ON THE LEGAL FRONT

The Meacham and Gulino Rulings:
Remnants of the Wards Cove Era

Art Gutman
Florida Institute of Technology

Eric Dunleavy
DCI Consulting

In August 2006 the 2nd Circuit ruled in two cases that have implications for adverse impact in the Age Discrimination in Employment Act (ADEA) (Meacham v. Knolls Atomic Power Laboratory [KAPL]; on August 14) and adverse impact in Title VII (Gulino v. New York State Education Department; on August 17). The Supreme Court invited the solicitor general to submit briefs expressing the views of the Bush administration in both bases. As this column was being written, the Supreme Court has decided to review Meacham.

The Meacham case is of particular interest to the TIP audience because it centers on the meaning of the Reasonable Factors Other than Age (RFOA) statutory defense recently supported by the Supreme Court in Smith v. City of Jackson (2005) in ADEA adverse impact cases. Specifically, this case will force the Court to consider whether the defendant has a burden of production or persuasion when the RFOA defense is invoked. The Gulino case is of particular interest to the TIP audience because the adequacy of content validity evidence in high-stakes testing is at the core of the claim. Additionally, the question of who is liable when one organization develops a test and another uses it is also a central issue in Gulino. Further, the Supreme Court hasn’t ruled on an adverse impact case under Title VII in almost 2 decades. Both cases are important and speak to unanswered questions dating back to the Wards Cove era.

Background Information

The two key cases from the Wards Cove era are Watson v. Fort Worth Bank (1988) and Wards Cove v. Atonio (1989), both Title VII cases. Recall that, during this period of time, courts were wrestling primarily with (a) which selection procedures were covered under an adverse impact theory of discrimination, and (b) exactly what plaintiff and employer burdens were under various circumstances. The question in Watson was whether subjective decision making may be challenged via adverse impact rules. There were three types of subjective decisions at issue (ratings of past performance, inter-
view ratings, and ratings of past experience). There were only eight justices in this case. All eight agreed that subjective decision making is subject to adverse impact rules. Indeed, speaking for all eight justices, Justice O’Connor, referencing *Griggs v. Duke Power* (1971), stated the following:

[I]f the employer in *Griggs* had consistently preferred applicants who had a high school diploma and who passed the company’s general aptitude test, its selection system could nonetheless have been considered “subjective” if it also included brief interviews with the candidates. So long as an employer refrained from making standardized criteria absolutely determinative, it would remain free to give such tests almost as much weight as it chose without risking a disparate impact challenge. If we announced a rule that allowed employers so easily to insulate themselves from liability under *Griggs*, disparate impact analysis might effectively be abolished.

However, at the same time, a plurality of four justices (O’Connor, Rehnquist, Scalia & White) argued that the rules for defending against adverse impact should be relaxed to the same standard as used in disparate treatment cases such as *McDonnell-Douglas v. Green* (1973). That is, instead of the job-relatedness defense for adverse impact, a burden of persuasion, the O’Connor plurality wanted to reduce the defense to a lighter burden of production, or an articulation (without proof) of a nondiscriminatory reason for the challenged practice. Or as stated by Justice O’Connor:

[W]hen a plaintiff has made out a prima facie case of disparate impact, and when the defendant has met its burden of producing evidence that its employment practices are based on legitimate business reasons, the plaintiff must “show that other tests or selection devices, without a similarly undesirable racial effect, would also serve the employer’s legitimate interest in efficient and trustworthy workmanship.”

O’Connor’s fear was that

In the context of subjective or discretionary employment decisions, the employer will often find it easier than in the case of standardized tests to produce evidence of a “manifest relationship to the employment in question.”

In response to the O’Connor plurality, Justice Blackman, speaking for Brennan and Marshall, issued two objections. First, Blackmun agreed with a brief written on behalf American Psychological Association (APA) in which Donald Bersoff (1988), citing both the 1985 *Standards for Educational and Psychological Tests* and the 1987 *SIOP Principles for the Validation and Use of Personnel Selection Procedures*, argued that subjective procedures are equally as amenable to psychometric scrutiny as are objective procedures. Blackmun also agreed that adverse impact was not a homogeneous scenario and that different types of proofs were used for different types of causes of
adverse impact. For example, he argued that the defense to adverse impact has varied with:

[T]he type and size of the business in question, as well as the particular jobs for which the selection process is employed. Courts have recognized ..., nationwide studies and reports... expert testimony... and psychologist’s testimony explaining job-relatedness...[etc.]

