safety is not concerned.

In summation, the literature provides information about how various law changes and rulings led to our current understanding of how physical ability tests are addressed in the courts. Title VII of the 1964 CRA as revised in 1972, prior to CRA-91 provided a relative framework; CRA-91 clarified that the plaintiff has the burden of proof to

demonstrate adverse impact, and that defendants must then demonstrate job relatedness and business necessity. Moreover, there are three primary methods by which physical ability is tested, each of which have ties to content, criterion, and construct validity. Several previous rulings such as *Hardy v. Stumpf* (1978), *Eison v. City of Knoxville* (1983), and *Dothard v. Rawlinson* (1977) have provided basic standards by which we can predict the outcome of a given case, but the intersection of validity and these standards creates a certain degree of ambiguity when attempting to determine which standards and procedures will prevail in court.

The Present Study

In the present study, I analyzed district, appellate, and Supreme Court level cases that focus on physical ability testing. The Civil Rights Act of 1991 led to a more unified structure by which adverse impact cases were handled. Therefore, only court cases occurring from 1992 on were included.

Hypotheses

Each of the three hypotheses compared one group to another with regard to the proportion of cases won by the plaintiff as opposed to those won by the defendant. Cases which were remanded or settled out of court were not included in the statistical analyses.

Hough et al. (2001) stated that each form of pure ability testing assessed yielded a higher score for men than for women, which would lead to higher instances of adverse impact against women. One method to prevent adverse impact, subgroup norming, was made illegal by CRA-91; as such, cases that include norm referencing are unexpected in the cases used in this study. Therefore, cases involving pure ability tests were compared

to cases involving work sample tests with regard to proportion of cases won by the plaintiff.

Hypothesis 1: Cases involving pure ability will be ruled in favor of the plaintiff more frequently than will cases involving work sample physical ability testing.

Hollar (2000) discussed the Spurlock Doctrine as placing a lighter burden of proving job relevance of a test in instances where public safety is a concern, which should lead to defendants more easily prevailing in cases involving public safety.

Hypothesis 2: Cases involving public safety will have a higher frequency of ruling in favor of the defendant than will cases that do not involve public safety.

Gutman et al. (2011) discussed that recent EEOC modifications to the definition and proof of disability with the ADAAA (2011) have broadened the definition of "disabled." Because of this shifting interpretation of disability, the courts may find more often that a given physical ability test ignores the potential for a disabled person to perform the job adequately while not being able to pass a physical ability test. That is, the disabled individuals may fail the physical ability test, but still have ability to perform the job well.

Hypothesis 3: ADA cases since 2011 will have a higher frequency of rulings in favor of the plaintiff than will those before 2011.

Method

The study at-hand is an expansion of previous research which analyzed and coded court cases from 1992 to 2006 (Starling, 2006). Because the Civil Rights Act of 1991 led to a more unified structure by which adverse impact cases were handled, only court cases

occurring from 1992 on were included. Similar to the 2006 study, this research used the Lexis-Nexis search engine to locate relevant cases. The current study focused on district level and higher court cases due the limited number of appellate level cases occurring from 2006-2014. This was in contrast to the 2006 study, which utilized only appellate and Supreme Court level cases. The search terms mimicked that of the previous study and included: *Physical Ability Test, Physical Agility Test, Physical Fitness Test, and Physical Capability Test.*

Coding Factors

The coding scheme used was adopted from Shoenfelt and Pedigo's (2005), which was based on Werner and Bolino's (1997) coding scheme. The coding scheme included 15 variables. It should be noted that each of the coded factors could be coded as *No Information* (NI) in instances where the case did not provide explicit information regarding the factor. For example, in cases that were based on gender discrimination, the race/ethnicity of the plaintiff was often unmentioned. Each of the 15 coded factors warrants further explanation regarding the rationale for how it was coded.

Gender of Plaintiff: Could be coded as either *male* or *female.*

Race/Ethnicity of Plaintiff: Could be coded as: *White, Afro-American, Hispanic, Asian,* or *Native American.*

Basis for Lawsuit (Claim) for Discrimination: This factor assessed the grounds on which the claim was made. Cases generally fell under the categories of *gender, race, or ADA* claims, but exceptions were found (e.g., *retaliation, legality of petition*).

Industrial, Professional, or Civil Service Work: This factor assessed the type of work that was being done by those in the position of the plaintiff. *Industrial* work entails

blue collar positions such as assembly line work, manual labor in a factory setting, or work on a production line. *Professional* work entails positions whose primary function is within an office setting or white collar jobs. *Civil service* work includes positions such as police officers, fire fighters, and military officers.

Class Action or Individual Plaintiff: This could be coded in three different ways. *Individual plaintiff* was used in instances where there was only one person filing the claim. *Multiple plaintiffs* was used in instances in which there were multiple people filing the claim, but they did not constitute a class. *Class action* was used in instances where the plaintiffs were a certified class.

Work Sample/Pure Ability or Other Type of Physical Ability Test: The physical ability test was coded as a *work sample* test in instances where the test mimicked the functions of the job or when it was a simulation. If the job in question involved tasks such as lifting boxes of a certain weight, the ability to lift a certain weight was still considered to be a work sample because the task mimics what is done within the job itself. The test was coded as *pure ability* where general physical performances were assessed (e.g., mile run). Any test which did not fit within these two categories was coded as *other.* A wellness test would exemplify the *other* category, as it assesses the general physical wellness of the applicant, rather than the pure ability of the individual.

