





The Uniform Guidelines: Love Them, Leave Them, or Work to Change Them?

Presenters: Lance Anderson, HumRRO, Chair
Rich Tonowski, EEO, Panelist
Joel Wiesen, Applied Personnel Research, Panelist
Suzanne Tsacouris, HumRRO, Panelist

Our Panel

- Rich Tonowski, PhD, Chief Psychologist, US EEO
- Joel Wiesen, PhD, Director, Applied Personnel Research
- Suzanne Tsacouris, PhD, President and CEO, HumRRO
- You!

2

Innovative. Responsive. Impactful. 

Uniform Guidelines on Employee Selection Procedures

- 1978 guidance on how to comply with Title VII of the Civil Rights Act.
- Unchanged despite legal and scientific developments
- Calls to replace or revise (cf. McDaniel, Kepes, and Banks, 2011)

3

Innovative. Responsive. Impactful. 

Questions

Topic	Question
What's changed?	1. What issues remain? What new issues will need to be addressed?
	2. What progress have we made in identifying alternative selection procedures?
VG?	3. Any psychometric reasons NOT to rely on VG when designing selection systems?
	4. Any troubling practical issues in relying on VG?
Alternatives?	5. Are there any fairness considerations in relying on VG?
	6. Should the government have a role in defining technical concepts? If not, who should?
Future?	7. Do the UGESPs add anything beyond the SIOP Principles? Do we need both?
	8. For what legal issues has UGESP been overtaken, or might be?
	9. Is UGESP up to dealing with assessment issues on the horizon?

4

Innovative. Responsive. Impactful. 

1. What issues remain? What new issues will need to be addressed?

- Generalization of validity
- Statistical demonstration of adverse impact
- Gradations of validity
 - Effect size
 - Utility
 - Validity-diversity trade-off
 - Validity and Big Data

5

Innovative. Responsive. Impactful. 

2. What progress have we made in identifying alternative selection procedures?

6

Innovative. Responsive. Impactful. 

3. Any psychometric reasons NOT to rely on VG when designing selection systems?

3.1

- Last year's accepted VG wisdom is out.
- Tidal wave of changes in validity findings
- AC's are more valid than *g* (r's of **.44 & .22**)!
 - Sackett, Shewach & Keiser (2017)
 - "Head-to-head" comparisons
 - Only studies w both AC and *g* using same criterion
 - Partialing out *g* has no effect on validity of AC

7

Innovative. Responsive. Impactful. 

3. Any psychometric reasons NOT to rely on VG when designing selection systems?

3.2

- Schmidt & Hunter (1998) on job knowledge
- For $r = .48$, they ultimately cite Hunter (1982), an unpublished, unreplicable (IPMAAC) talk.
- Hunter overlooked 5 studies with very low *r*'s.
- No managerial or supervisory studies were meta-analyzed.
- This sleuth work was done by Leaetta Hough

8

Innovative. Responsive. Impactful. 

3. Any psychometric reasons NOT to rely on VG when designing selection systems?

3.3

- Decreases in validity over time.
 - e.g., Dahlke, Kostal, Sackett & Kuncel (2018)
- *g* had no significant correlation w/ physician job performance (n=773)
- Personality-related SJT correlated .21
- 9 year gap between predictor and criterion
 - Lievens & Sackett (2012)
- Do we hire for short term or long term?

9

Innovative. Responsive. Impactful. 

3. Any psychometric reasons NOT to rely on VG when designing selection systems?

3.4.

- VG ignores utility
- Lower validity predictors can have higher utility
- e.g., when prescreen on *g* and not personality
 - Wiesen (2018)
- Court considered this in *Smith v Boston*

10

Innovative. Responsive. Impactful. 

3. Any psychometric reasons NOT to rely on VG when designing selection systems?

3.5.

- Some illogical results
- Substantial improvements on retesting seen
- Both cognitive (.45 s.d.) and non-cognitive
 - Hausknecht (2010) with 15,000+ applicants

11

Innovative. Responsive. Impactful. 

