PACIFIC COAST LABOR & EMPLOYMENT CONFERENCE
Seattle, WA
April 28, 2017

LEGAL ISSUES IN THE USE OF PRE-EMPLOYMENT TESTING

Christine E. Webber[1]
Cohen Milstein Sellers & Toll PLLC
1100 New York Ave., NW Suite 500
Washington, DC 20005
(202) 408-4600
cwebber@cohenmilstein.com

I. BACKGROUND ON PRE-EMPLOYMENT AND PROMOTIONS TESTING

A. A test is "a standardized series of problems or questions that assess a person's knowledge, skills, abilities, or other characteristics," in this context for use in hiring or promotion decisions. http://www.siop.org/workplace/employment%20testing/overview.aspx.

B. Pre-employment testing, which dates back to 1919, has become a prominent tool in the employment sector as a way to help organizations decide systematically—and, in theory, more accurately—which applicants will perform well in the job and remain with the company. To this end, testing has the capacity to be a cost-effective, consistent, and potentially useful tool, if the test is developed and applied accurately.

C. However, correlation rates seen today are no better than those found in the 1960s, suggesting that even with the social-science developments over the past 50-plus years, we've been unable to enhance workplace-predictive measures. A study published by the Cornell HR Review found, for example, that personality tests accounted for only about 5% of employment success. *See* http://www.cornellhrreview.org/personality-tests-in-employment-selection-use-with-caution/; *see also* Frederick P. Morgenson et al., *Reconsidering the Use of Personality Tests in Personnel Selection Contexts*, 60 Personnel Psychol. 683, 694-96 (2007) (discussing poor validity rates for personnel tests).

D. Despite the limited support for test usefulness, there has been an increase in pre-employment testing used by employers. The EEOC reported an increase in testing due to increased security concerns after 9/11, and increased concerns about workplace violence scares, as well as the proliferation of online application procedures which generate larger applicant pools.

---

[1] Miriam Nemeth of Cohen Milstein provided substantial assistance with preparation of this paper.

https://www.eeoc.gov/policy/docs/factemployment_procedures.html

E. Pre-employment tests also have the potential to serve as discriminatory screening-out tools to prevent minorities, women, disabled individuals, and older workers from participating meaningfully in our economy and must be carefully scrutinized to ensure that they provide equal opportunities for employment to all applicants.

F. Types of Pre-Employment Tests

1. Physical tests, such as ability tests for firefighters, hearing tests for commercial drivers, or drug tests used in positions affecting public safety.

2. Cognitive tests, which are often administered in writing, see, e.g., *Kerner v. City & Cty. of Denver*, No. 11-00256, 2015 WL 5698663, at *1, *3-4 (D. Colo. Sept. 29, 2015) (evaluating the alleged disparate impact of *a* test), but which may also include simulations or assessment centers.

3. Personality tests have been developed more recently. For example, these tests may measure honesty, dependability, stress tolerance, approaches to conflict resolution, emotional stability, openness to new experiences, and other traits related to workplace interactions or working styles.

a. These tests include formal measures, such as the NEO-Personality Inventory, Personality Characteristics Inventory, Hogan Personality Inventory, as well as tests commonly used in clinical mental health settings, such as the Minnesota Multiphasic Personality Inventory (MMPI). http://www.cornellhrreview.org/personality-tests-in-employment-selection-use-with-caution/

b. They also include employment selection tools developed by IO psychologists and companies to assist in employment selection. See, e.g., https://www.theguardian.com/science/2016/sep/01/how-algorithms-rule-our-working-lives; http://www.nytimes.com/2013/04/21/technology/big-data-trying-to-build-better-workers.html; Wonderlic.com; Michael A. McDaniel et al., Use of Situational Judgment Tests to Predict Job Performance: A Clarification of the Literature, 86 J. of Applied Psychol. 730, 731-32 (2001) (discussing different types of tests developed for the employment sector); Frederick P. Morgenson et al., supra I.C at 707 (discussing concerns with integrity tests).

II. LEGAL CHALLENGES TO POTENTIALLY DISCRIMINATORY TESTS

A. Pre-Employment Tests may be challenged under Title VII, the Americans with Disabilities Act (ADA), and the Age Discrimination in Employment Act (ADEA). Since the claims available to a litigant under each statute are similar, this paper will discuss the prototypical claims available under Title VII and will then discuss

special issues in ADA and ADEA testing claims.

B. Plaintiffs may challenge potentially discriminatory tests under both disparate impact and disparate treatment theories. Disparate impact arguments are substantially more common under Title VII and ADA screen-out claims, because plaintiffs are not required to demonstrate intent on the part of the employer. (Until recently, disparate impact theories were not available under ADEA claims, and disparate treatment arguments remain common under this statute. *See infra* Section V.B.)

C. The EEOC issued Uniform Guidelines on Employee Selection Procedures ("UGESP" or "the Uniform Guidelines") which address in depth the validation of tests or other selection procedures.