Standardized/Professionally Developed Test: This could be coded as either *yes* if the test was a standardized or professionally developed test, or *no* if it was not.

In-House or Consultant: This factor was coded based on who developed the physical ability test. If it was created by employees within the organization, this was

coded as *in-house*. If the test was created by an individual consultant or an outside consulting firm, this was coded as *consultant*.

Test Validation: This was coded as *yes* or *no* depending on whether the test had undergone a professional validation procedure. If the answer to this question was yes, the additional question of what type of validity was used was then answered. This second question was coded as either *criterion* or *content.*

Practice or Training Available: This was coded as *yes* or *no* depending on whether or not the applicant was able to either use training material, or practice taking the test itself.

Public Safety Issue: This was coded as *yes* or *no*. Public safety was considered an issue only in positions which the employee would be protecting the public. Thus, even though a truck driver has the potential for affecting the safety of the public, they are not responsible for protecting the public as an explicit part of their job.

Additional Selection Tests: This was coded as *yes* or *no*. Any other process/test that was used to assess an individual other than the physical ability test was considered an additional selection test. Thus, even an application blank would constitute an additional selection test.

Court Verdict: There were several possible verdicts the court could render. The verdict was coded as *plaintiff* or *defendant* in instances where the plaintiff or defendant won the case. However, cases could also be *remanded* to the lower courts, *settled* out of court, or there could be a *split decision* in which certain claims were won by the plaintiff and some by the defendant. In split decisions, the claim surrounding the physical ability

test was the primary focus. In these instances the coding would indicate a split decision, but also which side won on the issue of the physical ability test.

Decision Type: This could be coded as *selection* when the test was being used to determine which applicants would be selected for either initial hire or promotion. This could also be coded as *re-entry* in instances where the physical ability test is used for an employee who has already been hired, but is asked to take the test for either return from leave or as part of a newly instituted program by the organization.

Coding

Two Industrial/Organizational Psychology graduate students independently coded each of the cases found. These graduate students were trained in the use of the coding scheme (see explanation of coded factors below). The inter-coder agreement reliability of the coding factors between students was calculated by the number of instances where raters agreed divided by the total number of coded factors. This inter-coder agreement resulted in a 91.5% reliability between the two individuals responsible for coding when all cells in Table 1 were considered (after exclusion of *International Guards Union v. United States Department of Army,* 2007). That is, cells with no information were used. . When cells with no information were excluded, the inter-coder reliability was 89.7%. A sample of four of the 22 cases found in the 2006 study were re-coded to ensure reliability between raters before the coding of the new cases occurred. The coding factors can be found in Table 1.

Table 1

Coding Factor	**Definition**	**Code**
Job Analysis	Was a job analysis performed?	Yes, No
Gender of Plaintiff	Was the plaintiff male or female?	Male, Female
Race/Ethnicity of Plaintiff	What was the plaintiff's race/ethnicity?	Caucasian, Afro-American, Hispanic, Native American, Other
Basis for Lawsuit (Claim)	What did the plaintiff argue as the basis for discrimination?	ADA, Gender, Race, Age, Other
Industrial, Professional, or Civil Service Work	What type of job was in question in the lawsuit?	Industrial, Professional, Civil Service
Class Action or Individual Plaintiff	Was the plaintiff one person, or was this a class action lawsuit?	Class Action, Individual, Multiple
Work Sample/Pure Ability or Other Type of PAT	Did the test consist of on-the-job behaviors or other general physical conditions, such as strength or stamina?	Work Sample, Pure Ability, Other
Standardized/Professionally Developed Test	Was the test used in the selection procedure standardized/professionally developed?	Yes, No
In-House or Consultant	Was the test developed in-house or by a consultant?	In-House, Consultant
Test Validation	Was the test that was used for selection validated?	Yes, No
Practice or Training Available	Were training materials or practice time offered prior to testing?	Yes, No
Public Safety Issue	Was public safety an issue of concern related to job performance?	Yes, No
Additional Selection Tests	Were there additional tests that were used as part of the selection process?	Yes, No
Court Verdict	Did the court rule in favor of the defendant or the plaintiff?	Plaintiff, Defendant, Settlement, Summary Judgment, Remanded
Decision Type	Was the test used for hiring or was it used for re-entry to the position?	Selection, Reentry

Results

The LexisNexis search of the terms *Physical Ability Test, Physical Agility Test, Physical Fitness Test,* and *Physical Capability Test* yielded 59 related cases. After reviewing each of these cases, it was determined that 23 of these cases pertained directly to physical ability testing. However, because one of these cases focused on a collective bargaining agreement instead of the actual physical ability test, it was not included in the analysis (*International Guards Union v. United States Department of Army*, 2009). Using the same 22 cases coded by Starling (2006), the total number of cases used for analysis was 44. The court cases reviewed may be found in Appendix A. The 22 cases coded by Starling were reviewed and a sample of four cases were recoded to ensure that the coded factors were coded in the same manner. The inter-coder agreement for these four cases was 96.4%. The 22 new cases were coded using the factors listed in Table 1. It should also be noted that three factors were added to Starling's coding scheme: whether a job analysis had been completed, at what level of court the case was tried, and what the test was used for (selection or re-entry into an already held position). Job analysis was added as a factor because it was lacking from the original study. Court level was added so that appellate level and district level cases could be distinguished from one another. Decision type was added because several cases were used for purposes other than selecting applicants for hire or promotion. The previous 22 cases were not recoded because there was a high degree of reliability between the coding completed by Starling (2006) and a sample from these cases coded by the individuals who coded the present study. However, the original 22 cases were coded on the three additional factors added to the current

study. The results of the coding for all 44 cases can be found in Appendix B. A summary of the overall findings may be found in Appendix C.

