

NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

Effective Psychometric and Administrative Approaches to Maximize Diversity in Hiring Police Officers

Joel P. Wiesen, Ph.D.
 83rd Annual Convention of the
 New York State Psychological Association
 October 15, 2021


Wiesen (2021) - New York State Psychological Association Conference
1


NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

Presentation Topics

- Need for this presentation
- My twin goals
- Diversity is a legitimate goal
- Three psychometric approaches
- Three administrative approaches


Wiesen (2021) - New York State Psychological Association Conference
2


NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

Need for This Presentation

- PDs are often stymied in hiring black POs
 - More black POs not a panacea for social ills
 - More black POs is a step in the right direction
- Our exams often have adverse impact (AI)
 - Blacks tend to rank lower on exams, on average
 - Lower ranks tend not to be hired
- Re-design exams to hire more black POs?
 (PD = police department; PO = police officer)


Wiesen (2021) - New York State Psychological Association Conference
3


NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

My Twin Goals

- Increase **both**:
 - diversity in hiring
 - expected job performance


Wiesen (2021) - New York State Psychological Association Conference
4


NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

Diversity Is a Legitimate Goal

- Is striving for diversity in hiring a legitimate goal for a psychologist?
- Is choosing a test battery with maximum validity the only legitimate option?
- Do current test batteries for PO have reasonable levels of validity?

Wiesen (2021) - New York State Psychological Association Conference
5


NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

Diversity Is a Legitimate Goal

- Joint *Standards* (AERA, APA, NCME, 2014)
 “... The fact that the validity evidence supports the intended interpretation of test scores for use in applicant screening **does not mean that test use is thus required: Issues other than validity ... can play an important ... role in decisions** about test use.” (Page 21, emphasis added)

Wiesen (2021) - New York State Psychological Association Conference
6



Diversity Is a Legitimate Goal

- Joint *Standards* (AERA, APA, NCME, 2014)
“... different decision makers may make **different value judgments** about the impact of consequences on test use.” (Page 21, emphasis added)



Diversity Is a Legitimate Goal

- Joint *Standards* (AERA, APA, NCME, 2014)
“However, if ... **excluding some components** ... has a noticeable impact on selection rates for groups ... the intended interpretation of test scores ... would be rendered **invalid.**” (Page 21, emphasis added)



Diversity Is a Legitimate Goal

- Joint *Standards* (AERA, APA, NCME, 2014)
“... **evidence about consequences is relevant to validity** when it can be traced to a source of invalidity such as **construct underrepresentation** or construct-irrelevant components.” (Page 21, emphasis added)



Diversity Is a Legitimate Goal

- *Principles* (SIOP, 2018)
“... deficiency in measuring a ... construct ... or [omitting a construct] ... When judging [if omitting a construct] ... is problematic, professional judgment ... takes into account both psychometric and practical considerations, **including systematic bias** against subgroups...” (page 12-13, emphasis added)




Diversity Is a Legitimate Goal

- *Principles* (SIOP, 2018)
“When judging ... the ... [predictor, consider] ... whether factors such as age, race, ethnicity, ... **restrict accessibility [AKA opportunity to learn]** and affect measurement of the construct of interest.” (Page 12, emphasis added)



Diversity Is a Legitimate Goal

- *Principles* (SIOP, 2018)
“**Most organizations strive for a diverse** and inclusive workforce and equitable treatment of cultural and linguistic minorities.” (Page 23, emphasis added)


NYSPA CONVENTION
 OCTOBER 19 - 21, 2021

Diversity Is a Legitimate Goal

"... an optimal selection strategy may not be optimal for other employment functions, such as recruiting ... [consider] factors such as ... workforce diversity ... **When an organization focuses solely on selection ...the ... effectiveness of the overall HR system may suffer considerably.**"
 (Cascio & Aguinis, 2011, page 331, emphasis added)

Wiesen (2021) - New York State Psychological Association Conference 13



NYSPA CONVENTION
 OCTOBER 19 - 21, 2021

Diversity Is a Legitimate Goal

- "A critical goal for psychological science in the 21st century is to foster diversity, equity, and inclusion in occupational contexts."

(Burgoyne, Mashburn & Engle, 2021)


Wiesen (2021) - New York State Psychological Association Conference 14


NYSPA CONVENTION
 OCTOBER 19 - 21, 2021

Diversity Is a Legitimate Goal

- Federal *Uniform Guidelines* (UGESP, 1978)
 "... the user should include ... an investigation of **suitable alternative selection procedures** and suitable alternative methods of using the selection procedure which **have as little adverse impact as possible**"
 (7.3 Discrimination defined; B)


Wiesen (2021) - New York State Psychological Association Conference 15


NYSPA CONVENTION
 OCTOBER 19 - 21, 2021

Diversity Is a Legitimate Goal

- Federal Civil Rights Act of 1991
 "... **unlawful** ... for employment or promotion, to adjust the scores of, use different cutoff scores for, or otherwise **alter the results of**, employment related tests **on the basis of race, color, religion, sex, or national origin.**"
 (Sec. 106. Prohibition Against Discriminatory Use of Test Scores)