4. Any troubling practical issues in relying on VG?

4.1.

- No operational definitions for VG test areas
 - Do we say it does not matter what we measure?
- Little guidance on test content, especially for assessment centers
- Assessment center is a method not an ability or a construct, yet VG treats it as a construct

12

Innovative. Responsive. Impactful. 

4. Any troubling practical issues in relying on VG?

4.2.

- Meta-analysis will not give a negative weight to a test area.
- Project A gave a negative weight to a reading test for mechanics.

13

Innovative. Responsive. Impactful.



4. Any troubling practical issues in relying on VG?

4.3.

- If meta-analysis supports job knowledge tests, how do we determine the content of our job knowledge tests?
- Do we still need local validation for all promotional exams?

14

Innovative. Responsive. Impactful.



5. Are there any fairness considerations in relying on VG?

5.1.

- Strong evidence that criteria are biased
- Women paid less than men
- Tall people paid more
- Pretty people paid more than homely
- Unfair treatment in the work place
- Only biased tests can predict biased criteria "fairly".

15

Innovative. Responsive. Impactful.



5. Are there any fairness considerations in relying on VG?

5.2.

- 1 s.d. difference for g but .5 s.d. d for criteria
- $y = r * x$ (y is job perf, x is predictor)
 - Assume $r = .5$:
- $y = .5 * x$
- So a difference of 1 on x equates to .5 on y
- But we select based on tests, so the difference on y should be much less. Also, $r \ll .5$.
- This is a psychometric conundrum

16

Innovative. Responsive. Impactful.



5. Are there any fairness considerations in relying on VG?

5.3.

- Complete reliance on VG will stifle change
- No VG for new tests
 - Agnello, Ryan, & Yusko (2015)

17

Innovative. Responsive. Impactful.



5. Are there any fairness considerations in relying on VG?

5.4.

- Differential validity exists
 - Aguinis, Culpepper & Pierce (2010)
 - Berry & Sackett (2013)
 - VG generally ignores differential validity

18

Innovative. Responsive. Impactful.



5. Are there any fairness considerations in relying on VG?

5.5.

- Focusing on most valid predictor may be invalid.
- "... if ... excluding ... components that could readily be assessed has a noticeable impact on selection rates for groups of interest ... the intended interpretation of test scores as predicting job performance in a comparable manner for all groups of applicants would be **rendered invalid.**" [emphasis added]
 - AERA, APA, NCME (2014, page 21, col 1, par 1)

5. Are there any fairness considerations in relying on VG?

5.6.

- More minority false negatives
- "a given selection score ... will often result in proportionately more false negative decisions in groups with lower mean test scores"
 - AERA, APA, NCME (1999, page 79)

6. Should the government have a role in defining technical concepts? If not, who should?

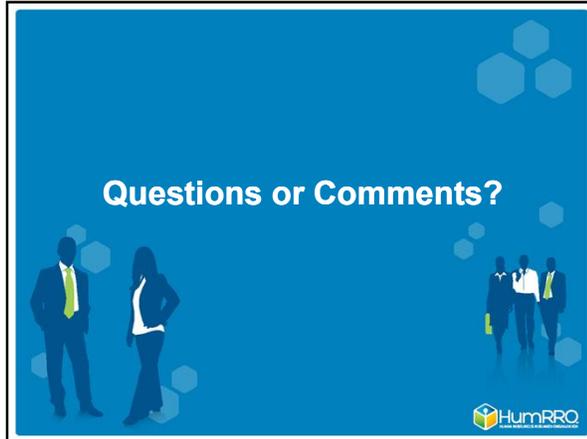
7. Do the UGESPs add anything beyond the SIOP Principles? Do we need both?

8. For what legal issues has UGESP been overtaken, or might be?

- UGESP is not regulation (except for record-keeping & federal contractors)
- Apply to age and disability cases (which recognize adverse impact theory)?
- Bottom line defense (old) and functionally integrated practice (new)
- Less discriminatory alternative: obligation now on plaintiff; evidence?
- Causation vs. correlation: questions for both adverse impact and validity.

