Legal Background Information Memorandum

Privileged & Confidential
Attorney-Client Communication

Date: 2012

Re: Pre-Employment Tests

Pre-Employment Tests and Disparate Impact

The EEOC has made it known that it has heightened its scrutiny of pre-employment tests in the past few years. Most recently, veterans' advocates also have questioned the tests' impact on returning veterans, especially in light of the Uniformed Services Employment and Reemployment Rights Act ("USERRA") and record deployments over the last 12 years.

Prior to the heightened scrutiny, the past history of appellate level pre-employment testing cases calls into question much of the emphasis on more generalized pre-employment testing that have not been tailored for a specific job. Many of the past cases demonstrate that such testing may be discriminatory. Since the tests must be shown to be "job related" and have a "business necessity" for each individual job, they can be very costly to validate and ultimately to legally defend.

Legal History

A disparate impact claim in the hiring process requires the plaintiff to demonstrate that a particular employment selection process causes a discriminatory impact on a race, gender, age, or disability. The same standard would be applicable under the USERRA for returning veterans. After a plaintiff demonstrates a disparate impact, the burden is on the defendant to demonstrate that the selection process is "job related" and "consistent with business necessity." *See* 42 *U.S.C.* § 2000e-(k)(1)(A)(i). If the defendant makes such a showing, the plaintiff can still prevail by demonstrating that an alternative selection process exists and the employer refuses to adopt the process. *See* 42 *U.S.C.* § 2000e-2(k)(1)(A); *see also Albemarle Paper Co. v. Moody,* 422 U.S. 405, 425 (1975). "Where two or more selection procedures are available which serve the

user's legitimate interest…, and which are substantially equally valid for a given purpose, the user should use the procedure which has been demonstrated to have the lesser adverse impact." 42 *U.S.C.* § 2000e-2(k)(1)(A); 29 *C.F.R.* § 1607.3(B).

Under the EEOC guidelines within 29 *C.F.R.* § 1607, any procedure having an adverse impact on a protected class constitutes discrimination unless that discrimination is justified. The use of any selection procedure which has an adverse impact on the hiring, promotion, or other employment or membership opportunities of members of any race, sex, or ethnic group will be considered to be discriminatory and inconsistent with the EEOC guidelines, unless the procedure has been validated in accordance with the guidelines.

Many of the generic pre-employment tests fail to meet the EEOC guidelines because they have not been found to be specific for the job at issue. "Tests used in the employee selection process must measure the person for the job and not the person in the abstract." *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971). The uniform guidelines on Employee Selection Procedures adopted by the [EEOC] provide that any test must be validated against the job or jobs for which it is a prerequisite. *See* 29 *C.F.R.* 1607.5(B). Likewise, Title VII requires that evidence offered to show the business necessity of an employment practice directly address the necessity of the practice for the particular job for which it is utilized. *See* 42 *U.S.C.* § 2000e-2(k)(1)(A)(i). Many legal scholars have written that the most common problem in a disparate impact case with a test is the failure to have the test professionally validated for the position in which it is intended.

EEOC Scrutiny

EECO release a Fact Sheet in 2007 on "Employer Best Practices for Testing and Selection. That fact sheet includes the following:

> If the selection procedure disproportionately screens out a protected group, the employer should ensure that the procedure is properly validated for the particular position in question and purposes for which it is used. The test or selection procedure must be job-related for the particular employer's purpose. While test vendor's documentation supporting the validity of a test may be helpful, the employer is still responsible for ensuring that its tests are valid under the Uniform

Guidelines on Employee Selection Procedures under Title VII for the particular position(s) in its workplace.

Even if the employer can validate (i.e., demonstrate that the selection procedure is job-related) the selection procedure for the particular position in question, it should determine whether there is an equally effective alternative selection procedure that has a less adverse impact and, if so, adopt the alternative procedure.

Further recommendations include ensuring that a test or selection procedure remains predictive of success in a job, employers should keep abreast of changes in job requirements and should update the test or selection procedures accordingly.

Employers should ensure that the tests and selection procedures are not adopted casually by managers who know little about these processes. A test or selection procedure can be an effective management tool, but no test or selection procedure should be implemented without an understanding of its effectiveness and limitations for the organization, its appropriateness for a specific job, and whether it can be appropriately administered and scored.