Wiesen (2021) - New York State Psychological Association Conference 16


NYSPA CONVENTION
 OCTOBER 19 - 21, 2021

Diversity Is a Legitimate Goal

- Is striving for diversity in hiring a legitimate goal for a psychologist?
 - Summary: **Yes, especially to reflect the community**
- My goal
 - Increase **both** diversity in hiring and expected job performance of Police Officers

Wiesen (2021) - New York State Psychological Association Conference 17



NYSPA CONVENTION
 OCTOBER 19 - 21, 2021

Psychometric Approaches

- (1) Do not rank on a M/C test of g
- (2) Select tests based on utility, not validity
- (3) Test job related abilities that have small group differences

(g = general mental ability, typically by an academic M/C test)
 (M/C = multiple choice)


Wiesen (2021) - New York State Psychological Association Conference 18


 NYSPA CONVENTION
 OCTOBER 19 - 21, 2021

(1) Do Not Rank Based on g

- g is typically measured with a M/C test
- M/C tests of g often cause AI
- Any composite with g will cause AI
- Evaluate carefully before letting one test in a battery create severe AI
- (Wiesen, 2018)


Wiesen (2021) - New York State Psychological Association Conference 19


 NYSPA CONVENTION
 OCTOBER 19 - 21, 2021

Do Not Rank Based on g

- But: g is the most valid predictor (**WRONG**)
- But: there is little beyond g (**WRONG**)

Wiesen (2021) - New York State Psychological Association Conference 20



 NYSPA CONVENTION
 OCTOBER 19 - 21, 2021

ACs More Valid Than g

- Schmidt & Hunter (1998) finding (**WRONG**)
- Drove employee selection for decades
- r of .51 for ability and .37 for ACs
 - Ability is the best predictor of job performance

(r = validity (correlation coefficient))
 (AC = Assessment Center)


Wiesen (2021) - New York State Psychological Association Conference 21


 NYSPA CONVENTION
 OCTOBER 19 - 21, 2021

ACs More Valid Than g

- “In contrast to Schmidt and Hunter’s ... reporting51 for ability and .37 for ACs, we found ... mean validity of **.22 for ability** and **.44 for ACs.**”
 (Sackett, Shewach & Keiser, 2017)
- Assessment centers have higher validity than tests of g , in general.
 - Why rank on a test with lower validity?


Wiesen (2021) - New York State Psychological Association Conference 22


 NYSPA CONVENTION
 OCTOBER 19 - 21, 2021

Low Validity of g for PO

- Aamodt (2004) meta-analysis
 - **$r = .24$ for job performance**
 - Corrected for all but predictor unreliability
 - Validity is higher for academy performance
 - Academy performance grade is largely based on written M/C tests

Wiesen (2021) - New York State Psychological Association Conference 23


 NYSPA CONVENTION
 OCTOBER 19 - 21, 2021

Law Enforcement Aptitude Battery

- Ability Test
 - Written Expression
 - Written Comprehension
 - Problem Sensitivity
 - Deductive Reasoning
 - Inductive Reasoning
 - Information Ordering

Wiesen (2021) - New York State Psychological Association Conference 24



Law Enforcement Aptitude Battery

- Cognitive Ability Test: used P/F and to rank
- Work Styles Questionnaire: used to rank
- Life Experience Survey used to rank



94% Unexplained Variance

- Explained variance = r^2
- For $r = .24$, $r^2 = .06$
- For $r = .24$, explained variance = .06
- Unexplained variance = $1 - \text{explained var.}$
- For $r = .24$, **94% of the variance in PO job performance is unexplained**



Validity of g Decreases with Time

- College GPA as proxy for g (and more)
- $r = .49$ at 1 year post college graduation
- $r = .33$ at 2-5 years post graduation
- $r = .12$ at 6 years post graduation (not signif.)
(Roth, BeVier, Switzer & Schippmann, 1996, all corrected r s)
- Not all studies of validity over time reach this conclusion



Personality More Valid with Time

- Consider the validity of conscientiousness as a predictor of medical school GPA
- $r = .18$ for year 1 of medical school
- $r = .45$ for year 7 of medical school
(Lievens, Ones & Dilchert, 2009)
- I am not aware of completely similar studies




How to Assure Sufficient g ?

- P/F use of a M/C test of g
- Measure g in other ways



P/F Use of a M/C Test of g


- Pros:
 - Assures a certain level of g
 - No adverse impact among those who pass
 - P/F use can have considerable utility
(Wiesen & Aguinis, 2010)
- Cons:
 - Lose some utility
 - Still have adverse impact for passing


NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

Measure g in Other Ways

- Use other tests of g
- Use proxies for g


Wiesen (2021) - New York State Psychological Association Conference 31


NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

M/C Tests of g Are Deficient

- Creative problem solving
- Correlation between g and leadership
- How other psychologists view g
- There are newer tests of intelligence


Wiesen (2021) - New York State Psychological Association Conference 32


NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

Creative Problem Solving

- Cannot fully measure creativity with a M/C test
 - Only open-ended questions measure originality
- Low correlation between creativity and g
 - Correlations from .07 to .29 (e.g., Kim, 2005; Sternberg, 2006, Tables 9, 11.1, 11.2)
- Creativity shows little adverse impact (Kaufman, 2010)