D. Defendants:

1. Ordinarily the entity that creates a test is not the proper defendant for suit, without some allegations showing an employment connection between the applicant and the test-creating entity or the test-creating entity and the employer with which the applicant is seeking to be hired. *See Chi. Reg'l Council of Carpenters v. Pepper Constr. Co.*, 32 F. Supp. 3d 918, 922-24 (N.D. Ill. 2014).

2. However, because of the high degree of control that the state may have over local school districts with respect to teacher hiring, state licensing exams for teachers have been subject to Title VII challenge. *Association of Mexican-American Educators v. California*, 231 F.3d 572, 580-81 (9th Cir. 2000).

3. In addition, of course, an employer can be held liable where it chooses to use a test with a disparate impact, even though the test was developed by a third-party consultant. *See James v. Stockham Valves & Fittings Co.*, 559 F.2d 310, 319, 339-40 (5th Cir. 1977).

III. DISPARATE IMPACT CHALLENGES

A. Background

1. The Supreme Court initially recognized that employment tests could be challenged under Title VII using a disparate impact theory in *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971); *see also Watson v. Fort Worth Bank & Tr.*, 487 U.S. 977, 987-88 (1988) (collecting cases); *Carroll v. Sears, Roebuck & Co.*, 708 F.2d 183, 189 (5th Cir. 1983) (employment tests are the "kind of employment practice to which the disparate impact model traditionally has applied").

2. Unlike disparate treatment claims, which require a showing of discriminatory intent, disparate impact claims do not require such a

showing, and instead focus on facially neutral practices which nonetheless have an adverse impact.

3. Title VII was subsequently amended to codify the approach articulated in *Griggs*. *See* 42 U.S.C. § 2000e-2(k)(1) *et seq.*

4. Disparate impact claims are tried to the Court, not a jury. 42 U.S.C. § 1981a(c).

B. Burden-shifting Framework. *See Albemarle Paper Co. v. Moody*, 422 U.S. 405, 425 (1975)

1. Plaintiffs demonstrate adverse impact of a specific employment practice on their protected class.

   a. Plaintiffs may challenge the adoption of a practice that has a disparate impact or the subsequent application of that practice to them. *See Lewis v. City of Chicago*, 560 U.S. 205, 214-15 (2010).

   b. Adverse impact must be statistically significant and requires the use of an expert. *See Kerner v. City & Cty. of Denver*, No. 11-00256, 2015 WL 5698663, at *3-4 (D. Colo. Sept. 29, 2015) (including a helpful discussion of statistical needs for disparate impact claims).

   c. Although the EEOC Guidelines reference a "4/5 rule of thumb," courts today more commonly rely on a formal test of statistical significance to determine if a test has a disparate impact. The four-fifths rule states that if the pass rate of the challenging group is less than four-fifths the pass rate of the majority group, then the test has a disparate impact on the group challenging the test.

2. Once adverse impact is demonstrated, which is usually fairly straightforward, the burden shifts to defendant to demonstrate that the challenged practice is "job-related" and "consistent with business necessity." 42 U.S.C. § 2000e-2(k)(1)(A)(i). If defendant meets that standard, the burden shifts back to plaintiffs.

   a. In *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971), the Supreme Court rejected the argument that a test was lawful simply because it was professionally developed. The Court instead held that § 703(h) requires that a test be "job related" in order to meet Title VII's nondiscrimination requirements. *Id.* at 435-36.

   b. *Griggs* interpreted Title VII and its implementing regulations as requiring a nexus between the test and the specific position at issue. 401 U.S. at 435-36 ("What Congress has commanded is that any tests must measure the person for the job and not the person in

the abstract.").

c. Defendants may show that a test is job-related by demonstrating the test's validity, i.e., that the test is predictive of success in the job position at issue. *See infra* Section VI; *cf. Albemarle Paper Co. v. Moody*, 422 U.S. 405, 425 (1975) (requiring "a manifest relationship to the employment in question"); *Nash v. Consol. City of Jacksonville, Duval Cty. Fla.*, 837 F.2d 1534, 1537 (11th Cir. 1988).

d. Business necessity goes beyond "job-relatedness" and requires a showing that there is a business reason that makes use of the test necessary. *Hamer v. City of Atlanta*, 872 F.2d 1521, 1533 (11th Cir. 1989); *Donnelly v. R.I. Bd. of Governors for Higher Educ.*, 929 F. Supp. 583, 593 (D.R.I. 1996), *aff'd on other grounds*, 110 F.3d 2 (1st Cir. 1997) (explaining that "'consistent with business necessity' requires something less than a showing that the challenged practice is essential to the conduct of the employer's business but something more than a showing that it serves a legitimate business purpose. What it appears to require is proof that the challenged practice is reasonably necessary to achieve an important business objective.")

e. Whether a test is valid is commonly a highly contested issue, and expert testimony is needed to resolve that issue.