For his part, Justice Stevens avoided the argument, concluding that the only question raised in this case (subjective decision making) had been answered.

Fast forward one year to *Wards Cove*. That case had very little in common with *Watson*. The issue in *Wards Cove* was cross-job disparities between minorities (Eskimos & Filipinos) overrepresented in unskilled jobs and Whites overrepresented in skilled jobs. In fact, Gutman has argued in several places that *Wards Cove* should have been a pattern or practice disparate treatment case in the image of *International Teamsters v. United States* (1977)1. Under pattern or practice rules, the appropriate defense to “stock” statistics (as opposed to “flow” statistics as in *Griggs*) is the same as in individual disparate treatment cases such as *McDonnell-Douglas v. Green* (i.e., the lighter burden of production). Therefore, the burden of production would have been appropriate for *Wards Cove* had the Supreme Court evaluated it as a pattern or practice case but not as an adverse impact case.

The rest is, as we say, history. With the arrival of Justice Kennedy, there were now five votes to turn the O’Connor plurality opinion in *Watson* into case law in *Wards Cove*. Congress then attempted to overturn *Wards Cove* (and five other 1989 Supreme Court rulings) in the Civil Rights Restoration Act of 1990 (CRRA-90). CRRA-90 was vetoed by President Bush and was nearly overridden (the veto was overridden in the House of Representatives but missed by a single vote in the Senate). The primary source of disagreement in 1990 was the *Wards Cove* ruling. Disagreements between the Democrats and Republicans on *Wards Cove* (and the other five cases) were ironed out in the next year, and the result was the Civil Rights Act of 1991 (CRA-91). As it relates to *Wards Cove*, CRA-91 demands two of the three things initially proposed by the O’Connor plurality in *Watson*: (a) identification of an employment practice(s) that (b) causes adverse impact. However, it overturned the *Wards Cove* ruling on the burden of production, deeming that if adverse impact is proven in accordance with the identification and causation principles, the defense must prove that the cause(s) of adverse impact is job related and consistent with business necessity, leaving the plaintiff the burden of proving there are other equally valid practices that produce less or no adverse impact. In other words, the plaintiff and employer burdens were revised to be “balanced.”

---

1 See for example On The Legal Front articles written in the January 2003 and January 2004 issues of TIP and Gutman (2005).
More recently, the Supreme Court evaluated adverse impact in the ADEA in *Smith v. City of Jackson* (2005), a case discussed in the July 2005 issue of *On The Legal Front*. Briefly, in the 1980s, adverse impact followed the same rules in ADEA as in Title VII (see for example *Geller v. Markham*, 1980 & *Leftwich v. Harris-Stowe*, 1983). However, in *Hazen v. Biggens* (1993), a disparate treatment case, the Supreme Court ruled that employer decisions may be motivated by ‘factors other than age…even if the motivating factor is correlated with age.’” Several circuit courts read *Hazen* to mean that adverse impact is an invalid claim in the ADEA as a matter of law, including the 5th Circuit in *Smith v. City of Jackson* (2005). To the surprise of many observers, the Supreme Court ruled that adverse impact is a valid ADEA claim. However, there were two caveats. First, the Supreme Court ruled that *Wards Cove* applies to ADEA claims because CRA-91 only overturned that ruling with respect to Title VII. Second the ADEA has the Reasonable Factors Other than Age (RFOA) statutory defense, which is a substantially lighter burden of persuasion than the job relatedness and business necessity defenses under Title VII. Therefore, as discussed in the April 2005 issue of *TIP*, the following summary of the differences between Title VII and the ADEA was provided:

<table>
<thead>
<tr>
<th></th>
<th><em>Griggs-Albemarle</em> (Title VII)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 1</td>
<td>Statistical evidence of an identified employment practice that disproportionately excludes protected group members</td>
</tr>
<tr>
<td>Phase 2</td>
<td>Proof that the challenged practice is job related and consistent with business necessity</td>
</tr>
<tr>
<td>Phase 3</td>
<td>Proof there is an equally valid, job-related practice with less or no adverse impact</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th><em>Smith v. City of Jackson</em> (ADEA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 1</td>
<td>Statistical evidence of an identified employment practice that disproportionately excludes protected group members</td>
</tr>
<tr>
<td>Phase 2</td>
<td>Proof that the challenged practice is supported by a Reasonable Factor Other Than Age (RFOA)</td>
</tr>
<tr>
<td>Phase 3</td>
<td>Proof that the factor cited is unreasonable or not the true reason for the employment practice</td>
</tr>
</tbody>
</table>

As depicted in Table 1, adverse impact follows the same prima facie (Phase 1) rules in both Title VII and the ADEA. However, unlike Title VII, which demands proof of job relatedness and consistency with business neces-
sity (Phase 2) forcing the plaintiff to prove there is an equally valid practice with less or no adverse impact (Phase 3), the ADEA permits the RFOA defense (Phase 2) forcing the plaintiff to prove that the reasonable factors advanced are not reasonable (Phase 3). The Meacham case represents an application of the Smith burden to different facts of an involuntary reduction in force (IRIF) case.

**The Meacham Case**

*Meacham* was tried by the Northern District of New York in 2002. The plaintiffs won at trial, and the 2nd Circuit upheld the district court ruling in 2004 (*Meacham I*). However, the Supreme Court vacated the ruling in light of *Smith v. City of Jackson* and the 5th Circuit, in a 2 to 1 ruling, overturned the district court ruling in its more recent (2006) review (*Meacham II*). There are three interesting aspects of this case.

First, the 2nd Circuit traditionally used *Wards Cove* principles to decide adverse impact in ADEA cases even after CRA-91. Meacham involved an IRIF combined with other procedures, most notably a voluntary separation plan (VSP) for individuals with 20 years or more of service and who lacked critical skills. In the IRIF component, 98% of the laid off employees were over age 40. KAPL articulated that the employees laid off were among the lowest rated on the key variables of criticality of skills and flexibility for retraining. Consequently, the plaintiffs proved adverse impact (Phase 1), and the defendants carried their burden of production (Phase 2). However, in Phase 3, the plaintiffs proved to the satisfaction of a jury that there were alternatives that were as suitable, but with less adverse impact, including a hiring freeze and an extension of the VSP to employees with less than 20 years of service. Thus, the plaintiffs won on a strict interpretation based on *Wards Cove* rules at the district court level and in *Meacham I*.

The second interesting aspect of this case relates to the Supreme Court’s ruling in *Smith* that *Wards Cove* rules apply to ADEA whereas, in reality, they do not. As noted above, under *Wards Cove* rules, after the defendant carries its burden of production the plaintiff may still prevail by proving there are alternative suitable practices with less or no adverse impact. This is not an option when the RFOA defense is used. For example, in *Smith*, Justice Stevens ruled:

While there may have been other reasonable ways for the City to achieve its goals, the one selected was not unreasonable. Unlike the business necessity test, which asks whether there are other ways for the employer to achieve its goals that do not result in a disparate impact on a protected class, the reasonableness inquiry includes no such requirement.

---

2 The reader is directed to opposing views on alternatives to adverse impact written by Jim Sharf and Jim Ouutz in the October 2007 issue of *TIP*. 

The Industrial-Organizational Psychologist 47
In other words, if the defendant invokes RFOA, the plaintiff can prevail only if it can be proven that the reasonable factors offered are unreasonable (see for example EEOC v. Allstate [2006], where the defense articulated five reasonable factors and each was successfully countered as unreasonable by the plaintiffs). In Meacham II, the 2nd Circuit ruled that KAPL’s use of criticality and flexibility for the IRIF was reasonable and that the plaintiffs could not prove otherwise.