To test each of the hypotheses, a *z test for differences between proportions between independent samples* was used. The formula for this test can be found below:

$$\frac{\frac{r_1}{n_1} - \frac{r_2}{n_2}}{\sqrt{\left(\frac{r_1 + r_2}{n_1 + n_2}\right)\left[1 - \left(\frac{r_1 + r_2}{n_1 + n_2}\right)\right]\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \mathrm{Z}$$

In this equation *r* refers to the number of cases won, whereas *n* means the total number of cases. The subscripts refer to each of the two groups being compared (e.g., cases involving pure ability tests compared to cases involving work sample tests). In each analysis, one group was compared to another with regard to frequency. Because the variables were assessed dichotomously (plaintiff win or defendant win), cases settled out of court, and those that were remanded were not included in the statistical analyses.

Hypothesis 1 stated that cases involving pure ability would be ruled in favor of the plaintiff more frequently than would cases which utilized work sample tests. Three of the 14 cases which utilized a work sample test found in favor of the plaintiff, while ten found in favor of the defendant. One case that used a work sample test was settled out of court (*Vasich v. City of Chicago*, 2013). Thus, three of 13 work sample cases (23.1%) were won by the plaintiff. Of the 21 cases which utilized a pure ability test, eight ruled in favor of the plaintiff, 12 ruled in favor of the defendant, and one was settled out of court (*United States v. Commonwealth of Massachusetts*, 2012). Thus, eight of 20 pure ability

cases (40.0%) were won by the plaintiff. The remaining nine cases were either categorized as *other* types of tests or the cases did not provide enough information to determine what type of test was being utilized. Hypothesis 1 was not supported (z = 1.01). The critical z value to demonstrate significance for a one-tailed test is 1.65 (p < .05), which was not reached.

Hypothesis 2 stated that cases involving positions of public safety would have a higher frequency of ruling in favor of the defendant than would cases that did not involve public safety. Hypothesis 2 was also not supported (z = -.09; p > .05). Seven of the 44 coded cases involved positions which did not involve public safety. Of these seven cases, five ruled in favor of the defendant (71.4%), while two ruled in favor of the plaintiff. There were 33 cases with positions that involved public safety which rendered verdicts useable for analysis. Of the cases that involved public safety, 23 found in favor of the defendant (69.7%), while ten found in favor of the plaintiff. Each of the remaining four decisions did not rule in favor of the plaintiff or defendant, but rather were remanded, or settled out of court.

Hypothesis 3 stated that ADA cases since 2011 would have a higher frequency of rulings in favor of the plaintiff than would those before 2011. Three cases based on ADA claims after 2011 were found (*Chicago Regional v. Thorne Associates*, 2012; *Kotwica v. Rose Packing*, 2010; *Spires v. Ingersoll Rand*, 2013). All three ruled in favor of the defendant; none ruled in favor of the plaintiff (0.0%). Eight ADA cases were found before 2011; four of these eight ruled in favor of the plaintiff (50%), while the other four found in favor of the defendant. Hypothesis 3 was not supported (z = -1.73; p > .05). Because the 2008 ADAAA was a more expansive extension of ADA than were the 2011

modifications, an additional analysis comparing ADA based cases before 2008 to ADA cases after 2008 was run. In this analysis, two of six ADA claims after 2008 were ruled in favor of the plaintiff (33.3%). Two of the five ADA cases before 2008 were ruled in favor of the plaintiff (40%), and three ruled in favor of the defendant. This analysis likewise yielded nonsignificant findings (z = -.231; p > .05)

Discussion

Selection processes are unique opportunities for employers to assess the abilities of their applicants to determine who best fits the job that is to be filled. Physical ability tests aid in this process for positions that have significant physical demands.

Pure ability tests measure abilities such as mile runs, push-ups, and sit-ups (Hoover, 1992). Because pure ability tests (i.e., tests that measure fitness rather than performance) consistently produce higher scores for male participants than for female participants (Hoover, 1992), this type of test will often result in adverse impact against women. Adverse impact refers to disproportionately lower rates of hire/promotion for those within a protected group (Gutman et al., 2011). In addition to females, Arvey, Nutting, and Landon (1992) stated that this adverse impact can be found against racial groups such as Asian and Hispanic applicants. Additionally, Ryan et al. (1996) stated that work sample tests, that is, tests that simulate the physical abilities used to do the job, are often viewed as less discriminatory by applicants. Because of all of these factors, Hypothesis 1 predicted that cases involving pure ability tests would be more likely to be found in favor of the plaintiff. Although 10 pure ability cases were found in favor of the plaintiff and 23 were found for the defendant, this hypothesis was not statistically supported. The number of cases which utilized work sample tests was inadequate to

assess this hypothesis. Further research should be completed with regard to this hypothesis, as there is high potential for support if using a larger sample yielding more power to detect a significant effect.