3. If the court finds defendant's burden is satisfied as to job-relatedness and business necessity, then the burden shifts to plaintiffs, who must demonstrate an alternative employment practice that satisfies the business need with less adverse impact than the challenged practice. See 42 U.S.C. §§ 2000e-2(k)(1)(A)(ii), 2(k)(1)(C).

a. The alternative should be "equally" or "substantially equally" effective as the challenged test. *See Wards Cove Packing Co. v. Atonio*, 490 U.S. 642, 661 (requiring that the practice be "equally effective" as the challenged practice); *Johnson v. City of Memphis*, 770 F.3d 464, 472 (6th Cir. 2014) (requiring "substantially equally valid results, but with . . . less discriminatory outcomes" (internal citations and quotation marks omitted). What this means in terms of proof and the range of factors to be considered remains an open question. *Cf. Watson v. Ft. Worth Bank & Tr.,* 487 U.S. 977, 998 (1988)(plurality) ("Factors such as the cost or other burdens of proposed alternative selection devices are relevant in determining whether they would be equally as effective as the challenged practice in serving the employer's legitimate business goals.").

b. An expert is likely necessary for this step as well, although the

tests examined and rejected by the expert may be discoverable. Fed. R. Civ. P 26(b)(4)(B). As such, some litigants prefer to hire separate consulting and testifying experts. *See* 1 Barbara T. Lindemann & Paul Grossman, Employment Discrimination Law 222-23 (4th ed. 2007).

c. Even a test that is concededly job related and consistent with business necessity may be successfully challenged by showing less discriminatory alternatives. *See, e.g.*, *Jones v. City of Boston*, 845 F.3d 28, 34–35 (1st Cir. 2016) (plaintiff conceded that drug testing was job related, but nonetheless survived summary judgment because there was dispute of fact over whether there were satisfactory alternatives with less of a disparate impact from false positives).

C. Other Legal Issues in Disparate Impact Challenges

1. Plaintiffs may challenge the test as a whole (even where the test comprises multiple segments) or selected components of the test. They may use aggregated data, from multiple instances in which a test was used (especially if the sample size is otherwise small), or disaggregated data to make their *prima facie* showing.

a. *Connecticut v. Teal*, 457 U.S. 440, 450-51 (1982) (holding that a court cannot look at just the "bottom line" hiring results and must instead examine each challenged component of the selection process at issue).

b. "'[I]ndividual components of a hiring process may constitute separate and independent employment practices subject to Title VII even if the overall decision-making process does not disparately impact the ultimate employment decisions involving a protected group.'" *Smith v. City of Boston*, 144 F. Supp. 3d 177, 193-94 (D. Mass. 2015) (quoting *Bradley v. City of Lynn*, 443 F. Supp. 2d 145, 158-59 (D. Mass. 2006)).

c. *See Kerner v. City & Cty. of Denver*, No. 11-00256, 2015 WL 5698663, at *6 (D. Colo. Sept. 29, 2015) (citing cases and explaining that "many courts observe that aggregated statistical data may be used where it is more probative than subdivided, or disaggregated, data")

d. While prior case authority had begun to require that plaintiffs demonstrate the disparate impact of the specific element of the selection process that was challenged, the Civil Rights Act of 1991 included a provision that eased the burden on plaintiffs if they "demonstrate[d] to the court that the elements of a respondent's

decisionmaking process are not capable of separation for analysis" then plaintiffs were permitted to analyze the selection process "as one employment practice." 42 U.S.C. § 2000e-2(k)(1)(B)(i). *Fisher v. Procter & Gamble Mfg.*, 613 F.2d 527, 544-546 (5th Cir. 1980) (finding discrimination in class action challenge to promotional process that included a written multiple choice test as well as subjective evaluations).

2. Different ways to score and use tests

   a. Tests may be used simply as a pass/fail, with some other process being used to select among those who pass; scores may be placed into bands, wherein everyone in the same band is deemed essentially equivalent; or employees may be selected in rank in order of their test scores (alone or with other criteria).

   b. Plaintiffs may challenge not simply the test itself, but how the test is used – i.e. whether the "cut-score" for passing was set at the correct level, whether a test which might be valid on a pass-fail basis is valid when used to select individuals in rank order from their scores, etc. *See* 29 C.F.R. § 1607.14(C)(9); *Guardians Ass'n of the N.Y.C. Police Dep't, Inc. v. Civ. Serv. Comm'n*, 630 F.2d 79, 100-01, 105 (2d Cir. 1980); *Chicago Firefighters Local 2 v. City of Chicago*, 249 F.3d 649, 656 (7th Cir. 2001) (promoting the use of banded scores when there are relatively narrow differences in skills among similar test scores, e.g., 200 questions vs. 199 questions correct).

IV. DISPARATE TREATMENT CHALLENGES

A. While people commonly think of tests – the quintessential facially neutral practice – as being subject to disparate impact challenges, they can be challenged on a disparate treatment basis too.

1. "[A] knowing use of a facially neutral employment practice that imposes a disparate adverse impact on a protected group may help to support an inference of intentional discrimination." *Melendez v. Ill. Bell. Tel. Co.*, No. 90-5020, 1992 WL 182234, at *5 (N.D. Ill. July 24, 1992) (collecting cases), discussed approvingly, 79 F.3d at 671.

