

Ways to Address Flawed Assumptions in Testing

Joel P. Wiesen, Ph.D.

Contact: jw@jpwphd.com

41st Annual SIOP Conference
New Orleans, LA; 4/30/2026

Wiesen (2026) SIOP Conference

1

1

Who Am I?

- 50+ years experience in testing
- 15 years at the MA civil service agency
- 35+ years as a consultant
- A scientist-practitioner

Wiesen (2026) SIOP Conference

2

2

Quick Takeaways

- Some test practices have weak support
- Might explain the ceiling on validity
- Suggest possible better practices

Wiesen (2026) SIOP Conference

3

3

Organization of Presentation

- Part 1: Brief historical perspective
- Part 2: 17 flawed or weak assumptions

- References on my website

<http://jpwphd.com/siop2026>

Wiesen (2026) SIOP Conference

4

4

Part 1: Historical Perspective

- Eugenics movement
 - NOT a fringe movement in early 1900s
 - Supported by leading psychologists
 - Found **Jews & Italians of low intelligence**
 - Federal restrictive Immigration Act of 1924
 - State law: sterilization, marriage restrictions

Wiesen (2026) SIOP Conference

5

5

Part 1: Historical Perspective

- Supported by leading psychologists
 - Terman (of Stanford-Binet fame)
 - Hall (first President of APA)
 - E. L. Thorndike (APA President)
 - Yerkes (Army Alpha and Beta)
 - Goddard
 - Bingham (invented SAT)
 - Watson (founder of Behaviorism)

Wiesen (2026) SIOP Conference

6

6

Part 1: Historical Perspective

- **Obvious flaw**
- English-language IQ tests were given to recent immigrants with limited English language facility

Wiesen (2026) SIOP Conference

7

7

Part 1: Historical Perspective

- Top psychologists supported eugenics
- Just a few dismissed the “findings”
- We need to be careful today: no repeats

Wiesen (2026) SIOP Conference

8

8

Part 2. Flawed Assumptions

- Assumption/Presumption
- Issue(s)
- Discussion
- Research Need
- Suggestions for Immediate Action

Wiesen (2026) SIOP Conference

9

9

Number Correct Predicts Best

Assumptions

Strengths compensate for weaknesses

All items equally predict overall ability

Wiesen (2026) SIOP Conference

10

10

Number Correct Predicts Best

Issues

Can pass despite failing crucial areas

SMEs not asked about item weight

Wiesen (2026) SIOP Conference

11

11

Number Correct Predicts Best

Discussion

Should these takers pass?

- 1) EMT aces anatomy but fails CPR
- 2) Nurse aces drugs but fails equipment
- 3) Police sergeant weak in juvenile law

Wiesen (2026) SIOP Conference

12

12

Number Correct Predicts Best

Discussion

Arguments against multiple hurdle

- 1) Results in high overall failure rate
- 2) High false negative rate due to error

Wiesen (2026) SIOP Conference

13

13

Number Correct Predicts Best

Research needed

Job analysis probe compensatory KSAPs

In past exams, how many passers had low scores on crucial KSAPs

Wiesen (2026) SIOP Conference

14

14

Number Correct Predicts Best

Immediate Action

Passing scores on all crucial KSAPs

Job analysis to probe compensatory and non-compensatory KSAPs

Wiesen (2026) SIOP Conference

15

15

Use z-scores to Equate

Assumption/Presumption

Good to equate using z-scores

Example 1 – Equate raters or panels
Applicants are rated by one of two panels
Different means and standard deviations
Use standardization to equate the panels

Wiesen (2026) SIOP Conference

16

16

Use z-scores to Equate

Assumption/Presumption

Example 2 – Combine scale scores
Combine scores on two tests that measure different KSAPs and have different means and standard deviations

Wiesen (2026) SIOP Conference

17

17

Use z-scores to Equate

Issue Example 1 – Equate raters/panels
What if one group has some superstars

After equating using z-scores:
Superstars get lower scores
Applicants graded w superstar suffer
Applicants graded w turkeys benefit

Wiesen (2026) SIOP Conference

18

18

Use z-scores to Equate

Issue Example 2 – Combine scale scores
Equating assumes similarly placed scores on the two scales are similar in meaning