The most frequent use of physical ability tests is for positions that involve public safety. It is of particular importance that those entrusted with protecting others have the physical abilities to effectively perform their jobs. Accordingly, Hollar (2000) stated that the Spurlock Doctrine places a lighter burden on defendants to demonstrate business necessity of a physical ability test when the test is used to assess those entering jobs that involve protection of the public. It would therefore stand to reason that positions involving public safety would have a higher frequency of ruling in favor of the defendant. However, Hypothesis 2, which stated that this would be found, was not supported.

Of the seven cases that did not involve public safety, only two were won by the plaintiff. In contrast, 10 of the 33 cases which involved public safety resulted in the plaintiff winning. One possible explanation for this finding is that because the Spurlock Doctrine is well established, those who choose to pursue legal action against employers of public safety positions do so knowing that the burden of proof lies more heavily on them (the plaintiff) than for positions which have no relation to public safety. Therefore, it may be that individuals choose to pursue litigation less frequently in instances that involve public safety than when the positions did not involve public safety. Further, the sample size of cases that did not involve public safety (seven) was very limiting in the statistical analysis. Additionally, some of these cases did not focus on the physical ability test itself. In fact, of the six cases that public safety was not involved in the position, four challenged the actual administration of the test rather than what was entailed within the

test (*Davis v. CDA*, 2010; *Kotwica v. Rose Packing*, 2010; *Norman v. Healthsouth Rehabilitation*, 2008; *Spires v. Ingersoll Rand*, 2013). Positions of public safety will inherently require physical ability testing to ensure that those hired have the requisite ability to protect others. However, positions which do not involve public safety may have less stringent requirements of their applicants because the tasks within these jobs may not be as physically demanding as jobs that concerned with public safety. This is one possible explanation for the lack of support for Hypothesis 2.

Given the limited number of cases related to physical ability tests, narrowing the search to ADA cases post 2011 was not fruitful in obtaining enough cases to adequately test the hypothesis. In total, only 11 cases were found involving ADA claims. Hypothesis 3, which stated that ADA cases after 2011 would be ruled in favor of the plaintiff more frequently than those before 2011, was not supported.

It should be noted that only two types of claims resulted in the plaintiff being successful in their claim: i.e., ADA and gender claims. It has been well demonstrated that women are adversely impacted by physical ability testing (Hoover, 1992). It may also be that disabled individuals are being discriminated against through the use of physical ability testing. Future research should further investigate the role that disability plays in both the administration and execution of physical ability tests. Testing must be related to essential job functions, but the test administration may need be adjusted in instances of disabled individuals to provide accommodation. In one case (*Chicago Regional v. Thorne Associates*, 2012), the focal point of the case was whether or not the individual was disabled, rather than issues surrounding the physical ability test. Therefore, further research may seek to separate ADA claims from Title VII claims, because in ADA cases

the plaintiff must first demonstrate proof of disability and ability to perform essential job functions with or without accommodation before a demonstration that a physical ability test is job related by the defendant (Gutman et al., 2011). This means that many ADA cases may never reach the second phase of the case, the part that actually pertains to the physical ability test. Additionally, further research should investigate the role that testing plays in cases where temporarily disabled individuals are returning to work. Testing for reentry to work is different from selection testing. In reentry testing, the focal point may not be the utility of the test, but rather the necessity of administration of the test. The court cases that involve ADA claims are unique and often involve issues of accommodation, administration, and whether the test assesses essential job functions. Thus, further investigation into ADA cases seems a possible direction for study on physical ability testing.

Another interesting finding through the course of this study relates to the comparisons across the court levels. Appellate and Supreme Court level cases yielded a much lower number of cases in which the plaintiff won. In fact, of the 26 Appellate and Supreme Court cases in which a verdict was rendered, only three verdicts were in favor of the plaintiff. Two of the appellate cases were remanded. This is in contrast to the eight of the 20 cases found for the plaintiff at the district level. Two of these cases were settled out of court.

Limitations

Several limitations were encountered throughout the course of the research. Since the work of Starling (2006), there were fewer appellate and Supreme Court level cases. From 1992-2006, 22 cases were found at the Supreme Court and appellate level; in

contrast, from 2006-2014, only four cases at the appellate level were located. Consequently, it was determined that district level cases should be included in the 2006-2014 analysis to obtain enough cases to perform adequate statistical analysis. The analysis in this study included only appellate and Supreme Court cases prior to 2006, and included appellate, Supreme Court, and district level cases from 2006-2014.

Additionally, the sample size as a whole was a limitation. Although a total of 45 cases were used for analysis, the hypotheses themselves substantially narrowed the number of analyzed cases to test each hypothesis. For instance, the total number of cases involving ADA claims was 11. This group of 11 cases had to be further divided into cases before 2011 and after 2011. It may have been more beneficial to assess the cases as a whole rather than limiting them to distinct groups through the hypotheses. That is, it may have been more practical to use hypotheses that divided all 44 cases into two groups rather than analyzing smaller portions of the groups. If a larger sample size had been obtained, the hypotheses chosen would have more easily statistically tested.

In this same manner, another limitation to the study was the number of cases involving ADA claims after 2011. Hypothesis 3 sought to compare ADA cases before 2011 to those after 2011. The number of physical ability testing cases involving ADA after 2011 was only three. It therefore stands to reason that a hypothesis that compared cases before the ADAAA (2008) to cases after ADAAA would have been more beneficial. This is because the ADAAA was a more massive expansion of the ADA than were the 2011 modifications, and because the number of cases would have been more adequate for statistical analysis. However, analysis of this additional hypothesis yielded results that were non-significant.