The third interesting aspect of this case relates to the RFOA defense itself. Generally, statutory defenses in all statutes are affirmative. In other words, the defendant must carry a heavier burden of persuasion, not production. For example, Part 1625.7(e) of the EEOC ADEA regulations states:

When the exception of “a reasonable factor other than age” is raised against an individual claim of discriminatory treatment, the employer bears the burden of showing that the “reasonable factor other than age” exists factually.

However, in Meacham II, the 2nd Circuit ruled that the burden in the RFOA is only productive and not a heavier persuasive burden. Accordingly:

Wards Cove explained that the plaintiff bears the burden of persuasion to defeat the employer’s “business necessity” justification because the plaintiff bears the ultimate burden under Title VII to “prove that it was ‘because of [his] race, color,’ etc., that he was denied a desired employment opportunity.” The analogous § 4(a) of the ADEA…is identical to that of Title VII…City of Jackson thus applies the reasoning and analysis of Wards Cove to disparate-impact claims under the ADEA, with the effect that an employer defeats a plaintiff’s prima facie case by producing a legitimate business justification, unless the plaintiff is able to discharge the ultimate burden of persuading the factfinder that the employer’s justification is unreasonable. Any other interpretation would compromise the holding in Wards Cove that the employer is not to bear the ultimate burden of persuasion with respect to the “legitimacy” of its business justification.

The same conclusion was drawn by the Eastern District of Missouri in EEOC v. Allstate (2006) and by the 10th Circuit in Pippin v. Burlington Resources (2006).

In the July 2005 issue of TIP, it was suggested that “there is no obvious reason to resurrect Wards Cove if the sole objective is to support the RFOA defense. RFOA stands alone as a Congressionally mandated statutory defense.” Meacham II illustrates this point. For example, as noted by the dissenting judge in this case (Pooler):

I respectfully dissent because I do not agree that Smith v. City of Jackson, 544 U.S. 228, 125 S. Ct. 1536, 161 L. Ed. 2d 410 (2005) requires vacatur of the district court judgment. The concerns animating my disagreement
with the majority are (1) the majority improperly conflates the analysis of proof of a reasonable factor other than age (“RFOA”) with the legitimate business justification analysis as it is used in a disparate impact analysis; (2) the majority errs by assigning to plaintiffs the burden of proving that a RFOA does not exist; and (3) the majority improperly reaches the asserted RFOA error because, although defendants pleaded an affirmative RFOA defense, they did not seek a charge or a verdict sheet question on that defense, thus requiring that we find fundamental error, which does not exist, to reach the claimed error.

In other words, Judge Pooler argued that by incorporating the RFOA statutory defense in Smith, the Supreme Court did not imply that the burden of defense for RFOA is productive. In our opinion, had the Supreme Court simply invoked the RFOA defense to adverse impact without reference to Wards Cove, lower courts would support the EEOC’s interpretation of the RFOA defense that it must be proven “factually.” That is to say, the heavier burden was expected to be the status quo.

The Gulino Case

Gulino involved two potential “employers,” the New York State Education Department (SED) and the New York City Board of Education (BOE), and two tests, the Core Battery and the LAST (Liberal Arts and Sciences Test). Ultimately, the BOE was deemed the employer. More importantly, both the Core Battery and the LAST produced adverse impact, but the Core Battery was properly validated using appropriate content validation procedures, whereas the LAST was not. Despite these facts, the district court judge found that both tests were job related and consistent with business necessity and, in the case of the LAST, that the plaintiffs failed to offer a “cost effective, practical alternative to the certification tests.” (Gulino v. Board of Education, 2003). The 2nd Circuit ruled that the BOE was the appropriate employer and overturned the district court’s ruling that the LAST was job related and consistent with business necessity.

The LAST covers five knowledge areas. Four of these areas are tested with multiple choice questions (science and math, historical and social science awareness, artistic expression and the humanities, and basic communication skills), and the fifth area is an essay test of written analysis and expression. Passing requires answering correctly two thirds of the multiple choice question and scoring at least 3 out of 5 on the essay test. The test is compensatory so that poor performance on one component can be overcome with good performance on another component. The passing rates were roughly 90–95% for Whites as compared to 50–60% for Blacks and 45–55% for Hispanics.