The EEOC's fact sheet included the following examples of litigation in pre-employment tests:

Title VII and Cognitive Tests**:** Less Discriminatory Alternative for Cognitive Test with Disparate Impact. *EEOC v. Ford Motor Co. and United Automobile Workers of America*, involved a court-approved settlement agreement on behalf of a nationwide class of African Americans who were rejected for an apprenticeship program after taking a cognitive test known as the Apprenticeship Training Selection System (ATSS). The ATSS was a written cognitive test that measured verbal, numerical, and spatial reasoning in order to evaluate mechanical aptitude. Although it had been validated in 1991, the ATSS continued to have a statistically significant disparate impact by excluding African American applicants. Less discriminatory selection procedures were subsequently developed that would have served Ford's needs, but Ford did not modify its procedures. In the settlement agreement, Ford agreed to replace the ATSS with a selection procedure, to be designed by a jointly-selected industrial psychologist, that would predict job success and reduce adverse impact. Additionally, Ford paid $8.55 million in monetary relief.

Title VII and Physical Strength Tests**:** Strength Test Must Be Job-Related and Consistent with Business Necessity If It Disproportionately Excludes Women. In *EEOC v. Dial Corp.*, women were disproportionately rejected for entry-level production jobs because of a strength test. The test had a significant adverse impact on women – prior to the use of the test, 46% of hires were women; after use of the test, only 15% of hires were women. Dial defended the test by noting that it looked like the job and use of the test had resulted in fewer injuries to hired

> workers. The EEOC established through expert testimony, however, that the test was considerably more difficult than the job and that the reduction in injuries occurred two years before the test was implemented, most likely due to improved training and better job rotation procedures. On appeal, the Eighth Circuit upheld the trial court's finding that Dial's use of the test violated Title VII under the disparate impact theory of discrimination.

Published Case Examples

Many other employers never made it through a formal process before they agreed to settle. CVS Caremark required an online personality test that was problematic when the Rhode Island ACLU filed a complaint with the state. CVS quickly and quietly agreed to eliminate much of the test.

The published appellate cases demonstrate that a generic one size fits all test will not pass scrutiny. For example, in the 2009 case of *U.S. v. City of New York*, 637 F. Supp. 2d. 77 (E.D.N.Y. 2009), a federal district court held that the test used for entry level firefighters had no business justification. The court found that the test did not actually test the abilities purported and the job analysis was deficient. The content of the test was not directly related to the job requirements. The expert who reviewed the test showed the questions used to measure a specific cognitive ability were actually correlated more highly with a different cognitive ability, so the test was not representative of the content or the procedures required for the job.

In the case of *Isabel v. City of Memphis*, 404 F.3d 404 (2005), the court found the cut off score on a written test developed by an industrial psychologist did not "approximate a candidate's potential job performance" and therefore was nothing more than an arbitrary decision that did not measure minimum qualifications.

In the case of *Melendez v. Illinois Bell Telephone*, *Not Reported in F. Supp*. 1992 WL 182234 (N.D. Ill.), the court declined to grant Illinois Bell's summary judgment motion because there was a prima facie case for adverse impact of a pre-employment test. The test at issue in the case was the Basic Scholastic Aptitude Test ("BSAT"), which was intended to measure math and verbal aptitudes. The plaintiff argued that the test had no correlation to the particular position for which he was applying (Manager of Urban Affairs) and it had minimal predictive value; there were less discriminatory means available for selecting qualified applicants; and the test

significantly under predicts the performance of Hispanic applicants and applicants older than 40. The validation study that was made 10 years earlier was not correlated to overall job performance for Hispanics or Whites.

A Connecticut case involving the Department of Corrections ("DOC") officer pre-employment physical fitness exam found the exam discriminated against women. *See Eatherling v. State of Connecticut*, 783 F. Supp. 2d 323 (2011). The DOC attempted to defend the exam by pointing to its validity for the positions of State Trooper Trainee and Public Safety Trainee and a number of cases that aggregated the data on different positions. However, the court rejected that argument because the other departments hire separately from the DOC and therefore the three positions were not comparable for the purposes of the tests validity.