There were additional limitations. In some instances, very little information was provided in the case record regarding the actual physical ability test and, because of this, many coded factors have incomplete data. In many of the cases, the central issue of the case was not whether the physical ability test was valid or developed correctly, but instead whether or not the individual should have been tested in the first place. For instance, in *Kotwica v. Rose Packing Company* (2010), the plaintiff was provided leave for a surgery. She was instructed by her physician to avoid certain physical tasks. Upon her return to work, she was asked to be seen by the company physician, who asked her to lift 50 pounds as demonstration of her capability to return to work. Because this violated her personal physician's orders, she did not complete the task and ultimately was fired, with the company physician stating that she did not complete the physical ability test. All of the limitations mentioned above ultimately led to difficulty in analyzing the hypotheses adequately. Finally, it is recommended that in future research evaluating court cases, the reason for finding in favor of the plaintiff or defendant be coded. This information would provide insight into important factors underlying the court decisions.

Conclusions

The limited number of cases available for testing the hypotheses was a detriment to the statistical power to detect patterns in the court cases reviewed. However, several interest findings did occur throughout the course of this study. Of the 44 cases related to physical ability testing, only 11 were won by the plaintiff whereas 28 were won by the defendant. This suggests that, overall, there is less likelihood that a plaintiff will be successful in a claim against a company based on physical ability testing.

However, the analysis of court findings surrounding physical ability testing is not a fruitless endeavor. The coding of these cases may, in fact, yield useful information for employers about how to best avoid litigation. Although none of the hypotheses were supported, the study provides information about issues that were not apparent prior to the research. For instance, prior to investigation of the cases, it was assumed that all cases would focus on the validity of the tests themselves. What was found was very different from this assumption in that many cases focused on the timing of administration *(Spires v. Ingersoll Rand*, 2013), the seniority policies regarding the administration of the test (*Davis v. CDA*, 2010), or even if the test violated union policies (*International Guards Union v. United States Department of Army*, 2007). This provides information to employers and Industrial Organizational Psychologists, suggesting that the validity of a test is important, but the administration and implementation of the test are likely equally important.

Appendix A: Court Cases

Appellate Level

Andrews v. City of Cookeville., 63 Fed. Appx. 804; (U.S. App. 2003).

Barrientos v. City of Eagle Pass., 444 F. Appx. 756 (U.S. App. 2011).

Brunet v. City of Columbus., 58 F. 3d 251; 1995 (U.S. App. 1995).

Brunet v. City of Columbus., 1 F. 3d 390; 1993 (U.S. App. 1993).

Danskine v. Miami Dade Fire Dep't., 253 F. 3d 1288; (U.S. App. 2001).

Davoll v. Webb., 194 F. 3d 1116; (U.S. App. 1999).

Dugan, Larry W. Durbin, Jesse R. Masters, Steven M. Ott., 532 (Fed. Appx. 805 2013).

Dyke v. O'Neal Steel., Inc. 327 F. 3d 628; 2003 (U.S. App. 2002).

Garcia v. City of Houston., 201 F. 3d 672; 2000 (U.S. App. 2000).

Holiday v. City of Chattanooga., 206 F. 3d 637; (U.S. App. 1999).

Howard v. City of Southfield., 95-1014; 1996 (U.S. App. 1996).

International Union, United Plant Guard Workers v. Lockheed Martin Util., 97-3495; (U.S. App. 1998).

James v. Sheahan., 137 F. 3d 1003; (U.S. App. 1997).

Jansen v. Cincinnati., 977 F. 2d 238; (U.S. App. 1992).

Koger v. Reno., 98 F. 3d 631; (U.S. App. 1996).

Lanning v. SEPTA., 308 F. 3d 286; (U.S. App. 2002).

Lanning v. SEPTA., 181 F. 3d 478; (U.S. App. 1999).

Merrit v. Old Dominion Freight Line., 601 F.3d 289 (U.S. App. 2010).

Peanick v. Morris., 96 F. 3d 316; (U.S. App. 1996).

Peightal v. Metropolitan Dade County., 58 F. 251; (U.S. App. 1995)

Pietras v. Board of Fire Comm'rs., 180 F. 3d 468; (U.S. App. 1999).

Stahl v. Bd. of County Comm'rs., 101 Fed. Appx. 316; (U.S. App. 2004)

Thomas v. City of Omaha., 63 F. 3d 763; (U.S. App. 1995).

Wright v. Illinois Dep't of Corrections., 204 F. 3d 727; (U.S. App. 1999).

Kotwica v. Rose Packing Company, Inc., F.3d 744 (U.S. App. 2010).

District Level

Chicago Regional Council of Carpenters, United Brotherhood of Carpenters and Joiners of America, and Rodolfo Rosas, Sr. v. Thorne Associates., 893 F. Supp, 2d 952 (E.D. Il. 2012).

Davis v. CDA Incorporated., 1:09 CV-406-WKW (M.D. S. Al. 2010).

Ernst, Dawn Hoard, Irene Res Pullano, Michelle Laha-Lih and Katherine Kean v. City of Chicago., 08 C 4370 (N.D. Il. 2013)

Cherie Easterling v. State of Connecticut Department of Correction., 783 F. Supp. 2d 323 (D. Co. 2011).

EEOC v. Akal Security., 08-1274-JTM-KMH (D. Ka. 2009).

