Test Fairness: Legal and Measurement Issues in Regulatory Testing

Presenters: John Weiner, PSI Services LLC
Keith Pyburn, Fisher & Phillips LLP

Fairness is a Longstanding Concern
• A deep-rooted issue in education, employment & credentialing
• What makes a test fair or unfair?
  - Many definitions and perspectives:
    • Social - equality
    • Legal - 14th amendment; case law
    • Psychometric - reliable and valid

Today’s Session Will Cover
• Psychometric approaches and issues in fairness
• Key laws, cases & implications for credentialing
• Legal requirements for ensuring fairness
• Considerations for regulatory testing programs
• Discussion of fairness issues
Psychometric Perspective

- Definitions of Fairness
- Psychometric Approaches & Examples
- Issues for regulatory testing

Definitions of Fairness

- Standards for Educational & Psychological Testing, AERA et al, 1999, Ch. 7)
- 4 types of “fairness”
  1. No bias (prediction)
  2. Equal treatment (use)
  3. Equal outcomes (pass rates)
  4. Equal opportunity to learn (precursors)
- Only types 1 & 3 relate to test properties

Psychometric Approaches

1. Fairness = equal prediction
   “...no bias exists if the regression equations relating a test and criterion are indistinguishable for the groups in question.”
   (Standards, p.79)

   - The focus is on validity...
3. Fairness = equal outcomes
   - Type of thinking behind “Golden Rule” case
   - The focus of many states and federal EEO agencies
   - NOT embraced by the scientific community
     “The idea that fairness requires equality in overall passing rates for different groups has been almost entirely repudiated in the professional testing literature” (Standards, p.74)

**Example 1: Fairness as Equal Outcomes**

Comparable mean scores result in comparable passing rates.

<table>
<thead>
<tr>
<th>Test Score</th>
<th>Cut score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 2</td>
<td></td>
</tr>
</tbody>
</table>

**Example 2: Unfairness = Unequal Outcomes**

Lower mean results in lower pass rate for group 2

“Golden Rule” view of unfairness.

<table>
<thead>
<tr>
<th>Test Score</th>
<th>Cut score</th>
</tr>
</thead>
<tbody>
<tr>
<td>2, 1</td>
<td></td>
</tr>
</tbody>
</table>
Example 3: Fairness as Unbiased Test

Correlation \( (r) \) between test score and criterion is the same

Test score means may be different

Example 4: Unfairness = Biased Test

\( r \) between test score and criterion is not the same for the groups

May differ in slope, intercept, or SE

Fairness Issues for Regulatory Testing

- Concern that focus has been on equal outcomes (pass rates)
  - Differences may be due to non-test factors
  - Does the test cause group differences or reflect them?
- Challenge is to incorporate other psychometric approaches that are not common to credentialing
  - CRV, differential prediction
Legal Perspective

- Courts focus on whether the credentialing decision involves “intentional discrimination,” and, if it does whether the use of such a generally prohibited factor can be justified under the “strict scrutiny” test.

- Absent evidence of “intentional discrimination,” the requirement need only have a “rational relationship” to legitimate state interests.

Legal Perspective - What is Required?

“A state cannot exclude a person from the practice of law or from any other occupation in a manner or for reasons that contravene the Due Process or Equal Protection Clause of the Fourteenth Amendment.”

Schware v. Board of Bar Examiners, 353 U.S. 232, 238-9 (1957)

Legal Perspective - What is Required?

“A State can require high standards of qualification, such as food moral character or proficiency in its law, before it admits an applicant to the bar, but any qualification must have a rational connection with the applicant’s fitness or capacity . . .”

Schware v. Board of Bar Examiners, 353 U.S. 232, 239 (1957)
Evidence of Discrimination

- Overt discrimination in state action (including licensing) is unconstitutional:
  
  It is by now well established that "all racial classifications reviewable under the Equal Protection Clause must be strictly scrutinized." *Adarand Constructors, Inc.*... This "standard of review . . . is not dependent on the race of those burdened or benefited by a particular classification." *Ibid.* . . . Thus, "any person, of whatever race, has the right to demand that any governmental actor subject to the Constitution justify any racial classification subjecting that person to unequal treatment under the strictest of judicial scrutiny." *


Evidence of Discrimination

- Is adverse impact evidence sufficient to prove "discrimination" in violation of constitutional standards?
- In other words, is an adverse (non-proportional) pass rate for a given group sufficient evidence of unfairness to support a finding the test is unconstitutional?

Evidence of Discrimination

- The Plaintiffs sued about a test for entry into the Police academy in D.C. and alleged the test was unconstitutional based on their:
  
  "claim that Test 21 bore no relationship to job performance and "has a highly discriminatory impact in screening out black candidates."

Evidence of Discrimination

“As the Court of Appeals understood Title VII, employees or applicants proceeding under it need not concern themselves with the employer's possibly discriminatory purpose but instead may focus solely on the racially differential impact of the challenged hiring or promotion practices. This is not the constitutional rule. We have never held that the constitutional standard for adjudicating claims of invidious racial discrimination is identical to the standards applicable under Title VII, and we decline to do so today.

Id. At 238, 239

Rationally Related?

• What makes a test not “rationally related” to a legitimate state interest?
• No logical connection to the state purpose for giving the test?
• A test where some have prior access to the test and others do not? (Cheating)
• A test that totally excludes a given group?

License or Employment?

• In concluding that Title VII applies in this case, the district court held that Defendants “interfer[ed]” with Plaintiffs’ employment opportunities with local school districts in California by requiring, implementing, and administering the CBEST. See AMAE I, 836 F.Supp. at 1551. We agree.

• Our conclusion is dictated by the peculiar degree of control that the State of California exercises over local school districts. In California, public schools are “a matter of statewide rather than local or municipal concern; their establishment, regulation and operation are covered by the [state] constitution and the state Legislature is given comprehensive powers in relation thereto.” Hall v. City of Taft, 47 Cal.2d 177, 179, 302 P.2d 574, 576 (1956).

The Golden Rule Case

- ETS settles licensing testing case involving a challenge to a test that disproportionately excluded minorities.
- Because it was a settlement it is not legal precedent.
- Practice of examining test construction to make sure it addresses “fairness” certainly remains a good practice.

State Laws

- Trend for states to require monitoring of pass rates for racial/ethnic minorities
  - e.g., Insurance licensure exams
    - MD (SB-701), TX (HB-1757)
    - Other states: IL, FL, VA, AZ, LA, WA, MO, DE
- Other laws?
- Implications for regulatory testing programs

Discussion Questions

- Are legal risks increasing for credentialing exams? What lies ahead?
- Has the fairness of your exam been challenged?
- How are you addressing test fairness?
- Does DIF analysis ensure test fairness?
- How does test security affect test fairness?
Thank you!

John Weiner
jweiner@psionline.com
www.psionline.com

Keith Pyburn
kpyburn@laborlawyers.com