2. *Clark v. Pennsylvania*, 885 F. Supp. 694, 708-09 (E.D. Pa. 1995) (prima facie case of disparate treatment established where evidence existed that defendant was aware that promotional exam would eliminate African American candidates).

3. Plaintiffs asserting disparate treatment must demonstrate that the defendant intentionally discriminated against them. *See Melendez v. Ill. Bell Tel. Co.*, 79 F.3d 661, 669-70 (7th Cir. 1996). The Seventh Circuit

has recently described the relevant question as, for example, "whether [the defendant] had an anti-female motivation for creating its . . . test." *Ernst v. City of Chicago*, 837 F.3d 788, 795 (7th Cir. 2016).

4. This is consistent with other case law finding that disparate treatment can be established with evidence that an employer was aware of the adverse impact a particular practice would have and chose to follow the practice anyway. Intent can be inferred in such circumstances. *See, e.g.*, *Bazemore v. Friday*, 478 U.S. 385, 394-97 (1986) (involving a claim for failure to change a program with known Title VII violations); *Corning Glass Works v. Brennan*, 417 U.S. 188, 205-06 (1974) (same); *see also Personnel Adm'r of Mass. v. Feeney*, 442 U.S. 256, 279 n.25 (1979) ("When the adverse consequences of a law upon an identifiable group are as inevitable as the gender-based consequences [here], a strong inference that the adverse effects were desired can reasonably be drawn."); *Columbus Bd. of Educ. v. Penick*, 443 U.S. 449, 464-65 (1979) ("Adherence to a particular policy or practice, 'with full knowledge of the predictable effects of such adherence upon racial imbalance in a school system is one factor among many others which may be considered by a court in determining whether an inference of segregative intent should be drawn.'").

B. *Ricci v. DeStefano*, 557 U.S. 557 (2009)

1. In *Ricci*, a city became aware that their test for firefighters had an adverse impact on minority candidates, and chose to discard the test results rather than rely on them for promotions. A white firefighter who had passed the test sued, alleging that the decision not to use the test after all was an act of intentional discrimination.

2. The Supreme Court found disparate treatment because the city had not shown a strong basis in evidence for believing it would be subject to disparate impact challenge. Since an adverse impact alone does not mean that plaintiffs prevail in disparate impact cases, an attempt to validate the test would be required before an employer could have a "strong basis in evidence" to find the employer was risking disparate impact liability.

3. Although some have suggested that *Ricci* places employers in a Catch-22 position, where using flawed test results may give rise to a disparate impact claim and discarding those same results may give rise to a disparate treatment claim, experience has not been so grim. Nor has *Ricci* served to gut disparate impact challenges, as some suggested. *See, e.g., United States v. City of New York ("Vulcan")*, 637 F. Supp. 2d 77, 83 (E.D.N.Y. 2009) (noting "this case presents the entirely separate question of whether Plaintiffs have shown that the … Exams … *actually had* a disparate impact" while *Ricci* instead considered whether the defendant had shown it would likely have been liable under disparate-impact

analysis). Nothing in *Ricci* prevents an employer from taking thorough steps to ensure a test is actually valid. Indeed, *Ricci*'s holding turned on a finding that the city had ample evidence of validity. *Vulcan*, in contrast, found the city liable where it had not adequately validated the exams it used.

V. SPECIAL ISSUES UNDER OTHER STATUTES

A. Americans with Disabilities Act (ADA), 42 U.S.C. § 12112(a), (b)(6)-(7); 29 C.F.R. § 1630.10

1. Pre-employment "medical examinations" are prohibited and must wait until after a conditional offer of employment has been made. 42 U.S.C. §§ 12112(d)(2)(A), (3); 29 C.F.R. § 1630.13.

a. The EEOC defines a "medical examination" as "a procedure or test that seeks information about an individual's physical or mental impairments or health." ADA Enforcement Guidance: Preemployment Disability-Related Questions and Medical Examinations (Oct. 10, 1995).

b. Factors courts may use in deciding whether an exam is medical (one factor alone may be sufficient) – whether the test: (a) is administered by a health care professional; (b) is interpreted by a health care professional; (c) is designed to reveal a physical or mental health impairment; (d) is invasive; (e) measures an employee's performance of a task or physiological responses to performing the task; (f) is normally given in a medical setting; and (g) uses medical equipment.

c. Psychological examinations can be unlawful medical exams if they provide evidence that would lead toward a mental health diagnosis or were designed to identify a mental disorder. *See, e.g.*, *Karraker v. Rent-A-Center, Inc.*, 411 F.3d 831, 837 (7th Cir. 2005) (holding that the MMPI constituted an improper medical examination because it could identify mental illnesses and thus "has the effect of hurting the employment prospects of one with a mental disability"); *Barnes v. Cochran*, 944 F. Supp. 897, 905 (S.D. Fla. 1996) (pre-employment examination performed by psychologist and including questions from the MMPI and other psychological tests that tended to result in disclosure of psychological disabilities was unlawful pre-employment medical examination).

d. Drug tests can count as medical examinations. *See Bates v. Dura Auto. Sys., Inc.*, 767 F.3d 566, 574-88 (6th Cir. 2014).