Wiesen (2021) - New York State Psychological Association Conference 33


NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

Correlation of g and Leadership

- Intelligence is the ability to deal with the demands of the real world in order to achieve success
- Leading others is a demand of the real world, for leaders
- $r = .19$ if g measured with M/C test
- $r = .6$ if g measured by ratings others make (Judge, Colbert & Ilies, 2004, Table 2)


Wiesen (2021) - New York State Psychological Association Conference 34


NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

How Other Psychologists View g

- “Most current large-scale testing is not well suited to [measure] higher order thinking skills, collaboration, information literacy, etc.” (Dolan, Goodman, Strain-Seymour, Adams & Sethuraman, 2011, page 5)

Wiesen (2021) - New York State Psychological Association Conference 35


NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

How Other Psychologists View g

- “Traditional tests of intelligence are not good proxies for rational thinking skills.” (Stanovich, West & Toplak, 2012)

Wiesen (2021) - New York State Psychological Association Conference 36



How Other Psychologists View g

- “Standard conventional tests only assess a narrow sampling of the abilities required for success in school and in life.”
(Sternberg, 2015)
 - His are ideas worth exploring



Newer Ways to Test Intelligence

- There are facets to g
- Some facets show smaller group differences
- Some composites of g yield less adverse impact, so more diverse hiring
- Facets of g are not equally valid for various different jobs
(e.g., Wee, Newman & Joseph, 2014)



Newer Ways to Test Intelligence

- There are some newer ways to test intelligence that show lower d values.
(e.g., Agnello, Ryan, Yusko, 2015)



Consider Other Measures of g

- Allow a college degree in lieu of a test of g
- Allow high school rank in lieu of a test of g
- Use these on a p/f basis or for ranking



Writing: An Essential g Ability

- Replace tests of writing ability with qualifying training courses with suitable final exams.
- Provide job aids for POs
 - Forms make it easier to describe incidents
 - OJT in report writing
- Some countries train POs for 2 years
 - In the US, PO require less training than barbers



(2) Choose Tests Based on Utility, Not Validity

- Utility \neq validity
- A **high** validity test can have **low** utility
- A **low** validity test can have **high** utility

NYS
21
NYSPA
CONVENTION
OCTOBER 15 - 17, 2021

Validity

- *SIOP Principles*:
“The degree to which **accumulated evidence** and theory **support specific interpretations** of scores from a selection procedure entailed by the proposed uses of that selection procedure.” (SIOP, 2018, page 50, emphasis added)

Wiesen (2021) - New York State Psychological Association Conference 43

NYS
21
NYSPA
CONVENTION
OCTOBER 15 - 17, 2021

Utility

- *SIOP Principles* mentions utility:
“**Projected productivity gains** or utility estimates for each employee and the organization due to use of the selection procedure...” (SIOP, 2018, page 33, emphasis added)
- We will focus here on **job performance**

Wiesen (2021) - New York State Psychological Association Conference 44

NYS
21
NYSPA
CONVENTION
OCTOBER 15 - 17, 2021

Utility ≠ Validity

- Ex. 1. Fewer applicants than job openings, so all applicants are hired
- Ex. 2. All applicants are equally qualified
- In both cases:
 - The validity of the test is irrelevant
 - The test has zero utility
- We need to evaluate utility

Wiesen (2021) - New York State Psychological Association Conference 45

NYS
21
NYSPA
CONVENTION
OCTOBER 15 - 17, 2021

What Drives Utility?

- Quality of applicants (Q)
 - Proportion of applicants who can do the job
- Selection ratio (SR)
 - Ratio of openings to applicants
- Validity (r)

(Cascio & Aguinis, 2011, pg 328)

Wiesen (2021) - New York State Psychological Association Conference 46

NYS
21
NYSPA
CONVENTION
OCTOBER 15 - 17, 2021

Decisions, Right and Wrong

Wiesen (2021) - New York State Psychological Association Conference 47

NYS
21
NYSPA
CONVENTION
OCTOBER 15 - 17, 2021

What Happens with Higher Q?

- We hire better people
- Less room for improvement over chance
 - Cannot do much better than hiring randomly
- **Utility is lower**

Wiesen (2021) - New York State Psychological Association Conference 48

NYSP 21 NYSPA CONVENTION OCTOBER 15 - 17, 2021

Expectancy Chart, $Q = .9$

Group	Chances of hires being successful ($r = .25$)	Chances of hires being successful ($r = .20$)
top 20%	95%	94%
top 40%	94%	93%
top 60%	93%	92%
top 80%	92%	91%
All	90%	90%

(Based on Taylor & Russell, 1939, page 577)

Wiesen (2021) - New York State Psychological Association Conference 49

NYSP 21 NYSPA CONVENTION OCTOBER 15 - 17, 2021

Expectancy Chart, $Q = .5$

Group	Chances of hires being successful ($r = .25$)	Chances of hires being successful ($r = .20$)
top 20%	64%	61%
top 40%	60%	58%
top 60%	56%	55%
top 80%	54%	53%
All	50%	50%