The *Eatherling* court specifically scrutinized the DOC cited cases in support of aggregating data across a number of positions for a one size fits all test. The court explained that the cases were not comparable because the cited cases were all related to similar positions within the same departments. For example, the cases of *Paige v. California*, 291 F.3d 1141 (permitting the aggregation of data on different promotional exams within the California Highway Patrol and permitting the aggregation of data on all non-white officers within the California Highway Patrol); *Pietras v. Board of Fire Comm'rs of Farmingville*, 180 F.3d 468, 474 (2d Cir. 1999) (permitting inclusion of results of full female firefighter on physical agility test with results of probationary female firefighters within the same fire department); *Eldredge v. Carpenters 46 N. California Counties Joint Apprenticeship and Training Comm.*, 833 F.2d 1334, 1339 (9th Cir. 1987) (aggregating admission rates to a single apprenticeship program over nine years of applications); *Lilly v. Harris-Teeter Supermarket*, 720 F.2d 326, 336 n.17 (4th Cir. 1983) (finding it preferable to aggregate data across years to increase statistical power and detect an ongoing "pattern or practice"); *Eison v. City of Knoxville*, 570 F. Supp. 11, 13 (E.D. Tenn. 1983) (pooling multiple administrations of a physical fitness test by a single police department). In the end, the court rejected the "DOC's contention that aggregation is appropriate where the same screening mechanism happens to be used by multiple employers hiring for different jobs."

The court found that the DOC presented no evidence showing that the exam was predictive of who can perform the essential physical functions of the job of a Corrections Officer ("CO"). All three of the defendant's experts admitted that they did not empirically demonstrate that the cut scores used by the DOC for the exam reliably predicted an individual's performance on particular job tasks as a CO.

In the case of *Bert v. AK Steel Corporation*, No. 1:02-CV-467 (2008) is but one example of the excessive costs in defending and litigating pre-employment tests. In approving a class action settlement the court cited the fact that the case had dragged on for four years, the plaintiffs fees and expenses had added up to $750,000 and they had not even reached a conclusion about whether the test had been properly validated. AK Steel agreed to pay the class members over $500,000.

Specific Types of Generalized Tests

Companies creating more generalized tests often claim broader applicability than can be legally validated for specific job duties. This is especially true for tests that claim to measure attitudes and personality traits which are alleged to be associated with safety and job performance. In some instances the tests may demonstrate some validity for a more labor intensive position, while having no validity for a sedentary office position where there is really no opportunity for the employee to engage in any injury prone behavior. In addition, a test found valid in 1990, may have a different outcome in 2012. In 2012 the employee pool has grown substantially with many more regional diverse cultures, physical and developmental disabilities and diverse perspectives that may seriously skew general validity studies of prior years. Under the EEOC guidelines, each employer has a responsibility to perform validity studies for the specific positions. Therefore, it is very hard to find a test applicable to all settings under all circumstances and trust that it is in true compliance.

Some of the generalized testing mimics questions, theories and methodology from psychological tests such as the Minnesota Multiphasic Personality Inventory ("MMPI"). The problem with tests such as the MMPI is that they questions, theories and methodologies have proven to be flawed over time. The issues with such testing is apparent in the MMPIs history. One of the

many flaws was that the original MMPI was developed in the early 1940s where the control group for the original test consisted of a small number of white, married, Midwesterners.

Over the years tests such as the MMPI have been criticized for the consistent disparity in scoring in racial minority populations. Presumably some of those same criticisms are applicable to those with physical and mental disabilities along with a possible veteran population. Some experts surmise that the differences in scoring reflects overriding experiences and socioeconomic variations unique to specific demographics, yet have little relation to any type of psychopathology or predictability of any specific unsafe or performance behavior.

Generalized questions such as the desire to be an entertainer, a passion for driving a car, or a perspective about how one feels about people in general may be viewed much differently in diverse populations and therefore will fail to properly measure things such as impulse control.