2. Screen Out Claims: Employers cannot use selection criteria or administration methods that tend to screen out disabled individuals unless

they meet the burden shifting framework for disparate impact claims. *See, e.g.*, *Morisky v. Broward Cty.*, 80 F.3d 445, 448-49 (11th Cir. 1996); *Littlefield v. Nevada ex rel. Dep't of Pub. Safety*, 195 F. Supp. 3d 1147, 1153-60 (D. Nevada 2016) (describing each step of a "screen out" claim); *Coale v. Metro-N. R.R. Co.*, No. 09-02065, 2014 WL 211507, at *1 (D. Conn. Jan. 7, 2014).

a. The main difference between "screen out" and "disparate impact" analysis is that a "screen out" claim may not require statistical evidence of adverse impact, permitting the use of other evidence to show that persons with disabilities are screened out due to their disability. Legislative history supports this distinction. 137 Congr. Rec. 15466 (Oct. 30, 1991) (statement of Sen. Harkin). Although statistics may not be required, usually some showing of causation is. *See, e.g.*, *Rohr v. Salt River Project Agric. Imp. & Power Dist.,* 555 F.3d 850, 862 (9th Cir. 2009); *Bates v. United Parcel Serv., Inc.,* 511 F.3d 974, 994 (9th Cir. 2007); *Cripe v. City of San Jose,* 261 F.3d 877, 889-90 (9th Cir. 2001)

b. Some courts continue to require statistical evidence, often reflexively equating disparate impact claims under the ADA with those under Title VII, without consideration of the differences in the text of the two statutes. *See, e.g.*, *Femino v. NFA Corp.*, 274 F. App'x 8, 10 (1st Cir. 2008); *Champ v. Baltimore Cty.*, 884 F. Supp. 991, 995 (D. Md. 1995), *aff'd*, 91 F.3d 129 (4th Cir. 1996); *Kintz v. United Parcel Serv., Inc.*, 766 F. Supp. 2d 1245, 1254 (M.D. Ala. 2011).

c. Other courts skip this threshold showing altogether and move directly to business necessity, particularly where "it is evident that the employer relied upon a disability in making an adverse employment decision." *Jeffrey v. Ashcroft*, 285 F. Supp. 2d 583, 588 (M.D. Pa. 2003); *see also Cremeens v. City of Montgomery*, No. 09-409, 2010 WL 3153721 (M.D. Ala. Aug. 9, 2010), *aff'd*, 427 F. App'x 855 (11th Cir. 2011); *Allmond v. Akal Sec., Inc.*, 558 F.3d 1312, 1317 (11th Cir. 2009) ("We, therefore, begin by reviewing the [test] for job-relatedness and business necessity."); *Wice v. Gen. Motors Corp.*, No. 07-10662, 2008 WL 5235996 (E.D. Mich. Dec. 15, 2008); *Ethridge v. Ala.*, 860 F. Supp. 808, 820 (M.D. Ala. 1994).

3. Business Necessity Defense is modified slightly: an employer must show that "the pertinent qualification standard is job-related for the position in question, is consistent with business necessity, and cannot be met by a person with the plaintiff's disability even with a reasonable accommodation." *Bender v. Norfolk S. Corp.*, 31 F. Supp. 3d 659, 667-68 (M.D. Pa. 2014) (citing *Verzeni v. Potter*, 109 F. App'x 485, 490 (3d Cir.

2004)); 29 C.F.R. § 1630.7, 1630.15(c).

B. Age Discrimination in Employment Act (ADEA), 29 U.S.C. § 623

1. Both Disparate Impact and Disparate Treatment claims are allowed. *See Smith v. City of Jackson*, 544 U.S. 228, 236-40 (2005) (permitting disparate impact claims); *Hazen Paper Co. v. Biggins*, 507 U.S. 604, 609 (1993) (permitting disparate treatment claims).

2. Disparate Impact:

a. "The scope of disparate impact liability under ADEA is narrower than under Title VII." *Smith*, 544 U.S. at 240.

b. *Wards Cove Packing Co. v. Atonio*, 490 U.S. 642 (1989), which requires employees to identify the specific employment practices responsible for statistical disparities suggesting an adverse impact, still applies to ADEA claims. *See also Wagner v. Pac. Mar. Ass'n*, No. 05-1729, 2007 WL 2407093, at *5 (D. Or. Aug. 21, 2007) (requiring statistical evidence to defeat summary judgment on a disparate impact claim).

c. Potential Applicants: Whether job applicants may bring disparate impact claims under the ADEA remains an open question subject to potential Supreme Court review next term. *Compare Villarreal v. RJ Reynolds Tobacco Co.*, 839 F.3d 958 (11th Cir. 2016) (en banc) (cert. petition filed Feb. 6, 2017) (holding that applicants may not raise a disparate impact claim under the ADEA) *with Rabin v. PricewaterhouseCoopers LLP*, --- F. Supp. 3d ---, 2017 WL 661354 (N.D. Cal. Feb. 17, 2017) (holding that applicants were able to bring disparate impact claims under the ADEA).

d. Defenses: Reasonable Factor Other Than Age

(1) Defendants do not need to show a business necessity. *See Smith*, 544 U.S. at 243.