(Based on Taylor & Russell, 1939, page 575)

Wiesen (2021) - New York State Psychological Association Conference 50

NYSP 21 NYSPA CONVENTION OCTOBER 15 - 17, 2021

Less Valid Test w/ Higher Utility

Assume:
 90% of applicants have the g to do the job
 - e.g., if the PD requires a college degree
 50% of applicants have the personality to do the job

Area	g	Personality
Assumed Validity	.25	.2
Assumed Q	.9	.5
Utility	5%	11%

Wiesen (2021) - New York State Psychological Association Conference 51

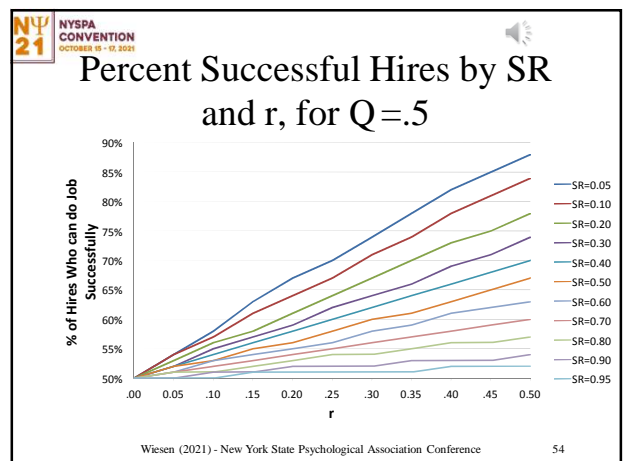
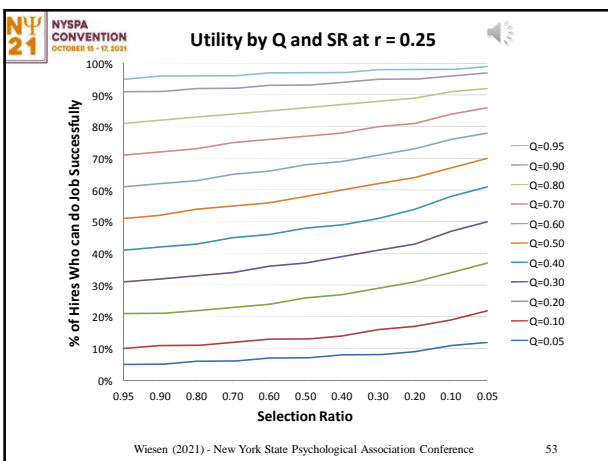
NYSP 21 NYSPA CONVENTION OCTOBER 15 - 17, 2021


Which More Important: Q , SR , r ?

- In theory: all three are very important
- In practice: Q and SR more easily changed
 - r is hard to change
- Better SR comes with worse adverse impact
- **Takeaway:**
Pay attention to recruitment (Q)

(SR = selection ratio)

Wiesen (2021) - New York State Psychological Association Conference 52





NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

(3) Test Job-Related Abilities With Small Group Differences

- Standardized mean score difference
- Standardized difference between the means of two groups (e.g., B & W)
- Denoted as d

Wiesen (2021) - New York State Psychological Association Conference 55



NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

How to Calculate d

- Standardized difference between the means of two groups

$$d = (MW - MB) / Sp$$
 - Where MB and MW are group means and Sp is the pooled estimated population standard deviation

Wiesen (2021) - New York State Psychological Association Conference 56



NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

What is d for g ?

- d of g is around 1.0, overall
 - White mean about 1 sd higher than Black mean
- Some find smaller d for job applicants (Bobko & Roth, 2013)

(sd = standard deviation)

Wiesen (2021) - New York State Psychological Association Conference 57



NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

Range of d

- Some KSAPs have small or zero d
- Some KSAPs have reverse d
- Some test modalities have low or zero d

(KSAP = knowledge, skill, ability, personal characteristic)


Wiesen (2021) - New York State Psychological Association Conference 58


NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

Low d , Job-Related Abilities

- Memory
- Creative and critical problem solving
- Oral communication
- Conscientiousness
- Ability to get along with others
- Structured oral exams
- New ways to measure intelligence

Wiesen (2021) - New York State Psychological Association Conference 59


NYSPA CONVENTION
 OCTOBER 15 - 17, 2021

Expect **Reverse** Impact

- Face recognition/memory for minority faces
- Implicit bias
 - Minority group members tend to be less biased than white people against members of their own minority group (Axt, Ebersole & Nosek, 2014, Table 1)

Wiesen (2021) - New York State Psychological Association Conference 60

NYS 21 NYSPA CONVENTION
OCTOBER 19 - 21, 2021

Expect **Reverse** Impact

- Knowledge of the community
- Community involvement
- Can make a logical argument for validity
 - Commitment to the community

Wiesen (2021) - New York State Psychological Association Conference 61

NYS 21 NYSPA CONVENTION
OCTOBER 19 - 21, 2021

Some Personality Areas

- Personality factors (e.g., conscientiousness) and facets with r 's in the .15 to .20 range and with small or zero d 's (e.g., Hough & Johnson, 2013)
- Extroversion and emotional stability (Ployhart & Holtz, 2008, Table 1)

