Avoiding Misuse of Tests: Strategies for Defensible Use of Exam Results for Making Hiring Decisions

Foster Pepper Civil Service Conference 2013

Carl Swander, Ph.D.
Ergometrics and Applied Personnel Research
Lynnwood, WA

Use of Tests

- **Research is Clear**
  - Testing for employment screening is powerful
  - Tests can be used defensibly
  - There are many different factors to consider in the design of testing process

- **Lessons Learned**
  - Tests must be valid
  - Valid tests can be used incorrectly
Overview

- Review of Test Principles
  - Definitions
  - Use of tests
  - Types of tests
  - Research
  - Strategies for implementation and use
- Examples of What Can Go Wrong
- Provide Ways to Enhance Defensibility

What Can Go Wrong

- Chicago Firefighter Lawsuit – 78 Million
- New York Firefighter Lawsuit – 128 Million
- Other Examples
  - Buffalo, Corpus Christi, New Jersey
  - Years of challenges, DOJ, EEOC and other issues
  - Many, many, many civil service challenges and appeals
- Most likely hiring personnel that are not the best
Definitions

- **Adverse Impact**
  - Typically defined as the 80% (4/5ths) rule
  - Minority group pass rate should be at least 80% of the majority group
  - Not illegal but a common element in most legal situations

- **Validity**
  - Tests must be valid
  - Criterion and content approaches

Validation of Tests

- **Content Validation**
  - Demonstrating a nexus between the test and important job demands
  - Job analysis (JA), test blueprints, clear linkages from test to JA

- **Criterion Validation**
  - Demonstrating that the test is statistically related to job performance
  - Keys: Power and Repeatability
Criterion Validity

1. New York
   - 128 Million in Damages
   - Adverse Impact
   - Minimal Validity Evidence
   - Use of Test Results is Discrimination
     - Violates the 1991 Civil Rights Act (Title VII)
   - Lesson Learned
     - Validation is critical
     - Reflect actual job demands
Types of Tests

- Personality
- Cognitive Ability
  - Job knowledge
- Job Simulation
  - Work sample

Personality Tests

- Abundance of Research
- Poor Validity Evidence; Response Distortion
- Should be used to identify problems rather than select for performance (e.g., MMPI)
Personality Tests

- Example Questions (IPIP, 2013)
  - Don't mind being the center of attention.
  - Often feel blue.
  - Am full of ideas.

- Potential problems
  - Lack of validity and repeatability in job performance
  - Should not be used to rank order or band candidates

Cognitive Ability Tests

- Research demonstrates high validity
  - (Schmidt & Hunter, 1998; Schmidt & Hunter 1981; Hunter & Hunter, 1984)

- High Adverse Impact
  - (Chan & Schmitt, 1997; Hunter & Hunter, 1984)

- Used in the majority of selection processes
Cognitive Ability Tests

- Examples
  - General cognitive ability
  - Reading
  - Writing
  - Math
  - Critical thinking

- Potential Problems
  - Adverse impact
  - Limited coverage of the job domain

Predictor Combination

- Personality plus cognitive ability
  - (Ryan, Polyhart & Friedel, 1998; Sackett & Ellingson, 1997; Sacket & Roth, 1996; Schmitt, Rogers, Chan, Sheppard, & Jennings, 1997)

- Not very effective – personality has to be given large weight to impact adverse impact
  - (Ryan et al., 1998; Schmitt et al., 1997)
Job Simulations

- Job simulations/situational judgment tests are widely studied and shown to be valid

Job Simulations

- Demonstrate less adverse impact than cognitive ability without sacrificing validity
  - (Robertson & Kandola, 1982)
- Comply with EEOC guidelines directing that alternative methods of selection be used when they demonstrate less adverse impact and have comparable validities
  - (EEOC, 1978; Strong & Najar, 1999)
Simulation Examples

- FrontLine
- FireTEAM
- ECOMM

Taking a Deeper Look

- What types of tests are you using?
- Personality – Can have trouble with validity
  - Make sure no adverse impact
  - Look out for weighted personality/cognitive predictors that have adverse impact (low validity)
- Cognitive Ability
  - Ensure validity evidence is clear
- Job simulations have clear advantages
Taking a Deeper Look

- Job simulations have clear advantages
  - High validity
  - Lower adverse impact
  - Can cover a wide range of KSAOs
  - Are more widely accepted by candidates

Test Use/Scoring

- Valid tests can be used inappropriately
  - Cut Scores
  - Rank Order
  - Banding
  - Final Decision Making
Cut Scores

- Why not use 70?
  - Arbitrary

- Science behind your cut-off score
  - Validation information
  - SME input (Angoff)
  - Applicant data
  - Job analysis data

- Clearly identify why those below will not be successful

Chicago

- Written tests – basic skills (cognitive)
- Adverse impact
- Content validation
- Passing score = 65
- Only process score $\geq 89$ – “well qualified”
- Inappropriate use of rank order/banding
Rank Order

- The process of strict hiring from top of the list to the bottom
- Banding is a Variant
  - Grouping of candidates that are rank ordered
  - Groupings are determined by the SEM of the Exam
- Information Needed to Rank Order
  - Strong criterion validation
  - Clear documentation that higher levels of a KSA are related to increase in job performance

Final Decision Making

- Validity of Tests
- How Tests are Combined
- Cut Off, Rank Order, Band, Both
- Clearly Defined Strategy
  - Selection process
  - Final decision making
Taking a Deeper Look

- Make sure there are not alternate predictors with equal validity and less adverse impact
- What type of validation has been conducted?
- Is the test job related?
- How can you use the test?

Conclusions

- Make sure your tests are valid
- Make sure there are not alternate predictors with equal validity and less adverse impact
- Tests must be validated for specific use
- Do this to be defensible but also to hire the best