EEOC v. Lyon-Dell-Citgo Refining., 835; 13 (S.D. Te. 2008).

Ellis v. Chertoff., 4:07CV00041 JMM (E.D. Ar. 2008)

Eudy v. The City of Ridgeland, Mississippi and Matthew Bailey., 464 F. Supp. 2d 580 (S.D. Mi. 2006).

Gilbert v. Village or Cooperstown., 3:09-CV-754 (N.D. N Yo. 2011).

Godfrey v. City of Chicago., 12 C 08601 (N.D. Il 2013).

Godoy v. Habersham Country., 2:04-CV-211-RWS (N.D. Ge. 2006).

Hunter v. Santa Fe Protective Services., 822 F. Supp. 2d 1238 (M.D. Al 2011).

International Guards Union of America, Local #106 v. C&D Security, Inc., 07-CV-523 MCA/ACT (D. N Me. 2007).

Norman v. Healthsouth Rehabilitation Centers of Louisville., 06-465-C (W.D. Ke. 2008).

Spires v. Ingersoll Rand and Trane U.S. Inc., C/A No.: 3:11-2530-TLW-SVH (D. S. Ca. 2013).

Starkey v. City of Burnsville, Minnesota., 07-1948 DWF/SRN (D. Mi. 2008).

Turner v. Arkansas Children's Hospital., 4:10CV00746 SWW (E.D. Ar. 2011).

United States of America v. The Commonwealth of Massachusetts., 781 F. Supp. 2d 1 (U.D. Ma. 2011).

Vasich v. City of Chicago., 11 C 04843 (N.D. Il. 2013).

Appendix B: List of Cases Summarized

Case	Gender	Race	Lawsuit	Work	Action	PAT Type	Stnd	In-House/ Consultant	Validated	Practice	Safety	Additional Tests	Verdict	Level	Job Analysis	Selection Type
Andrews v. City of Cookeville	Male	NI	Age	Civil	Individual	Other	NI	NA	NI	NI	Yes	Yes	Remanded	Appellate	NI	Selection
Brunet v. City of Columbus	Both	NI	Gender	Civil	Class	Work Sample	Yes	Consultant	Yes	NI	Yes	Yes	Defendant	Appellate	Yes	Selection
Brunet v. City of Columbus	Both	NI	Gender	Civil	Class	Work Sample	Yes	In-House	Yes	NI	Yes	Yes	Split (PAT for defendant)	Appellate	Yes	Selection
Cindea v. Jackson Twp.	Male	NI	Age	Civil	Individual	Work Sample	NI	NI	NI	NI	Yes	Yes	Defendant	Appellate	NI	Selection
Danskine v. Miami Dade Fire Dep't	Male	NI	Gender	Civil	Class	Other	NI	NI	NI	NI	Yes	Yes	Defendant	Appellate	NI	Selection
Davoll v. Webb	Both	NI	ADA	Civil	Class	Work Sample	NI	NI	NI	NI	Yes	Yes	Split (PAT for plaintiff)	Appellate	NI	Reentry
Dyke v. O'Neal Steel, Inc.	Male	NI	ADA	Industrial	Individual	Other	NI	NI	NI	NI	Yes	Yes	Defendant (summary judgment)	Appellate	NI	Selection
Garcia v. City of Houston	Male	Hispanic	Race	Civil	Individual	Pure ability	NI	NI	NI	NI	Yes	Yes	Defendant	Appellate	NI	Selection
Holiday v. City of Chattanooga	Male	NA	ADA	Civil	Individual	Pure ability	NI	NI	NI	NI	Yes	Yes	Plaintiff	Appellate	NI	Selection
Howard v. City of Southfield	Male	Afro-American	Race	Civil	Individual	Other	NI	NI	NI	NI	Yes	NI	Defendant	Appellate	NI	Selection
International Union v. Lockheed	NI	NI	NI	Industrial	Individual	Pure ability	NI	NI	NI	NI	Yes	Yes	Defendant	Appellate	NI	Reentry
James v. Sheahan	Female	NI	Gender	Civil	Individual	Pure ability	NI	NI	NI	NI	Yes	Yes	Defendant	Appellate	NI	Selection
Jansen v. Cincinnati	Male	White	Race	Civil	Class	Other	NI	NI	NI	NI	Yes	Yes	Remanded	Appellate	NI	Selection
Koger v. Reno	Male	NI	Age	Civil	Individual	Pure ability	NI	NI	NI	NI	Yes	Yes	Defendant	Appellate	NI	Selection
Lanning v. SEPTA (2002)	Female	NI	Gender	Civil	Class	Pure ability	Yes	NI	Yes	Yes	Yes	Yes	Defendant	Appellate	Yes	Selection
Lanning v. SEPTA (1999)	Female	NI	Gender	Civil	Class	Work Sample	Yes	Consultant	Yes	NI	Yes	Yes	Remanded	Appellate	Yes	Selection
Peanick v. Morris	Male	Native American	Race	Civil	Individual	Pure ability	NI	NI	NI	NI	Yes	Yes	Defendant	Appellate	NI	Selection
Peightal v. Metro Dade County	Male	White	Race	Civil	Individual	Other	NI	In-House	No	NI	Yes	Yes	Defendant	Appellate	NI	Selection
Pietras v. Board of Fire Comm'rs	Female	NI	Gender	Civil	Individual	Pure ability	No	NI	No	NI	Yes	Yes	Plaintiff	Appellate	NI	Reentry
Stahl v. Bd. of County Comm'rs	Female	NI	Gender	Civil	Individual	Pure ability	NI	NI	NI	Yes	Yes	Yes	Defendant	Appellate	NI	Reentry
Thomas v. City of Omaha	Female	NI	Gender	Civil	Individual	Other	NI	NI	NI	NI	Yes	Yes	Defendant	Appellate	NI	Selection
Wright v. Illinois Dep't of Corrections	Male	NI	ADA	Civil	Individual	Work Sample	NI	NI	NI	NI	Yes	Yes	Defendant	Appellate	NI	Selection