Wiesen (2021) - New York State Psychological Association Conference 62

NYS 21 NYSPA CONVENTION
OCTOBER 19 - 21, 2021

Some Personality Areas

- Certain personality areas: achievement orientation, conscientiousness, surgency/dominance/potency facet of extroversion, and adjustment (Hough, Oswald & Ployhart, 2001, Table 2)
- Effort (care and persistence)
 - $r = .28-.33$ with college GPA (Briley, Domiteaux & Tucker-Drob, 2014)

Wiesen (2021) - New York State Psychological Association Conference 63

NYS 21 NYSPA CONVENTION
OCTOBER 19 - 21, 2021

Integrity Tests

- d around zero for race (Ones & Viswesvaran, 1998)
- Validity high ($r = .41$)
 - Highest incremental validity over g (Schmidt & Hunter, 1998)

Wiesen (2021) - New York State Psychological Association Conference 64

NYS 21 NYSPA CONVENTION
OCTOBER 19 - 21, 2021

Personality Validity Tends to Sum

- **Validities for uncorrelated variables sum**
 - Validities do not average
- Personality has low correlation with g
- Personality validities should tend to sum (Schmitt, 2014)
- Low validity tests sometimes have higher utility than higher validity tests

Wiesen (2021) - New York State Psychological Association Conference 65

NYS 21 NYSPA CONVENTION
OCTOBER 19 - 21, 2021

Structured Oral Exams

- Highly valid
 - the most valid, $r = .57$ (Aamodt, 2016, Table 5.2, page 194)
- d of zero (Levashina, Hartwell, Morgeson & Campion, 2014, Table 3, page 254)

Wiesen (2021) - New York State Psychological Association Conference 66



Rank on Structured Oral: Pros

- Validity may be higher than a test of g
- Content validity higher than test of g
- Can measure much more than a M/C g test (e.g., Cascio & Aguinis, 2011, p. 268, par. 4)
- Likely high candidate acceptance
 - Depends on the questions and interviewers
- Expect low or zero d



Rank on Structured Oral: Cons

- Difficult to maintain security of questions
- Lower reliability than a M/C test
- Inter-panel differences must be addressed
- Relatively costly to administer
- Candidate demand for transparency may threaten test security



Administrative Approaches

- (4) Consultants project the number of diversity hires
- (5) Focus on quality over quantity in recruiting
- (6) Use residency preference in hiring



(4) Consultants Project the Number of Diversity Hires

- Before the exam, consultants say:
We'll do all we can to reduce adverse impact
- When the exam results in few minority hires the appointing authority is left holding the bag



Watchword "No Surprises"

- Projections help to avoid surprises
- Empower PDs to make informed decisions on test areas/methods
- Improve PD decisions on consultant hiring



What to Project

- Number of hires by protected class
- Adverse impact
- Expected job performance (utility)

NYSP 21 NYSPA CONVENTION OCTOBER 15 - 17, 2021

Benefits of Projections

- Avoid surprises
- Empower managers to make tradeoffs
- Managers make decisions about:
 - Time/Cost
 - Expected number of diversity hires
 - Expected level of job performance

73

NYSP 21 NYSPA CONVENTION OCTOBER 15 - 17, 2021

Information Needed to Make Projections

- Number of applicants, by race/ethnic group
- Number of job openings
- *ds* for areas tested
- Intercorrelations of areas tested
- How subtest scores will be combined
- How candidates will be ranked

74

NYSP 21 NYSPA CONVENTION OCTOBER 15 - 17, 2021

Consultants' Likely Objections

- Not all variables are under their control
- Projections will be based on estimates
 - Number of applicants
 - Number of minority applicants
 - Number of job openings
- Variance of AI is high
- Projections can be wrong

75

NYSP 21 NYSPA CONVENTION OCTOBER 15 - 17, 2021

Build Projections into the RFP

- Collect the projections systematically
- Compare proposals on relevant outcomes
- Projections are rare but should be standard
- We have the data and theory to predict AI

(RFP = Request for proposals)
(AI = adverse impact)

76

NYSP 21 NYSPA CONVENTION OCTOBER 15 - 17, 2021

Proposal Comparison Form

- Form to structure data collection/summary

77

NYSP 21 NYSPA CONVENTION OCTOBER 15 - 17, 2021

Proposal Comparison Form

Police Officer Selection System Proposal Evaluation Form		
Topic	Selection System Approach 1	Selection System Approach 2
1. Projected Number of Hires		
Projected number of whites hired		
Projected number of blacks hired		
Projected number... etc.		

78

NYS 21 NYSPA CONVENTION OCTOBER 15 - 17, 2021

Proposal Comparison Form

Police Officer Selection System Proposal Evaluation Form

Topic	Selection System Approach 1	Selection System Approach 2
2. Projected Adverse Impact		
Ad Impact: black		
Ad Impact: Hispanic		
Etc.		