(2) Plaintiffs cannot respond that other reasonable methods not resulting in a disparate impact were available. *See id.* 243.

(3) There is no ADEA claim for discrimination based on too much experience. *See Jianqing Wu v. Special Couns., Inc.*, 54 F. Supp. 3d 48, 54-56 (D.D.C. 2014).

3. Disparate Treatment

a. Plaintiffs must prove that age is the "but-for" cause of the adverse employment decision. *Gross v. FBL Fin. Servs., Inc.*, 557 U.S.

167 (2009); *Zawacki v. Realogy Corp.*, 628 F. Supp. 2d 274, 280 (D. Conn. 2009) ("In a disparate treatment case, liability depends on whether the protected trait (under the ADEA, age) actually motivated the employer's decision.") (citing *Hazen Paper Co. v. Biggins*, 507 U.S. 604, 610 (1993)).

b. Most courts continue to apply the *McDonell Douglas* burden-shifting framework to these claims, with the addition of the but-for requirement. *McDonnell Douglas Corp. v. Green*, 411 U.S. 792 (1973); *Vahey v. Philips Elecs. N. Am. Corp.*, 461 F. App'x 873, 874 (11th Cir. 2012) ("Because the Supreme Court did not overrule this precedent in *Gross*, we review [the plaintiff's] claims under both *McDonnell Douglas* and *Gross*."); *Anderson v. Consol. Rail Corp.*, 297 F.3d 242, 249 (3d Cir. 2002).

VI. VALIDATION AND JOB-RELATEDNESS

A. "The term 'validity' describes the extent to which a candidate's performance on a test relates to his or her performance on the job. A test is 'valid' if a candidate's test performance can be used to make a better prediction about how well he or she will perform on the job than might be possible without the test." *United States v. City of Erie*, 411 F. Supp. 2d 524, 535 (W.D. Pa. 2005). "Validity" is "job-relatedness." *Id.*; *see also Ernst v. City of Chicago*, 837 F.3d 788, 796, 799 (7th Cir. 2016) ("Validity is the extent to which a study accurately measures what it sets out to measure.").

B. The EEOC's Uniform Guidelines on Employee Selection Procedures (29 C.F.R. § 1607.1 *et seq.*) apply to all selection procedures "used as a basis for making employment decisions" and serve as a "framework for determining the proper use of tests and other selection procedures" to assist potential employers in complying with federal antidiscrimination laws. 29 C.F.R. §§ 1607.1(B), 1607.2(B) & (C). They are the starting point for determining whether a test is "valid."

1. Although amendments were discussed in the 1990s, they were never undertaken, and the Guidelines have not been updated since their promulgation in 1978. They are not binding, but they are given "great deference" by courts. *See Albemarle*, 422 U.S. at 431; *Griggs*, 401 U.S. at 433-34.

2. In addition to the Uniform Guidelines and case law, courts also give substantial consideration to scientific standards including the Society for Industrial and Organizational Psychology (SIOP) Principles (http://www.siop.org/_principles/principles.pdf) and the Standards for Educational and Psychological Testing (jointly produced by the APA, AERA & NCME).

3. When current scientific standards (such as the SIOP Principles or APA

Standards) differ from the Uniform Guidelines, courts often follow the more current principles. *See NEA v. South Carolina*, 434 U.S. 1026 (1978).

4. The Guidelines do not apply to ADA, Rehabilitation Act, or ADEA claims, to which case law instead applies. See 1 Barbara T. Lindemann & Paul Grossman, Employment Discrimination Law 174 (4th ed. 2007).

C. Although the Guidelines receive deference, there are several notable areas in which courts have rejected the Guidelines in favor of other standards:

1. Bottom-line Defense: The Guidelines promote a "bottom line" approach to adverse impact, meaning that if the overall selection process does not adversely impact a protected group, courts do not need to examine the validity of the individual components. *See* 29 C.F.R. § 1607.4(C). This approach has been rejected by the Supreme Court. *See Connecticut v. Teal*, 457 U.S. 440, 456 (1982); *Stout v. Potter*, 276 F.3d 1118, 1122 (9th Cir. 2002) (that there was no adverse impact in the final result did not preclude an adverse impact challenge to the dispositive interview stage earlier in the process); *see also supra* Section II.D.4.