Case	Gender	Race	Lawsuit	Work	Action	PAT Type	Stnd	In-House/ Consultant	Validated	Practice	Safety	Additional Tests	Verdict	Level	Job Analysis	Selection Type
Barrientos v. City of Eagle Pass	Male	NI	Gender	Civil	Individual	Pure Ability	NI	In-house	NI	NI	Yes	NI	Defendant	Appellate	NI	Selection
Chicago Regional v. Thorne Associates	NI	NI	ADA	Industrial	Multiple	Work Sample	Yes	Consultant	NI	NI	No	NI	Defendant	District	NI	Selection
Davis v. CDA	Male	NI	ADA, Age	Civil	Individual	Pure Ability	Yes	In-house	NI	Yes	Yes	NI	Plaintiff	District	NI	Reentry
Dugan v. Amtex	Both	NI	Age	Civil	Multiple	Pure Ability	Yes	In-house	NI	NI	Yes	NI	Defendant	Appellate	NI	Reentry
Earnst v. City of Chicago	Female	NI	Gender	Civil	Multiple	Pure ability	Yes	Consultant	Yes	NI	Yes	Yes	Plaintiff	District	Yes	Selection
Easterling v. State of Connecticut	Female	NI	Gender	Civil	Individual	Pure Ability	NI	In-house	NI	NI	Yes	NI	Plaintiff	District	Yes	Selection
EEOC v. Akal Security	Female	NI	Gender	Civil	Multiple	Pure Ability	NI	In-house	NI	NI	Yes	Yes	Plaintiff	District	NI	Selection
EEOC v. Lyon-Dell	Male	NI	ADA	Industrial	Individual	Work Sample	NI	Consultant	NI	Yes	No	NI	Defendant	District	NI	Selection
Ellis v. Chertoff	Female	Afro American	Race	Industrial	Individual	Work Sample	Yes	Consultant	NI	NI	No	Yes	Defendant	District	NI	Selection
Eudy v. City of Ridgeland	Male	NI	Retaliation	Civil	Individual	Work Sample	NI	NI	NI	NI	Yes	NI	Plaintiff	District	NI	Selection
Gilbert v. Village of Cooperstown	Female	NI	Sexual Harassment	Civil	Individual	Pure Ability	NI	NI	NI	NI	Yes	NI	Split Decision (Defendant)	District	NI	Selection
Godfrey v. City of Chicago	Female	African American	Gender	Civil	Class Action	Pure Ability	NI	NI	NI	NI	Yes	Yes	Plaintiffs	District	NI	Selection
Godoy v. Habersham County	Male	Hispanic	Race	Civil	Individual	NI	NI	NI	NI	NI	Yes	NI	Defendant	District	NI	Selection
Hunter et. al. v. Santa Fe Protective Services	Female	NI	ADEA	Civil	Multiple	Pure Ability	Yes	In-house	Yes	Yes	Yes	Yes	Defendant	District	NI	Selection
Kotwica v. Rose Packing	Female	NI	ADA	Industrial	Individual	Work Sample	NI	NI	NI	NI	No	No	Defendant	Appellate	NI	Reentry
Merritt v. Old Dominion Freight	Female	NI	Gender	Industrial	Individual	Work Sample	Yes	Consultant	No	No	No	No	Plaintiff	Appellate	No	Reentry
Norman v. Healthsouth Rehabilitation	Male	NI	ADA	Industrial	Individual	Pure Ability	NI	Consultant	NI	No	No	No	Plaintiff	District	No	Selection
Spires v. Ingersoll Rand	Female	NI	ADA	Industrial	Individual	NI	NI	NI	NI	No	No	No	Defendant	District	No	Reentry
Starkey v. City of Burnsville	Female	NI	ADA	Civil	Individual	Work Sample	Yes	Consultant	NI	No	Yes	No	Defendant	District	No	Reentry
Turner v. Arkansas Children's Hospital	Male	African American	Race	Industrial	Individual	Work Sample	No	In-house	No	No	No	Yes	Defendant	District	No	Selection
United States v. Commonwealth of Massachusetts	Both	NI	Gender	Civil	Individual	Pure Ability	NI	NI	NI	NI	Yes	NI	Settlement	District	Yes	Selection

Case	Gender	Race	Lawsuit	Work	Action	PAT Type	Stnd	In-House/ Consultant	Validated	Practice	Safety	Additional Tests	Verdict	Level	Job Analysis	Selection Type
Vasich v. City of Chicago	Female	NI	Gender	Civil	Class Action	Work Sample	NI	NI	NI	NI	Yes	NI	Settlement	District	NI	Selection