Therefore, the lack of a proper validation study for a specific demographic pool of applicants for a specific employer trying to fill specific positions is legally problematic. The various cultural or disability demographics have grown over the years so words such as "power" may easily be conflated with freedom, or taken as meaning the power of economic self determination, thereby rendering many of the answers meaningless. Furthermore, attempting to measure "impulse control" may draw the most serious scrutiny by those protecting employees with mental and developmental disabilities who have already had there share of quiet legal victories with similar tests.

More recently, certain questions may have a skewed result for returning military veterans. Twelve years of the largest deployment of citizen soldiers since World War II, provides a larger and more distinct pool of potential employment applicants that were never factored in many tests over the years. Many of these applicants spent years in a war that cannot be compared to any others, where they were taught the only way to keep themselves and their comrades alive was to break all the rules. From a returning veteran's perspective some very mundane things may legitimately give them a feeling of power – to save their lives. It doesn't necessarily mean that returning veterans are any less safety conscious than others, it just means they may have a very

different perspective. Soldiers with multiple deployments may perceive certain issues differently, especially as they relate to safety because of the way they were required to conduct themselves.

Tests similar to the MMPI have successfully been viewed as an invasion of privacy and a prohibited psychological test, especially for those with mental illnesses and/or developmental disabilities. Target Corporation paid out over $2 million for its psychological test when the plaintiff claimed the test probed into his private thoughts and deepest feelings that were not job related. *See Saroka v. Dayton Hudson*, 235 Cal. App. 3d 654, (1991). In *Karraker v. Rent-A-Center*, 316 F. Supp.2d 675 (CD Ill. 2004), the court found alleged violations of the ADA for its personal interests and personality trait tests. Rent-A-Center was not successful in arguing that their test simply measured whether someone works well in groups or is comfortable in a face paced office. The court agreed with the plaintiffs, that the test was a pre-employment medical examination violated the ADA. More importantly, the court noted that the use of the test was prohibited because it measured traits that are not relevant to the specific job position applied for.

While some tests with psychological components may be valid for positions in law enforcement because of their unique primary duty of public safety, those same tests would not be appropriate in most occupations or for most job positions. *See Miller v. City of Springfield*, 146 F.3d 612 (8th Cir. 1998) (Testing upheld for law enforcement applicants as a business necessity because officers are directly in charge of public safety.)

Target for Costly Litigation

Prior to the time of the EEOC's stricter scrutiny, Larry R. Seegull and Emily J. Caputo wrote a summary of the pitfalls of pre-employment testing for the American Bar Association in January 2006. Their summary provides a good insight into potential for future problems. They argue that such tests will provide an increase in the potential for litigation and therefore employers should consider avoiding the use of tests. They emphasize that "the failure to ensure the tests are statistically valid, reliable and devoid of culture and ethnic bias could be costly." Tests must be job-related and of a business necessity.

For example, Seegull and Caputo emphasize that tests measuring areas such as extraversion might be justified for "people in sales, but not for people seeking positions that do not focus heavily on interaction with others. Similarly, an integrity test might be justifiable for financial and security positions but not for some other positions. The failure to monitor workplace statistics on injuries, attrition, theft, turnover and production to determine whether the use of the tests has resulted in their intended purpose in successful employee selection may result in liability for the test's discriminatory impact.

The ABA article's authors are not alone in their concerns about the problems with the testing. Most legal experts have valid concerns that tests do not provide enough benefit in the screening process to outweigh the fact they many times they are simply an easy target for lengthy and expensive litigation. Margaret Spence, a workers compensation consultant, joked that for "some reason the workers compensation industry keeps trying re-invent the mouse trap and employers are the ones who keep getting caught over and over again." Spence cited a recent case that highlighted her fear. "An employer was sold on the "new thing" to prevent accidents and workers compensation claims, a "fit for duty test" that did not evaluate the true picture of the job duties. As a result of the test, and contrary to previous experience in successful hiring of qualified candidates, the employer experienced a significantly reduced female applicant pool. The test resulted in years of protracted litigation that concluded with the employer appealing to the U.S. Court of Appeals, losing and subsequently paying millions of dollars in damages and agreeing to embark on developing nondiscriminatory hiring criteria.