2. Four-Fifths Rule: The Guidelines articulate the four-fifths test for adverse impact, 29 C.F.R. § 1607.4(D), where disparate impact cannot be shown if the minority group was selected at a rate equal to or exceeding four-fifths the rate of the majority group. *See, e.g.*, *Smith v. City of Boston*, 144 F. Supp. 3d 177, 192 (D. Mass. 2015) (describing the rule). However, courts generally rely on this test as only a starting point. *See, e.g.*, *Cox v. City of Chicago*, 868 F.2d 217, 220 n.1 (7th Cir. 1989) ("The 80% Rule is one starting point for a disparate impact analysis."); *Evans v. City of Evanston*, 695 F. Supp. 922, 925 (N.D. Ill. 1988) (explaining that violation of this "guideline . . . does ordinarily establish a prima facie case of disparate impact discrimination"), *vacated on other grounds*, 881 F.2d 382 (7th Cir. 1989). Courts commonly rely upon showings of statistical significance rather than the four-fifths rule. *United States. v. City of New York ("Vulcan")*, 637 F. Supp. 2d 77, 97-99 (E.D.N.Y. 2009)

3. Identifying Less Discriminatory Alternatives: The Guidelines require that employers conduct an independent search for alternative, less discriminatory pre-employment tests. 29 C.F.R. §1607.3(B); *see also Brunet v. City of Columbus*, 1 F.3d 390, 412 (6th Cir. 1993) (faulting the district court for failing to determine if defendant met its burden of exploring alternatives); *Ashton v. City of Memphis*, 49 F. Supp. 2d 1051, 1061 (W.D. Tenn. 1999) (discussing the city's obligation under the Guidelines to ensure no other equally job-related test would have less adverse impact). However, most courts instead follow *Albemarle*, which imposes the burden of locating such tests only on the plaintiffs, in the third part of the burden-shifting model of proof. 422 U.S. at 425.

D. Validity: Both Guidelines-based and non-Guidelines-based approaches to validity recognize three approaches that can be used to establish validity. Defendants need meet only one standard, and no approach is preferred over the others (although which test is most appropriate may depend on the circumstances). *See Ass'n of Mexican-Am. Educators v. California*, 231 F.3d 572, 584-86 (9th Cir. 2000); *Williams v. Ford Motor Co.*, 187 F.3d 553, 544-45 (6th Cir. 1999); 29 C.F.R. § 1607.5(A).

1. Criterion-Related Validity

a. Validity is determined by comparing performance on the test to performance in the job position. *United States v. City of Erie*, 411 F. Supp. 2d 524, 535–36 (W.D. Pa. 2005). "[E]mpirical data showing that the test is predictive of, or significantly correlated with, important elements of job performance" (i.e., statistics) is necessary. *Id*. An employer must show that the test scores related in a meaningful way with some measure of job performance. *Id*.

b. Can use either predictive or concurrent validation. *See, e.g.*, *Ernst v. City of Chicago*, 837 F.3d 788, 796-99, 800 (7th Cir. 2016) (describing these two models).

(1) Predictive: a sample group is tested before they begin work and selected without regard to scores, then their subsequent job performance is evaluated and correlated with their test scores to determine the accuracy of the test.

(2) Concurrent: same as predictive, except incumbent employees are tested while they are already working on the job.

(3) Predictive is preferred, but both are acceptable. *See United States v. City of Chicago*, 549 F.2d 415, 433 n.21 (7th Cir. 1977).

c. This validity analysis is essentially necessary for any test that attempts to predict future job performance, but it is not necessary for other types of tests (e.g., minimum competency tests). *See, e.g.*, *Ass'n of Mexican-Am. Educators*, 231 F.3d at 585; *Fickling v. N.Y. State Dep't of Civ. Serv.*, 909 F. Supp. 185, 191 (S.D.N.Y. 1995).

d. Three Major Legal Issues

(1) Selection of Criteria

(a) "Criteria" means key aspects of job performance

(i) In the ADA context, Plaintiffs should focus on selection procedures that measure "essential function" of the job. *See* 42 U.S.C. § 12111(8); *Jeffrey v. Ashcroft*, 285 F. Supp. 2d 583, 590-92 (M.D. Pa. 2003)

(ii) Courts may scrutinize which criteria are selected for testing, if not all areas of job performance are tested. *See, e.g.*, *Boston Chapter, NAACP, Inc. v. Beecher*, 371 F. Supp. 507, 517-18 (D. Mass. 1974), *aff'd*, 504 F.2d 1017 (1st Cir. 1974) (testing only a few of the duties required for successful job performance was not "convincing" evidence of job-relatedness).

(b) Job analyses are important to justify or attack selection of criteria. *See, e.g.*, *Albemarle*, 422 U.S. at 431-33, *Vulcan Soc'y v. Civ. Serv. Comm'n*, 490 F.2d 387, 394 & n.8 (2d Cir. 1973); *Fickling v. N.Y. State Dep't of Civ. Serv.*, 909 F. Supp. 185, 190-91 (S.D.N.Y. 1995); 29 C.F.R. § 1607.14(C)(2).

(i) Although defendants are not required to provide formal validation studies or job analyses, if they do, these studies or analyses must comply with technical standards established in federal regulations. *See Ernst v. City of Chicago*, 837 F.3d 788, 796 (7th Cir. 2016).

(ii) Experts are necessary here.