Wiesen (2021) - New York State Psychological Association Conference 79

NYS 21 NYSPA CONVENTION OCTOBER 15 - 17, 2021

Proposal Comparison Form

Police Officer Selection System Proposal Evaluation Form

Topic	Selection System Approach 1	Selection System Approach 2
3. Projected Job Performance		
Option 1. Mean using SAT-type scale		
Opt. 2. % hires who will be successful		
Opt. 3. Other job performance measure		

Wiesen (2021) - New York State Psychological Association Conference 80

NYS 21 NYSPA CONVENTION OCTOBER 15 - 17, 2021

Proposal Comparison Form

Police Officer Selection System Proposal Evaluation Form

Topic	Selection System Approach 1	Selection System Approach 2
6. The basis for ranking candidates		
Justification for selection of consultant		

Wiesen (2021) - New York State Psychological Association Conference 81

NYS 21 NYSPA CONVENTION OCTOBER 15 - 17, 2021

Form on Web

- URL for Excel file for evaluating proposals <http://jpwphd.com/form>

Wiesen (2021) - New York State Psychological Association Conference 82

NYS 21 NYSPA CONVENTION OCTOBER 15 - 17, 2021

Two Examples of Projections

- Predict AI based on assumptions for:
 - Number of applicants
 - Number of openings
 - d (standardized group difference between means)
 - Normal distribution

Wiesen (2021) - New York State Psychological Association Conference 83

NYS 21 NYSPA CONVENTION OCTOBER 15 - 17, 2021

Example: Assumption One

- 10,000 applicants
 - 9,000 White
 - 1,000 Minority
- 500 openings
- Selection ratio = .05 (i.e., 5% or 1 in 20)
- What if we double number of applicants?

Wiesen (2021) - New York State Psychological Association Conference 84

NYSP 21 NYSPA CONVENTION
OCTOBER 15 - 17, 2021

Example: Assumption Two

- 20,000 applicants
 - 18,000 White
 - 2,000 Minority
- 500 openings
- Selection ratio = .025 (i.e., 2.5% or 1 in 40)

Wiesen (2021) - New York State Psychological Association Conference 85

NYSP 21 NYSPA CONVENTION
OCTOBER 15 - 17, 2021

Project # of Diversity Hires

Total # of Applicants	# of Minority Applicants	# of Minority Hires	Adverse Impact Ratio
10,000	1,000	4	.08
20,000	2,000	3	.06

Wiesen (2021) - New York State Psychological Association Conference 86

NYSP 21 NYSPA CONVENTION
OCTOBER 15 - 17, 2021

Result of Additional Recruitment

- Adverse impact likely to be worse
- Fewer minority hires likely
- “The effects of group differences [AI] are greater as an organization becomes more selective.”
(Sackett & Ellingson, 1997, page 711)
- Sometimes intensive recruitment is **misguided**: ineffective and costly.

Wiesen (2021) - New York State Psychological Association Conference 87

NYSP 21 NYSPA CONVENTION
OCTOBER 15 - 17, 2021

(5) Focus on Quality Over Quantity

- Need high quality applicants
- Need high proportion of minority applicants
- Need highly qualified minority applicants

Wiesen (2021) - New York State Psychological Association Conference 88

NYSP 21 NYSPA CONVENTION
OCTOBER 15 - 17, 2021

Naive View of Testing

- Initial view: Selection tests work
- Tests are an effective and fair way to hire
- Tests will identify really good employees
- Recruit lots of applicants and hire the best
- Experienced view: Tests do not work well
 - Too many hiring errors

Wiesen (2021) - New York State Psychological Association Conference 89

NYSP 21 NYSPA CONVENTION
OCTOBER 15 - 17, 2021

Critique of Naive View of Testing

- Tests work but only to a modest extent
 - Tests improve a little over hiring by chance
- Recruitment should focus on quality
- Recruit widely, test, and choose high scorers **usually does not work**
- With $r = .24$, 1/3 to 2/3 of hires will fail
 - Under plausible assumptions

Wiesen (2021) - New York State Psychological Association Conference 90

NYS 21 NYSPA CONVENTION
OCTOBER 15 - 17, 2021

Recruitment Recommendations

- Develop recruitment sources
 - Make long term investment
- Need to find many minority applicants
 - Ideally higher qualified minority applicants
- Evaluate recruitment efforts systematically
(BTW - I am not a recruitment expert)

Wiesen (2021) - New York State Psychological Association Conference 91

NYS 21 NYSPA CONVENTION
OCTOBER 15 - 17, 2021

Develop Recruitment Sources

- College majors, such as:
 - Psychology
 - Sociology
 - Philosophy
 - History
- Offer paid internships
 - Goal: build interest

Wiesen (2021) - New York State Psychological Association Conference 92

NYS 21 NYSPA CONVENTION
OCTOBER 15 - 17, 2021

Develop Recruitment Sources

- Recruit before there is an exam scheduled
- Jr. High and High Schools
 - Courses in introduction to policing
- Summer and year-round programs
 - Internships (paid?)
 - Classroom
 - Office work
 - Project work

Wiesen (2021) - New York State Psychological Association Conference 93

NYS 21 NYSPA CONVENTION
OCTOBER 15 - 17, 2021

Recruitment Idea

- Consider raising the minimum qualifications (AKA entrance requirements) to increase applicant quality
 - Reliability
 - Ability to get along with others
 - Good attendance