The most startling aspect of the district court ruling was that the judge gave reason after reason why the validation process by NES (National Eval-
uation Systems), the test maker, did not meet professional standards. She stated that “defendants have not demonstrated that the LAST was properly validated according to professional guidelines and standards.” She pointed to failure to satisfy criteria established for content validity in the Uniform Guidelines. Interestingly, she also cited several instances in which the APA Standards were violated, including Standard 3.7 (documentation procedures used to develop, review, and try out items), Standard 3.8 (samples representative of intended population), and Standard 7.3 (failure to conduct differential item analysis). She stated that “efforts to obtain sufficient technical information for the committee to evaluate the tests similar to what the committee received from ETS were unsuccessful for NES tests. As a result,…the committee can make no statements about their soundness or technical quality.” And she also stated that there is “insufficient documentation upon which to make a determination regarding the validity of the LAST for the uses to which it was put by defendants.”

In other words, the judge documented why the LAST should have failed the test for job relatedness and consistency with business necessity. Nevertheless, she interpreted the Supreme Court’s ruling in Watson as having lowered the bar for test validation, and thus, a lack of validity evidence did not equate to inadequate validity evidence, even in the case of a standardized test. Accordingly:

Unhappily for plaintiffs, however, the Supreme Court lowered the bar for defendants in disparate impact suits. In Watson v. Fort Worth Bank & Trust, 487 U.S. 977, 108 S. Ct. 2777, 101 L. Ed. 2d 827 (1988), the Court explained: Our cases make it clear that employers are not required, even when defending standardized or objective tests, to introduce formal “validation studies” showing that particular criteria predict actual on-the-job performance. In [New York City Transit Authority v. Beazer, 440 U.S. 568, 99 S. Ct. 1355, 59 L. Ed. 2d 587 (1979)], for example, the Court considered it obvious that legitimate employment goals of safety and efficiency permitted the exclusion of methadone users from employment with the New York City Transit Authority; the Court indicated that the manifest relationship test was satisfied even with respect to non-safety-sensitive jobs because those legitimate goals were significantly served by the exclusionary rule at issue in that case even though the rule was not required by those goals.

The judge then concluded that the “LAST is manifestly related to the legitimate educational goals enunciated by SED.”

For its part, the 2nd Circuit gave numerous reasons why the LAST was not job related. For present purposes, two are worth nothing. First, the court noted that there was nothing in the Watson ruling that lowered the bar for test validation. Rather, the passage from Watson relied upon by the district court
judge was part of a plurality ruling, and the reference to New York City v. Beazer (1979) was a subset of “outlier” cases in which validation is not necessary because it was “obvious” that an exclusionary rule prohibiting methadone users from becoming transit authority police officers was a valid requirement. Thus, the question became whether an exclusionary rule was different from a standardized test. Or in the words of the court:

[I]t is not clear that the quoted portion of the Watson opinion purported to overrule earlier Supreme Court cases that require employers to conduct validation studies that are at least consistent with the EEOC Guidelines. We think that Watson as a whole is more reasonably read as simply pointing out that some tests measure abilities that are abstract, yet so clearly consistent with legitimate business needs, that formal validation may be either functionally impossible or inadequate as a[*386] measure of the test’s job relatedness. The examples from Beazer discussed in the quoted passage illustrates a subset of disparate impact cases in which the job relatedness of an employment practice is so patent that formal validation is unnecessary. …Second, courts should not rely on this portion of Watson[*71] because that language comes from a section of the Watson opinion that was joined by only four of the eight participating justices.

Ironically, if anything, the reference to Beazer supports the Blackmun plurality opinion in Watson that there was no need to alter the rules for adverse impact claims because different methods for determining job relatedness had already been established in prior Supreme Court and lower court cases.