9. Is UGESP up to dealing with assessment issues on the horizon?

- Big Question: Big Data
- Internet Applicant Rule
 - Proposed for UGESP Q&As
 - Adopted by OFCCP, opposed by EEOC
- Is Web-scraping a covered form of validity?
- Recruitment confounded with selection
 - UGESP does not apply to recruitment
 - What if recruitment amounts to selection?



Annotated Bibliography

- 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.
 - "Advances in ... from psychometric, neurocognitive, cross-cultural, and cognitive literatures are reviewed."

Innovative. Responsive. Impactful.



Annotated Bibliography

- Aguinis, H., Culpepper, S. A., & Pierce, C. A. (2010). Revival of test bias research in preemployment testing. *Journal of Applied Psychology*, 95, 648–680.
 - Researched and revived interest in what was once thought to be a dead issue.

Innovative. Responsive. Impactful.



Annotated Bibliography

- American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
 - These "Joint Standards" are much more demanding than the previous 1999 version, the 2003 SIOP *Principles*, or the 2017 draft of the *Principles*.

Innovative. Responsive. Impactful.



Annotated Bibliography

- Berry, C. M. & Sackett, P. R. (2013) The Role of Range Restriction and Criterion Contamination in Assessing Differential Validity by Race/Ethnicity. *Journal of Business Psychology*, 28, 345–359.
 - "The first differential validity study to simultaneously control for range restriction and individual differences in course choice, and answers a call to investigate potential causes of differential validity."

Innovative. Responsive. Impactful.



Annotated Bibliography

- Dahlke, J. A., Kostal, J. W., Sackett, P. R., & Kuncel, N. R. (2018). Changing Abilities vs. Changing Tasks: Examining Validity Degradation With Test Scores and College Performance Criteria Both Assessed Longitudinally. *Journal of Applied Psychology*. <http://dx.doi.org/10.1037/apl0000316>
 - "... changes to the determinants of performance over time have much stronger effects on validity degradation than does the self-life of cognitive test scores."

Innovative. Responsive. Impactful.



Annotated Bibliography

- Hausknecht, J. P. (2010). Candidate persistence and personality test practice effects: Implications for staffing system management. *Personnel Psychology, 63*, 299-324.
 - "... failing candidates pursued alternative response strategies upon retesting and generated dimension-level practice effects that reached .40 to .60, whereas passing candidates generally replicated their initial profiles."

Innovative. Responsive. Impactful.



Annotated Bibliography

- Lievens, F. & Sackett, P. R. (2012). The Validity of Interpersonal Skills Assessment Via Situational Judgment Tests for Predicting Academic Success and Job Performance. *Journal of Applied Psychology, 97*, 460-468.
 - "Four cohorts of medical students ($N = 723$) ... from admission to employment. Procedural knowledge about interpersonal behavior at the time of admission was valid for ... job performance (9 years later) and showed incremental validity over [invalid] cognitive factors."

Innovative. Responsive. Impactful.



Annotated Bibliography

- 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.
 - "We posit that 2 factors contribute to the differences in [Schmidt & Hunter, 1998] findings: (a) ACs being used on populations already restricted on cognitive ability and (b) the use of less cognitively loaded criteria in AC validation research."

Innovative. Responsive. Impactful.



Annotated Bibliography

- Wiesen, J. P. (2018). Tools to Increase Diversity, Utility, and Validity in Hiring Police Officers. SIOP 2018 Conference, Chicago, IL, April 20, 2018.
 - This "master tutorial" presents 15 tools (most novel or little-used) to help police departments (PDs) hire ethnically diverse academy classes while maintaining and even enhancing expected job performance (in 200+ PowerPoints).
<http://appliedpersonnelresearch.com/papers/>

Innovative. Responsive. Impactful.