(2) Method of Measuring Criteria

(a) Supervisory ratings and the criteria on which they are based cannot be too vague or subjective. *See, e.g.*, *Robinson v. Union Carbide Corp.*, 538 F.2d 652, 662 (5th Cir. 1976) (finding criteria such as adaptability, demeanor, maturity, and social behavior to be too subjective and open to prejudice), *modified on other grounds*, 544 F.2d 1258 (5th Cir. 1977).

(b) Work samples, if used, must be validated before the test on which they rely can be validated. *See Ernst*, 837 F.3d at 802-04 (applying validation tests to

work samples and concluding that they were not validated).

(3) Degree of Correlation Necessary to Establish Validity

(a) The Uniform Guidelines do not provide a minimum correlation coefficient for all employment situations, 29 C.F.R. § 1607.14(B)(6), but the greater the adverse impact, the higher the correlation coefficient (or validity) needs to be.

(b) IO psychologists have suggested that a correlation "of approximately .20 is often high enough to be useful," .30 is good, and "a correlation of .40 is ordinarily considered very good." Amicus Curiae Brief of the Executive Committee of the Division of Industrial and Organizational Psychology, Am. Psych. Ass'n, at A-3, *United States. v. Georgia Power Co.*, 474 F.2d 906 (5th Cir. 1973).

(c) Some courts require more than a .30. *Compare Melendez v. Ill. Bell Tel. Co.*, 79 F.3d 661, 669 (7th Cir. 1996) (requiring a higher correlation) *with Williams v. Ford Motor Co.*, 187 F.3d 533, 546 (6th Cir. 1999) (finding a .30 correlation coefficient to be "statistically significant and sufficient to establish job-relatedness").

(d) A greater degree of adverse impact may lead a court to require a higher correlation before finding validity. *Guardians v. Civil Service Comm'n of New York*, 630 F.2d 79, 105-06 (2d Cir. 1980).

(e) Reliability of the test over time is also an issue. *See Ernst*, 837 F.3d at 801-02 ("All tests must be statistically examined for evidence of reliability before the test developer can establish the validity of the test." (internal citations and quotation marks omitted)); 29 C.F.R. § 1607.14(B)(5).

2. Content Validity

a. Defendant must demonstrate that the content of the test—often a "work-sample," simulation, or knowledge test—represents important aspects of job performance through "an evaluation of the extent to which the content of the test is adequately matched to the 'content of the job.'" *United States v. City of Erie*, 411 F. Supp. 2d 524, 535–36 (W.D. Pa. 2005).

b. Five factors for good content validation: suitable job analysis, competence in test construction, test construction related to job content, test content representative of job content, and a scoring system that selects those who can better perform the job. See *Guardians Ass'n of the N.Y.C. Police Dep't, Inc. v. Civ. Serv. Comm'n*, 630 F.2d 79, 95-106 (2d Cir. 1980) (providing a helpful discussion of content validity studies). The job analysis is key to proving content validity. *City of Erie*, 411 F. Supp. 2d at 535-36.

c. These tests do not need to measure every job skill or even the most common skills (e.g., a landing-simulation test for pilots can be content-valid even though landing constitutes only a small part of their job duties).

3. Construct Validity

a. This approach is used the least in current case law, but we may see it become more prevalent as tests assessing elements of an applicant's personality become more popular.

b. Most theoretical – need to "establish that a construct is required for job success and that the selection device measures that same construct." *United States v. City of Erie*, 411 F. Supp. 2d 524, 536 (W.D. Pa. 2005). The validation data must show that the test (a) "measures the degree to which candidates have identifiable characteristics (i.e., 'constructs')" and (b) "validly relates the constructs to the performance of . . . important work behavior(s)." *Id*. Part (b) is where most disputes occur.

c. Often need to use a criterion-related approach to show that the construct is related to job performance. *City of Erie*, 411 F. Supp. 2d at 535-36; 29 C.F.R. § 1607.14(D)(4).

d. This test is often used where another construct-based test has been validated, and Defendants want to argue that their new test is also valid because it measures the same construct. *City of Erie*, 411 F. Supp. 2d at 535–36.

4. Differential or Single-Group Validity: these approaches determine whether a test is valid for one group but not for another, i.e., that lower test scores for minority candidates do not correlate with worse job performance in the same way as is found for white employees. The Guidelines recommend completing these studies (deemed "unfairness" studies) where feasible. 29 C.F.R. § 1607.14(B)(8). Whether the phenomenon of differential or single-group validity actually exists is deeply questioned in psychological research, but some courts have required these approaches. *Clady v. Los Angeles*, 770 F.2d 1421, 1431

(9th Cir. 1985).

5. Situational Validity vs. General Validity: While some testing companies promote "general validity"—i.e., that a test is valid regardless of the job for which it is used—the Civil Rights Act of 1991 clarified that the showing of job relatedness must be specific to the "position in question." At a minimum, if an employer intends to adopt a test based on its validation with another employer, there must be ample evidence to establish a high degree of similarity in the job requirements and circumstances in order to be able to "port" validity over. On the other hand, if an employer has validated a test for use with a particular job at one location, it need not re-validate the test to use it at another geographic location.