Second, the court reiterated it’s five-part test for content validity initially established in Guardians v. Civil Service (1980). Accordingly:

(1) [T]he test-makers must have conducted a suitable job analysis[;] (2) they must have used reasonable competence in constructing the test itself[;] (3) the content of the test must be related to the content of the job …[;] (4) the content of the test must be representative of the content of the job[; and] [there must]be (5) a scoring system that usefully selects from among the applicants [*385] those who can better perform the job.

Obviously, based on the district court judge’s own references to the reasons why the LAST was not properly validated, the 2nd Circuit concluded that the LAST failed the Guardians test and is not job related and consistent with business necessity.

Conclusions

It is not surprising to us that the Supreme Court would choose to review Meacham II, particularly after the Smith ruling. Frankly, the Supreme Court has run this race before. Specifically, in Public Employees Retirement System of Ohio v. Betts (1989), the Supreme Court itself placed a burden of pro-
duction on the BFBP (bona fide benefit plan) statutory ADEA defense ruling that, if invoked by the defense, initiates the plaintiff’s burden to prove that the benefit plan is “subterfuge to evade the purposes of the Act.” Congress subsequently passed the Older Workers Benefit Protection Act of 1990 that, among other things, overturned the Betts ruling, thereby placing the burden of proof of BFBP on the defendant. We suspect the Supreme Court is likely to do the same thing in relation to Meacham II.

On the other hand, the reasons for reviewing Gulino (if in fact that happens), is somewhat opaque. There are no obvious disagreements among circuit courts on the proper standards for conducting a content validity study. Certainly, the five-part rule established by the 2nd Circuit in Guardians is consistent with both the Uniform Guidelines and generally accepted principles in our profession. There is, however, one possibility relating to the 2nd Circuit’s opinion in Gulino that is worrisome. Specifically, the court wrote:

[T]his Circuit remains bound by the validation requirements expressed in earlier Supreme Court precedent, namely, Albemarle Paper and Griggs, and as interpreted by Guardians. Further bolstering that conclusion is the fact that this case is much more factually analogous to Albemarle Paper and Griggs than to Watson. Albemarle Paper, Griggs, and Guardians addressed the use of standardized tests in making employment decisions. Watson, on the other hand, addressed the applicability of Title VII to subjective employment practices, such as evaluations by superiors, in making employment decisions. The testing-related cases delineate the appropriate standard for assessing job relatedness.

This quote could be interpreted to mean that cases such as Griggs v. Duke Power (1971) and Albemarle v. Moody (1975), which involved standardized tests, should be validated in one fashion, whereas cases such as Watson, which involved subjective ratings of past performance, interview ratings, and ratings of past experience, should be validated differently. This is just a guess on our part, and one we hope is wrong. The opinion of the 1988 Bersoff brief makes sense, and in accordance with the SIOP Principles, there is no reason to believe there are any differences in the manner in which subjective and objective forms of assessment are amenable to psychometric scrutiny.

The fact that the Court asked for the solicitor general to weigh in suggests that Gulino will be added to the docket eventually. Perhaps the less interesting issue in Gulino concerning who the “employer is” could be of interest to the Court in spite of the fact that this issue hasn’t stirred controversy for a long time. Alternatively, Gulino could be of interest to the Court because of the educational context surrounding the case, particularly given that educational reform and teacher certification have recently resurfaced as issues of national concern. According to the Center for Constitutional Rights (CCR), thousands of teachers have been demoted from their jobs as a result of the
National Teachers Exam (NTE). Poor performance on the NTE may also result in the loss of teaching licenses, seniority, retention rights, tenured teaching positions, and even salary reductions. Perhaps the tests under review and the jobs of interest in *Gulino* are the real focus of the Court and not necessarily anything controversial about adverse impact theory under Title VII.

**References**


**Cases Cited**

Geller v. Markham (CA2 1980) 635 F.2d 1027.
Guardians v. Civil Service Commission (CA2 1980) 630 F.2d 79.
Gulino v. State Education Department (CA2 2006) 461 F.3d 134.
Meacham v. Knolls Atomic Power Laboratory (CA2 2004) 381 F.3d 56.
Meacham v. Knolls Atomic Power Laboratory (CA2 2006) 461 F.3d 134.