

Reenvisioning Entry-Level and Promotional Selection to Improve Both Diversity and Job Performance

Joel P. Wiesen, Ph.D.
w@jpwphd.com & <http://jpwphd.com/papers>

47th Annual Conference
Society for Police and Criminal Psychology
Online; October 1, 2021

Wiesen (2021) - Society for Police and Criminal Psychology Conference

1

Why New Selection Approaches?

- Exams have adverse impact on minorities
 - Entry-level Police Officers (POs)
 - Promotion to Sergeant and above
- Re-evaluate current selection tools/methods
 - Consider new selection tools/methods
- Better job performance & more diverse POs

Wiesen (2021) - Society for Police and Criminal Psychology Conference

2

More Detailed Presentations

- Cannot fully survey literature in 15 minutes
- Some highlights in this presentation
- For more on this topic see:
<http://jpwphd.com/papers>

Wiesen (2021) - Society for Police and Criminal Psychology Conference

3

Selection of Entry-Level POs

- Problem: M/C g tests have adverse impact
- Solutions:
 - Do not rank based on g
 - Measure g in other ways
 - Use g on a pass-fail basis, if at all
 - Rank using other valid KSAPs

g = “general mental ability” (typically academic?)

Wiesen (2021) - Society for Police and Criminal Psychology Conference

4

Do Not Rank Based on g

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

Wiesen (2021) - Society for Police and Criminal Psychology Conference

5

ACs More Valid Than g

- Schmidt & Hunter (1988) 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) - Society for Police and Criminal Psychology Conference

6

ACs More Valid than g

- Sackett, Shewach, Keiser (2017)
“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.**”
- Assessment centers seem to have higher validity than tests of g , in general.
 - Why not rank on the test with highest validity?

Low Validity of g for PO

- Aamodt (2004) meta-analysis
 - $r = .27$, over corrected
 - Over corrected, for predictor unreliability
 - **$r = .24$ for job performance**
 - Corrected for all but predictor unreliability
- In general, test validity has not improved in our lifetimes
(Cascio & Aguinis, 2008, page 141)

Validity of g Decreases with Time

- 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
- All corrected r 's
- Roth, BeVier, Switzer & Schippmann (1996)

Personality More Valid with Time

- $r = .18$ to $r = .45$ for GPA, from year 1 to year 7 of medical school

(Lievens, Ones & Dilchert, 2009)

Utility

- SIOP *Principles*:
“Projected productivity gains or utility estimates for each employee and the organization due to use of the selection procedure” (SIOP, 2017, page 46)
- We will focus here on **job performance**
- Can consider diversity in evaluating utility
(Cascio & Aguinis, 2011, page 331)

Utility of Personality Can Be Higher than g

- The utility of using a low r test can be high
- The utility of using a high r test can be low
- Utility is the bottom line for an organization

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)

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**

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)

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)

Less Valid Test w/ Higher Utility

- Test of g: $r = .25$, $Q = .9$
 - i.e., 90% of applicants have the g to do the job
 - Utility of g is **5%**
- Test of personality : $r = .20$, $Q = .5$
 - Utility of personality is **11%**
- Rank based on personality!
 - E.g., if the PD requires a college degree

Low d, Job-Related Abilities

- Face recognition and memory
- Creative and critical problem solving
- Oral communication
- New ways to measure intelligence
- Structured oral exams
 - $d =$ standardized difference between means of two groups

Other Possible Test Areas

- Knowledge of the community
- Community involvement

Writing: An Essential *g* Ability

- Tests of writing ability can be replaced by qualifying training courses with suitable final exams.

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*
- There is statistical research that shows that P-F use of measures of *g* can have good utility
(Wiesen & Aguinis, 2010)

Promotion to Sergeant & Above

- Problems:
 - Low mean scores on M/C tests
 - M/C tests often have adverse impact
 - Emphasis on rote memorization
- Solutions:
 - Assure knowledge in other ways
 - Do not rank based on M/C tests
 - Rank on valid KSAPs other than a M/C test

Low Mean Scores on M/C tests

- What do the M/C tests measure
 - Important police knowledges
 - Must know, cannot look up on the job
 - Thousands of pages of SOPs, law, etc.
- Mean scores often low (in the 80's)
- Incumbents often fail the knowledge test
- We promote people with critical K gaps

Reenvision Promotional Tests

- Now: Issue reading list and then test
 - Rank based on rote knowledge of sources:
 - rules, procedures, and guidelines, and sometimes police science or other textbooks
 - But candidates don't learn all the sources
 - Sources do not teach problem solving
 - Hard to learn supervision from a book

Reenvision Promotional Tests

- Train to criteria on written sources
 - SOPs
 - Directives
 - Law
 - Textbooks (practical in orientation)
- Train to criteria on strategy and tactics
- Train to criteria on supervision, leadership

Ideas for Implementation

- Self-paced learning of sources
- Mastery tests to earn course credit
- Retake course if needed
- Many short courses on the sources

Time to Train POs

- USA: Average of **21 weeks** in academy
- Europe: Several **years** to become a PO
- Average # hours required to a **barber** is greater than # hours required to be a PO
- **Nurses train for 2-4 years.** Why not POs?

Source: <https://www.cbsnews.com/news/police-training-weeks-united-states/>

Source: <https://www.cnn.com/2016/09/28/us/jobs-training-police-trnd/index.html>

If We Are Too Conservative

- If we continue as we have in the past, the adverse impact will continue unabated.
- For list of references cited see:
 - <http://jpwphd.com/papers>

Joel P. Wiesen, Ph.D,
jpw@jpwphd.com
(617) 244-8859 (land line)

References

- Aamodt, M. G. (2004). *Research in Law Enforcement Selection*. Boca Raton, FL: Brown Walker Press.

References

- Cascio, W. F. & Aguinis, H. A. (2008). Staffing twenty-first century organizations. *Academy of Management Annals*, 2, 133–165.
- Cascio, W. F. & Aguinis, H. (2011) *Applied Psychology in Human Resource Management* (7th ed.) Saddle River, NJ: Prentice Hall.

References

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



References

- Roth, P. L., BeVier, C. A., Switzer III, F. A. & Schippmann, J. S. (1996). Meta-analyzing the relationship between grades and job performance. *Journal of Applied Psychology, 81*, 548-556.



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

- Taylor, H. C. & Russell, J. T. (1939). The relationship of validity coefficients to the practical effectiveness of tests in selection. *Journal of Applied Psychology, 23*, 565-578.



References

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