Appendix C: Summary of Findings

Lawsuit Basis	Type of Work	PAT Type	In-house or Consultant	Position Involves Public Safety	Verdict	Was Job Analysis Performed?
5 Age (11.4%)	8 Industrial (18.2%)	21 Pure Ability (47.7%)	9 Consultant (20.5%)	7 No (15.9%)	11 Plaintiff (25%)	7 Yes (15.9%)
17 Gender (38.6%)	36 Civil (81.8%)	14 Work Sample (31.8%)	9 In-House (20.5%)	37 Yes (84.1%)	29 Defendant (65.6%)	5 No (11.4%)
11 ADA (25.0%)	0 Professional (0.0%)	10 Other (22.7%)	26 No Information (59.1%)			32 No Information (72.3%)
8 Race (18.2%)						
3 Other (6.8%)						

References

Arvey, R. D., Nutting, S. M., & Landon, T. E. (1992). Validation strategies for physical ability testing in police and fire settings. *Public Personnel Management, 21,* 301-311.

Chicago Regional Council of Carpenters, United Brotherhood of Carpenters and Joiners of America, and Rodolfo Rosas, Sr. v. Thorne Associates., 893 F. Supp, 2d 952 (E.D. Il. 2012).

Davis v. CDA Incorporated., 1:09 CV-406-WKW (M.D. S. Al. 2010).

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

Equal Employment Opportunity Commission, & Equal Employment Opportunity Commission. (1978). Uniform guidelines on employee selection procedures.*Federal register*, *43*(166), 38295-38309.

Eison v. Knoxville (1983) 570 F. Supp. 11.

Goldstein, I. L., & Ford, J. K. (2002). *Training in organizations* (4th ed.). Belmont, CA: Wadsworth Publications

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

Gutman, A., Koppes, L. L., & Vodanovich, S. J. (2011). *EEO law and personnel practices* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Hardy v. Stumpf (1978) 567 P. 2d 1342.

Hollar, D. E. (2000). Physical ability tests and Title VII. *The University of Chicago Law Review, 67,* 777-803.

Hoover, L. T. (1992). Trends in police physical ability selection testing. *Public Personnel Management, 21,* 29-40.

Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. *International Journal of Selection and Assessment, 9,* 152-194.

International Guards Union of America, Local #106 v. C&D Security, Inc., 07-CV-523 MCA/ACT (D. N Me. 2007).

Lanning v. SEPTA., 181 F. 3d 478; (U.S. App. 1999).

Lyons, B. D., Hoffman, B. J., Michel, J. W., & Williams, K. J. (2011). On the Predictive Efficiency of Past Performance and Physical Ability: The Case of the National Football League. *Human Performance*, *24*(2), 158-172.

Kotwica v. Rose Packing Company, Inc., F.3d 744 (U.S. App. 2010).

Maher, P.T. (1984). Police physical ability tests: Can they ever be valid? *Public Personnel Management Journal,* 173-183.

McDaniel, M. A., Kepes, S., & Banks, G. C. (2011). The uniform guidelines are a detriment to the field of personnel selection. *Industrial Organizational Psychology, 4,* 494-514.

Norman v. Healthsouth Rehabilitation Centers of Louisville., 06-465-C (W.D. Ke. 2008).

Picano, J. J., Williams, T. J., & Roland, R. R. (2006). Assessment and selection of high-risk operational personnel. *Military Psychology: Clinical and Operational Applications, Guildford Press, New York*, 353-70.

Ryan, A. M., Greguras, G. J., & Ployhart, R. E. (1996). Perceived job relatedness of physical ability testing for firefighters: Exploring variations in reactions. *Human Performance, 9,* 219-240.

Sackett, P. R., & Lievens, F. Personnel selection, (2008). *Annual Review of Psychology*, 59, 419-450.

Sady, K., & Dunleavy, E. (2013). Exploring the uniform guidelines on employee selection decisions. *OFCCP Blog Spot.* Retrieved from http://dciconsult.com/exploring-the-uniform-guidelines-on-employee-selection-procedures/

Shoenfelt, E.L., & Pedigo, L.C. (2005). A review of court decisions on cognitive ability testing, 1992-2004. *Review of Public Personnel Administration, 25,* 271-287.

Sothmann, M.S., Gebhardt, D.L., Baker, T.A., Kastello, G.M., & Sheppard, V.A. (2004). Performing requirements of physically strenuous occupations: validating minimum standards for muscular strength and endurance. *Ergonomics, 47,* 864-875.

Spires v. Ingersoll Rand and Trane U.S. Inc., C/A No.: 3:11-2530-TLW-SVH (D. S. Ca. 2013).

Starling, P.G. (2006). Physical ability testing: a review of court cases 1992-2006 (Unpublished master's thesis). Western Kentucky University, Bowling Green, KY.

United States of America v. The Commonwealth of Massachusetts., 781 F. Supp. 2d 1 (U.D. Ma. 2011).

Vasich v. City of Chicago., 11 C 04843 (N.D. Il. 2013).

Wax, A. L. (2011). Disparate Impact Realism. *Wm. & Mary L. Rev.*, *53*, 621-714.

Wards Cove Packing Company v. Atonio (1989) 490 U.S. 642.

Werner, J. M., & Bolino, M. C. (1997). Explaining US courts of appeals decisions involving performance appraisal: Accuracy, fairness, and validation. *Personnel Psychology*, *50*(1), 1-24.

Williams, K.Z., Schaffer, M.M., & Ellis, L.E. (2013). Legal risks in selection: an analysis of processes and tools. *Journal of Business and Psychology*, 28, 401-411.