# Thinking Outside the Box in Merit Selection

Joel P. Wiesen, Ph.D.

wiesen@appliedpersonnelresearch.com

MAPAC Fall Conference

Princeton, NJ

11/16/06

# Two Most Vexing Problems

- Adverse Impact
- Ceiling on Validity

# Adverse Impact

- Seen with cognitive ability tests
- DOJ sometimes favors random selection
  - Perhaps with a low cut point
- Search for alternative selection procedures has led to innovations in personnel selection

# Ceiling on Validity

- Rarely observe validity over .50

# New Ideas

- New ways to use test scores
  - Reduce adverse impact
  - Maintain validity
- New selection tools/approaches

- Based on 2006 IPMAAC presentation

# New Ways To Use Test Scores

- Propose: Occupational Diversity Models
  - A family of models
  - Depart from weighted component model
- Greatest Strength Model (SGM)
- Many variations of ODM
  - Greatest Two Strengths Model
  - Drop Lowest Score Model

# Why Look for New Models?

- Pesky facts about adverse impact
  - Expectation of a sum = Sum of the expectations
    - Mean of sum will equal the sum of the means
    - Mean difference between groups will be no less than mean difference for the component with the greatest adverse impact
    - Adverse impact is additive (in terms of means)

# Occupational Diversity Models

- Proposed new ways to use test data
- Employees contribute based on strengths
- Cookie-cutter model may be wrong for some jobs
- Team orientation

# Team Orientation

- Team members not all equal
- Some can rebound (e.g., Marcus Camby)
- Some can score, but have some weaknesses
- Build on strengths
- Compensate for weaknesses

# Greatest Strength Model

- Step 1. Give several tests
- Step 2. Put tests on common metric
- Step 3. Determine highest score
- Step 4. Fail candidates with any low score
- Step 5. Rank candidates based on their highest scores

# Firefighter Example

- Written (M/C) test of cognitive ability
- Structured oral interview
- Physical performance test (PPT)

# Greatest Strength Model: Firefighter Example

| Person | Written | Oral | PPT | Highest Grade |
|---|---|---|---|---|
| A | 80 | 90 | 95 | 95 |
| B | 90 | 70 | 80 | 90 |
| C | 75 | 65 | 80 | 80 |

# Evaluating the GSM

- Adverse Impact
- Validity

# Approaches to Evaluation

- Data from the real world
  - Now seeking such data
- Simulation study
  - Easier to do
  - Faster and more comprehensive evaluation

# Simulation Study

- Create imaginary applicants
- Create test/job data with known correlations
- Evaluate data two ways:
  - GSM
  - Conventional approaches

# Simulation Study Methodology

- Specify intercorrelations
- Generate data with these intercorrelations
- Create gender and EEO groups
- Create mean score differences
- Evaluate adverse impact in appointments
- Evaluate validity

# Specify Intercorrelations

| | Oral | PPT | Gender | EEO Gp. | Job Perf. |
|---|---|---|---|---|---|
| M/C Cog | .2 | 0 | 0 | 0 | .35 |
| Oral | | 0 | 0 | 0 | .35 |
| PPT | | | 0 | 0 | .35 |
| Gender | | | | 0 | 0 |
| EEO Gp. | | | | | 0 |

# Create Mean Score Differences

- Gender: 1.25 s.d. on PPT
- EEO Group: 1 s.d. on written cognitive test

# Adverse Impact in Appointments

- Will vary by selection ratio
- Lower selection ratios yield higher impact
- Assume we hire top 3% of applicants
  - Extreme example
  - Realistic for Massachusetts firefighters

# Adverse Impact in Appointments

| M/C Cog Only | | | | | |
|---|---|---|---|---|---|
| EEO | Gender | | | | |
| .08 | 1.04 | | | | |

Note: Based on 50,000 cases.

Key: EEO stands for EEO group.

GSM stands for Greatest Strength Model.

# Adverse Impact in Appointments

| M/C Cog Only | | Composite | | | |
|---|---|---|---|---|---|
| EEO | Gender | EEO | Gender | | |
| .08 | 1.04 | .30 | .21 | | |

Note: Based on 50,000 cases.

Key: EEO stands for EEO group.

GSM stands for Greatest Strength Model.

# Adverse Impact in Appointments

| M/C Cog Only | | Composite | | GSM | |
|---|---|---|---|---|---|
| EEO | Gender | EEO | Gender | EEO | Gender |
| .08 | 1.04 | .30 | .21 | .73 | .75 |

Note: Based on 50,000 cases.

Key: EEO stands for EEO group.

GSM stands for Greatest Strength Model.

# Adverse Impact in Appointments

| Sample of 1,000 | M/C Cognitive Only | | Composite | | GSM | |
|---|---|---|---|---|---|---|
| | EEO | Gender | EEO | Gender | EEO | Gender |
| 1 | .03 | 1.00 | .15 | .30 | .71 | 1.67 |
| 2 | .07 | .88 | .43 | .20 | .62 | .50 |
| 3 | .07 | .76 | .11 | .25 | .28 | .77 |
| 4 | .03 | 1.14 | .25 | .20 | .33 | .60 |
| 5 | .11 | 1.50 | .25 | .15 | .53 | .37 |
| 6 | .00 | .88 | .50 | .30 | .56 | .79 |
| 7 | .03 | 1.31 | .15 | .43 | .71 | 1.00 |
| 8 | .03 | 1.00 | .50 | .25 | 1.00 | .83 |
| 9 | .15 | .76 | .30 | .15 | .56 | .79 |
| 10 | .15 | .76 | .30 | .15 | .56 | .79 |
| Average | .07 | 1.00 | .29 | .24 | .59 | .81 |

# Evaluate Adverse Impact

- Much lower adverse impact with GSM

# Evaluate Validity

- We have job performance data!
- We have GSM grade
- We can calculate a composite score based on M/C cognitive, oral, and PPT
- Can compute criterion-related validity

# Evaluate Validity

| M/C Cog Only | | |
|---|---|---|
| .31 | | |

Note: Based on 50,000 cases.