The EEOC administrative reviews, state human rights reviews, and courts are filled with decades of the most expensive and time consuming litigation related to pre-employment testing. Those disputes always involve numerous studies, experts reviews and plaintiff attorney's attempting to mitigate their enormous expenses by finding various ways to increase the overall monetary damages.

Times have changed since the initial creation of many pre-employment tests. The EEOC has made changes in its priorities in the past few years and most recently created an agenda that includes the scrutiny of recruitment and hiring practices. Specifically, the EEOC has released its

agenda stating that it will "target class-based recruitment and hiring practices that discriminate against racial, ethnic and religious groups, older workers, women and people with disabilities." That agenda goes on to include placing a "special emphasis on emerging issues in equal employment law including demographic changes, developing theories." The law firm of Barnes & Thornburg LLP wrote: "Employers should expect to see a very active EEOC…"

Just as importantly, the EEOC and other related governmental agencies have the resources to fully audit and attack the tests for bias, which may be found in the most unexpected places. Under those circumstances, even if the employer proves the tests' validity the employer may lose because of the cost of defending its usage.

Other Legal Scrutiny

In addition to the laws previously mentioned, there are state privacy and human rights laws that include the protection of additional class of individuals which may not be included in a broad based validity study. The legal news also reports the National Labor Relations Board ("NLRB") as being very active in scrutinizing company policies and practices that have a chilling effect on employee rights to collectively organize. The NLRB has looked at all aspects of the employer/employee relationship, and of late has a special emphasis on company rules and policies that tend to be an impediment to an employee's right to organize and bargain collectively. Tests that market themselves as measuring whether a candidate is likely to follow company rules and policies would fall perfectly under the NLRB's current mission to eliminate any employment practices that have a chilling effect on the employee's right to challenge such rules and policies under the National Labor Relations Act. As many legal experts have stated, the current Administration's National Labor Relations Board has overturned over 50 years of precedent by attacking the most well established employment policies and rules as being a violation of the National Labor Relations Act.

Yet another layer of concern is for any employer with a federal contract. Most recently the Office of Federal Contract Compliance Programs "OFCCP" has taken the cue from the EEOC and began audits of companies for its hiring practices. The OFCCP recently filed a complaint against Leprino Foods about a pre-employment test that adversely effected minority applicants

and was not shown to be valid or job related. Leprino could not substantiate its claim that the test measuring math, workplace observation and information location skills were related to the essential functions of the applied-for position of a laborer. The OFCCP sought at least 17 job offers for the original applicants, back wages and interest for 270 class members. In addition, OFCCP sought the cancellation of federal contracts and the company debarment of future government contracts. The complaint was filed in May and the company was at the settlement table by July with a $550,000 check.

Similar incidents took place with the OFCCP with complaints against Kraft Foods and Gerber Products stemming from pre-employment tests used which had an adverse impact on a class of applicants. The OFCCP found there was insufficient evidence of validity for the specific positions at issue and subsequently settled for $227,500 and $900,000 respectively.

The May 2012 Newsletter of Vedder Price wrote, "The message is also clear to those in the contractor community: Tests used during the hiring process will be scrutinized during a compliance review, and tests that adversely impact minority candidates may lead to OFCCP findings of discrimination and substantial liability if the contractor cannot demonstrate that skills tested are required to perform the job in question and are properly vetted through a validation study."

Conclusion

As the Vedder Newsletter warns, while the EEOC's uniform guidelines on pre-employment tests allow an employer to use a test based on the publisher's validation studies, the guidelines emphasize that an employer "must show that the skill requirements and job responsibilities of the jobs used in the test publishers general validity studies are sufficiently similar to its own job skill requirements and duties." Yet, even having a specific validation study does not protect against liability for a pre-employment test because "an employer who uses a test with any type of an adverse impact must investigate whether there is an alternative process that meets the business needs by a less discriminatory process."

Each employer can easily find themselves with costly legal issues if they fail to ensure a test's validity for each individual job position. The validity of a test will never remain stagnant

because there are changing job requirements and an ever changing workforce. Considering all the information cited above, a generic pre-employment test is ripe for all types of legal scrutiny and thus the risks for liability would seem to outweigh any rewards for its screening potential.