Wiesen (2021) - New York State Psychological Association Conference 94

NYS 21 NYSPA CONVENTION
OCTOBER 15 - 17, 2021

Develop Recruitment Sources

- Some departments prefer to hire academy trained applicants
 - Tends to lead to adverse impact
- Encourage community groups to fund academy training for some community members

Wiesen (2021) - New York State Psychological Association Conference 95

NYS 21 NYSPA CONVENTION
OCTOBER 15 - 17, 2021

Evaluate Recruitment Sources

- Bring all the power of psychological research to the topic of recruitment

Wiesen (2021) - New York State Psychological Association Conference 96



(6) Grant Residency Preference

- Residency preferences are historic
 - Absolute, residents are hired first
 - Preference points, added to final exam score
- Absent residency preference, majority of POs can come from outside the city
- Use where poor city schools are surrounded by rich suburban schools
 - Unequal educational systems



What Type of Preference?

- Absolute
 - Hire residents before non-residents
 - A management decision
- Provide points
 - Based on SME judgment concerning validity
 - Equal to the SD on the M/C test of g , if such a test is used for ranking



Residency Preference: Pros

- Content valid
 - Know the community, dedicated to community
- Maintain use of traditional, valid predictors
- Historical precedent
- Very effective to increase diversity
- Practical to implement
- Candidates accept it as fair



Residency Preference: Cons

- May lower the utility of selection procedure
- Can be hard to determine residency
- Wealthier candidates can maintain two residences (i.e., fake residency)



Other Possible Preferences

- In the top X% of HS graduating class
- Predetermined number by zip code
- Graduate from a local high school
- Worked in the community for X years



A Dozen Takeaways

1. It is possible to increase both diversity of hires and expected job performance
2. Current M/C tests of g are poor predictors of PO job performance
3. A high proportion of PO hires are expected to fail on the job



A Dozen Takeaways

4. Current M/C tests of g often show severe adverse impact
5. Use of M/C tests of g to rank candidates, even in a composite, will result in adverse impact
6. Tests with low validity can have high incremental utility



A Dozen Takeaways

7. Use valid KSAPs & tests w/ low d
8. Use valid KSAPs and tests with reverse d
9. Consultants should project the number of diversity hires and level of expected job performance to enable PDs to make decisions about test content and mode, and the method of use of test scores.



A Dozen Takeaways

10. Recruitment is extremely important and must focus on quality over quantity (but quantity of minority applicants is important)
11. Residency preference can ameliorate differences in school district funding
12. We must try new approaches or we will not progress



Do We Face a Moral Issue?

- Psychologists in forefront of the **eugenics** movement; Immigration Acts of 1923/4
- Limit entry to the USA by country: quotas
- Favored England and Western Europe
- Southern and Eastern Europe not favored
- **Asians/Jews thought to be very low IQ**
- **How could psychologists be so wrong?**



Do We Face a Moral Issue?

- Are I/O psychologists making errors today similar to those made by the psychologists of 100 years ago who supported eugenics and restrictive immigration laws?
- M/C tests of g for PO **explain ONLY 6% of the variance** in job performance and cause severe adverse impact.



Topics Not Covered

- Statistical formulas
- Numeric examples
- Other approaches to reduce adverse impact
- Real life applications
- Some of this is on my website:
<http://jpwphd.com/papers>



Q&As

- Feel free to contact me at any time about this topic
 - (617) 244-8859
 - jpw@jpwphd.com
- See my website for references (and more):
<http://jpwphd.com/papers>

References

- Aamodt, M. G. (2004). *Research in Law Enforcement Selection*. Boca Raton, FL: Brown Walker Press.
- Aamodt, M. G. (2016) *Industrial-Organizational Psychology: An Applied Approach* (8th ed.) Boston: Cengage Learning.

References

- Agnello, P., Ryan, R. & Yusko, K. P. (2015) Implications of modern intelligence research for assessing intelligence in the workplace. *Human Resource Management Review* 25, 47–55.

References

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

References

- Axt, J. R., Ebersole, C. R., & Nosek, B. A. (2014). The rules of implicit evaluation by race, religion, and age. *Psychological Science*, 25, 1804-1815.

References

- Bobko, P. & Roth, P. L. (2013). Reviewing, Categorizing, and Analyzing Literature on Black-white Mean Differences for Predictors of Job Performance: Verifying Some Perceptions and Updating/correcting Others. *Personnel Psychology*, 66, 91-126.

References

- Briley, D. A., Domiteaux, M., & Tucker-Drob, E. M. (2014). Achievement-relevant personality: Relations with the Big Five and validation of an efficient instrument. *Learning and Individual Differences, 32*, 26-39.

References

- Burgoyne, A. P., Mashburn, C. A. & Engle, R. W. (2021). Reducing adverse impact in high-stakes testing. *Intelligence, 87* (101561), 1-7.

References

- Cascio, W. F. & Aguinis, H. (2011). *Applied Psychology in Human Resource Management*. Boston: Pearson.