Key: EEO stands for EEO group.

# Evaluate Validity

| M/C Cog Only | | GSM |
|---|---|---|
| .31 | | .40 |

Note: Based on 50,000 cases.

Key: EEO stands for EEO group.

GSM stands for Greatest Strength Model.

# Evaluate Validity

| M/C Cog Only | Composite | GSM |
|---|---|---|
| .31 | .52 | .40 |

Note: Based on 50,000 cases.

Key: EEO stands for EEO group.

GSM stands for Greatest Strength Model.

# Evaluate Validity

| Sample of 1,000 | M/C Cog. Only | Composite | GSM |
|---|---|---|---|
| 1 | .29 | .49 | .41 |
| 2 | .33 | .52 | .42 |
| 3 | .30 | .55 | .43 |
| 4 | .25 | .46 | .36 |
| 5 | .34 | .53 | .43 |
| 6 | .32 | .55 | .40 |
| Average | .30 | .51 | .39 |

# Evaluate Validity

- Higher validity for GSM than M/C cognitive
- M/C cognitive was the standard for generations

# Two Most Vexing Problems

- Adverse Impact
- Ceiling on Validity

# Ceiling on Validity

- Consider other models of job performance
- New ideas on tests and their uses
- “New” KSAPs may have unexpected relationships with criterion

# Models of Job Performance

- Compensatory model
- Wiesen's Occupational Diversity Models
  - Greatest Strength Model
  - Drop Lowest Score Model
  - Many other possible models
- Parse abilities more finely and look for non-linear solutions to regression equations

# New Ideas on Tests

- New ways to use test scores
  - Reduce adverse impact
  - Maintain validity
- New selection tools/approaches (Wiesen, 2004)

# New Selection Tools/Approaches

- Alternate ways to pass the first hurdle
- More use of life/work experience
- Other types of tests
- Consider stability of personality traits
  - 75% of variation in weekly job performance is within person rather than between person. (Stewart and Nandkeolyar, 2006)

# More Ways to Pass First Hurdle

- Written cognitive ability test
- High school rank
- Score on statewide HS graduation test
- College degree
- Honorable discharge from military
- Allow retaking test
- Several week course

# Life and Work Experience

- Volunteer experience as Firefighter
- Paid experience as Firefighter
- Recommendations from teachers

# Other Types of Tests

- Several week course on fire subjects
- Face recognition tests (esp. for police)
- Short term memory test
- Peripheral vision test
- Spatial orientation (esp. for firefighter)
- Balance
- Oral comprehension of various dialects
- Fine motor coordination (e.g., paramedics)

# Other Types of Tests

- Mackworth Clock Test
  - Attentional capacity (e.g., Hollenbeck et al. 1995)
- Affect intensity (e.g., Larson, 1987)
- Education & experience evaluations
  - Citizenship behaviors
  - Altruistic behaviors
  - Ability to deal with interruptions

# Stability of Personality Traits

- Cognitive ability is stable
- Within-person variability of personality
  - Sociability may vary from day to day
  - Responsibility may wax and wane
  - e.g., Beal et al. (2005), Fleeson et al. (2002)
- Can our current models handle this within person variability?

# Summary

- New ways to combine test scores
  - Greatest Strength Model (GSM)
  - Wiesen Occupational Diversity Models
- Reduce adverse impact
- Maintain Validity
- New measurement tools and approaches

# Final Thoughts

- Field is still young and developing
- Call for collaboration in simulations
  - Students
  - Researchers
  - Practitioners
- Call for real life applications
  - Police Officer
  - Firefighter

Copies of this presentation are available at:
http://appliedpersonnelresearch.com/pubs.html

# References


- Beal, DJ, Weiss, HM, Barros, E & MacDermid, SM (2005) An Episodic Process Model of Affective Influences on Performance. *Journal of Applied Psychology*, *90*, 1054-1068.
- Fleeson, W, Malanos, AB & Achille, NM (2002) An Intraindividual Process Approach to the Relationship Between Extraversion and Positive Affect: Is Acting Extraverted as "Good" as Being Extraverted? *Journal of Personality and Social Psychology*, *83*, 1409-1422.

# References


- Hollenbeck, JR, Ilgen, DR, Tuttle, DB & Sego, DJ. (1995) Team Performance on Monitoring Tasks An Examination of Decision Errors in Contexts Requiring Sustained Attention. *Journal of Applied Psychology*, *80*, 685-696.
- Larsen, R.J. (1987) The Stability of Mood Variability A Spectral Analytic Approach to Daily Mood Assessments. *Journal of Personality and Social Psychology*, *52*, 1195-1204.

# References


- Stewart, G.L. & Nandkeolyar, A.K. (2006) Adaptation and intraindividual variation in sales outcomes: exploring the interactive effects of personality and environmental opportunity. *Personnel Psychology*, *59*, 307-332.
- Wiesen, J. (2004, October) *Adverse Impact: Theory and Practical Approaches.* Paper presented at the Fall meeting of MAPAC, New York, NY.