References

- Dolan, R. P., Goodman, J., Strain-Seymour, E., Adams, J. & Sethuraman, S. (2011). *Cognitive Lab Evaluation of Innovative Items in Mathematics and English Language Arts Assessment of Elementary, Middle, and High School Students; Research Report*. San Antonio, TX: Pearson.

References

- Hough, L. M., & Johnson, J. W. (2013). Use and importance of personality variables in work settings. In I. B. Weiner (Ed.-in-Chief) & N. Schmitt & S. Highhouse (Vol. Eds.), *Handbook of psychology: Vol. 12. Industrial and organizational psychology* (pp. 211-243). New York: Wiley.

References

- Hough, L.M., Oswald, F.L. & Ployhart, R.E. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: issues, evidence and lessons learned. *International Journal of Selection and Assessment, 9*, 152-194.

References

- Judge, T. A., Colbert, A. E., & Ilies, R. (2004). Intelligence and leadership: A quantitative review and test of theoretical propositions. *Journal of Applied Psychology, 89*, 542-552.

References

- Kaufman, J. C. (2010). Using Creativity to Reduce Ethnic Bias in College Admissions. *Review of General Psychology, 14*, 189–203.

References

- Kim, K. H. (2005). Can Only Intelligent People Be Creative?; A Meta-Analysis. *The Journal of Secondary Gifted Education, 16*, 57–66.

References

- Levashina, J., Hartwell, C. J., Morgeson, F. P. & Campion, M. A. (2014). The structured employment interview: Narrative and quantitative review of the research literature. *Personnel Psychology, 67*, 241–293.

References

- Lievens, F., Ones, D. S., & Dilchert, S. (2009). Personality scale validities increase throughout medical school. *Journal of Applied Psychology, 94*, 1514-1535.

References

- Ones, D. S. & Viswesvaran, C. (1998). Gender, age, and race differences on overt integrity tests: Results across four large-scale job applicant data sets. *Journal of Applied Psychology, 83*, 35-42.

References

- Ployhart, R. E., & Holtz, B. C. (2008). The diversity-validity dilemma: Strategies for reducing racioethnic and sex subgroup differences and adverse impact in selection. *Personnel Psychology*, *61*, 153-172.

References

- Roth, P. L., BeVier, C. A., Switzer, F. S. & Schippmann, J. S. (1996). Meta-Analyzing the Relationship Between Grades and Job Performance. *Journal of Applied Psychology*, *81*, 548-556.

References

- Sackett, P. R. & Ellingson, J. E. (1997). The effects of forming multi-predictor composites on group differences and adverse impact. *Personnel Psychology*, *50*, 707-721.

References

- Sackett, P. R., Shewach, O. R. & Keiser, H. N. (2017). Assessment centers versus cognitive ability tests: challenging the conventional wisdom on criterion-related validity. *Journal of Applied Psychology*, *102*, 1435-1447.

References

- Schmidt, F. L. & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262-274.

References

- Schmitt, N. (2014). Personality and Cognitive Ability as Predictors of Effective Performance at Work. *Annual Review Organizational Psychology and Organizational Behavior*, *1*, 45-65.

References

- Stanovich, K. E., West, R. F., & Toplak, M. E. (2012). Judgment and decision making in adolescence: Separating intelligence from rationality. In V. F. Reyna, S. B. Chapman, M. R. Dougherty, & J. Confrey (Eds.), *The adolescent brain: Learning, reasoning, and decision making* (pp. 337-378). Washington, DC: American Psych. Assn.

References

- Sternberg, R. J. (2006). The Rainbow Project: Enhancing the SAT through assessments of analytical, practical, and creative skills. *Intelligence, 34*, 321–350.

References

- Sternberg, R. J. (2015). Successful intelligence: A model for testing intelligence beyond IQ tests. *European Journal of Education and Psychology, 8*, 76-84.

References

- Schmitt, N. (2014). Personality and Cognitive Ability as Predictors of Effective Performance at Work. *Annual Review Organizational Psychology and Organizational Behavior, 1*, 45-65.

References

- SIOP (2018) *Principles for the Validation and Use of Personnel Selection Procedures*. Washington, DC: American Psychological Association.

References

- Uniform Guidelines on Employee Selection Procedures. (1978). 29 C.F.R. Part 1607. (Also: 44 FR 11996, March 2, 1979 and 45 FR 29530, May 2, 1980.)

References

- Wee, S., Newman, D. A. & Joseph, D. L. (2014). More than g: Selection quality and adverse impact implications of considering second-stratum cognitive abilities. *Journal of Applied Psychology*, 99, 547–563.

References

- Wiesen, J. P. (2018). *Master Tutorial: Tools to Increase Diversity, Utility, and Validity in Hiring Police Officers*. 2018 Annual SIOF Conference; Minneapolis, MN.

References

- Wiesen, J. P. & Aguinis, H. (2010). *New Methods for Reducing Adverse Impact and Preserving Validity*. Symposium presentation at the 25th Annual Convention Society for Industrial and Organizational Psychology, Atlanta, GA.

Related Print and Audio Links

<http://jpwphd.com/papers>

Feel free to contact me about this topic:

- w@jpwphd.com
- (617) 244-8859 (land line)