Equating SDs may magnify trivial differences on a KSAP that has a small variance

Wiesen (2026) SIOP Conference

19

19

Use z-scores to Equate

Discussion Example 1 – Equate panels

If add a superstar to 15 candidates

- True superstar can go from $z = 5.2$ to 3

- Top score can go from $z = 2.2$ to 1.2

Wiesen (2026) SIOP Conference

20

20

Use z-scores to Equate

Discussion

Example 1 – Equate panels

Serious loss to miss a true superstar

Applicant scores no longer independent of each other if standardize

Wiesen (2026) SIOP Conference

21

21

Use z-scores to Equate

Discussion

Example 2 – Combine scale scores

Lose all score meaning when equate

Rare for scores to have no meaning

Wiesen (2026) SIOP Conference

22

22

Use z-scores to Equate

Research Need

Criteria for when to equate raters

Better ways to equate raters

Better ways to train raters

Better ways to combine 2 scale scores

Wiesen (2026) SIOP Conference

23

23

Use z-scores to Equate

Suggestions for Immediate Action

All raters grade some applicants, and equate based on those grades

For two different KSAPs, use scale anchors (e.g., minimally acceptable) and equate based on the anchors

Wiesen (2026) SIOP Conference

24

24

100 Items Are Sufficient

Assumption/Presumption

100 items provide good reliability

100 items allow for good KSAP coverage

Wiesen (2026) SIOP Conference

25

25

100 Items Are Sufficient

Issue

Low reliability of measurement

- 30 KSAPs, each with 3 questions?

An important KSAP may have 20 items

- Least important KSAP may have 1 item

Wiesen (2026) SIOP Conference

26

26

100 Items Are Sufficient

Discussion

Dangers of using only a few items:

Low reliability: High score error

Little assurance test taker is competent

Wiesen (2026) SIOP Conference

27

27

100 Items Are Sufficient

Research Need

Calculate KSAP reliability on past exams

Calculate pass rate with multiple hurdles

Wiesen (2026) SIOP Conference

28

28

100 Items Are Sufficient

Suggestions for Immediate Action

Reliably measure each critical KSAP

Use enough items to sample the domain

Wiesen (2026) SIOP Conference

29

29

Weight by Number of Items

Assumption/Presumption

Weight by number of items is most valid

Number of items should reflect import

Wiesen (2026) SIOP Conference

30

30

Weight by Number of Items

Issue

Some KSAPs may have very few items

Low-weight KSAPs are unreliable

The same number of items for areas of large and small scope is illogical

Wiesen (2026) SIOP Conference

31

31

Weight by Number of Items

Discussion

Court appalled at 1 item for 600 page book

Candidates don't study low weight KSAPs

Wiesen (2026) SIOP Conference

32

32

Weight by Number of Items

Research Need

Guidelines for determining # of items

Consider reliability

Consider scope of subject matter

Wiesen (2026) SIOP Conference

33

33

Weight by Number of Items

Suggestions for Immediate Action

Consider scope in setting number of items

Include enough items for reliable measurement of each KSAP

Wiesen (2026) SIOP Conference

34

34

Written Scenarios Mimic Job

Assumption/Presumption

Written situational items test KSAPs as used on job

Wiesen (2026) SIOP Conference

35

35

Written Scenarios Mimic Job

Issue

Superficial reflection of job

On the job, information must be sought, and responses generated, not recognized

Wiesen (2026) SIOP Conference

36

36

Written Scenarios Mimic Job

Issue

Examples for police sergeant:

- 1) A subject is verbally hostile but not physically aggressive, and officers appear frustrated.
- 2) You observe an officer becoming overly focused on one suspect while ignoring potential nearby threats.

Wiesen (2026) SIOP Conference

37

37

Written Scenarios Mimic Job

Discussion

It is becoming easier to create video items

GenAI can help create video items

Wiesen (2026) SIOP Conference

38

38

Written Scenarios Mimic Job

Research Need

Metric to evaluate similarity to job

Metric to evaluate quality of video items

Wiesen (2026) SIOP Conference

39

39

Written Scenarios Mimic Job

Suggestions for Immediate Action

Create a SIOP award for students who develop video items with GenAI

GenAI situational items workshop at next year's conference

Wiesen (2026) SIOP Conference

40

40

High Rater Reliability = Valid

Assumption/Presumption

High within-panel rater reliability indicates ratings are accurate

Wiesen (2026) SIOP Conference

41

41

High Rater Reliability = Valid

Issue

New data:

Two BARS rater panels grading the same exercise had reliabilities of **0.86 and 0.90** within-panel, but correlation between panels was only **0.48** (Wiesen, 2025, 2026)

Wiesen (2026) SIOP Conference

42

42

High Rater Reliability = Valid

Discussion

Low between-panel correlation suggests neither panel is very valid

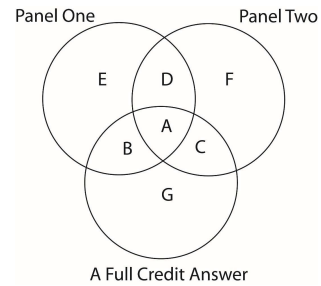
No longer accept high within-panel rater agreement as good indication of accuracy

Wiesen (2026) SIOP Conference

43

43

Figure 1. Validity and Invalidity



Wiesen (2026) SIOP Conference

44

44

High Rater Reliability = Valid

Research Need

Why do trained panels differ in grading?

Can panels articulate support for large differences in grades?

Wiesen (2026) SIOP Conference

45

45

High Rater Reliability = Valid

Research Need

Are between-panel rating differences related to the nature of the exercises, the areas rated, the anchors, or something else?

Wiesen (2026) SIOP Conference

46

46

High Rater Reliability = Valid

Suggestions for Immediate Action

Rate some/all candidates by two panels

Look at between-panel agreement

Follow up on disparate ratings of panels

Wiesen (2026) SIOP Conference

47

47

Quote sources → Defensibility

Assumption/Presumption

Using exact wording from sources (e.g., textbooks) avoids confusion and strengthens defensibility of items

To win item appeals, use source wording

Wiesen (2026) SIOP Conference

48

48

Quote sources → Defensibility

Issue

Such “quote” items often test rote memory rather than application of the knowledge

Wiesen (2026) SIOP Conference

49

49

Quote sources → Defensibility

Issue

Source wording is often unclear out of context

Wiesen (2026) SIOP Conference

50

50

Quote sources → Defensibility

Discussion

Recent court cases lost due, in part, to rote memory items

But, applicants ask how they can study for applications that are not in sources

Wiesen (2026) SIOP Conference

51

51

Quote sources → Defensibility

Research Need

How best to wean applicants from quotes?

Wiesen (2026) SIOP Conference

52

52

Quote sources → Defensibility

Research Need

How to defend key when applying source to a new situation?

Will SMEs agree on key?

Best balance of fluid and crystallized items

Wiesen (2026) SIOP Conference

53

53

Quote sources → Defensibility

Suggestions for Immediate Action

Include a certain % of understanding and application: test fluid as well as crystallized knowledge

Wiesen (2026) SIOP Conference

54

54

Quote sources → Defensibility

Suggestions for Immediate Action

Offer promotion training courses
(This topic deserves its own session.)

Wiesen (2026) SIOP Conference

55

55

Quote sources → Defensibility

Suggestions for Immediate Action

Use many situational items

Issue Rule that recognizes test
development SME interpretation of the
sources

Wiesen (2026) SIOP Conference

56

56

Hire From Top = Perform Well

Assumption/Presumption

Those at the top of the list will be well
qualified

The test will identify the best applicants
- Even if applicants are a mixed bag

Wiesen (2026) SIOP Conference

57

57

Hire From Top = Perform Well

Issue

With typical validity many top scorers will
not be top performers

Assume:

Validity = 0.30

50% of applicants can do the job

Top 10% on exam are hired

Wiesen (2026) SIOP Conference

58

58

Hire From Top = Perform Well

Issue

Many top scorers will not be top
performers

Results:

71% of hires can do job

29% of hires cannot do job

If hire randomly, 50% of hires can do job

Wiesen (2026) SIOP Conference

59

59

Hire From Top = Perform Well

Issue

More results:

Only 22% of hires are in top 10% on job

Most of top 10% on job are not hired

3% of hired are in the bottom 10% on job

Wiesen (2026) SIOP Conference

60

60

Hire From Top = Perform Well

Issue

Validity often below 0.3, esp. for police

Issue

Critical KSAPs sometimes ignored (e.g., face memory for police officer, esp. cross ethnic group)

Wiesen (2026) SIOP Conference

61

61

Hire From Top = Perform Well

Issue

Promoting people who barely pass happens in small departments

Wiesen (2026) SIOP Conference

62

62

Hire From Top = Perform Well

Discussion

Need more valid tests

Wiesen (2026) SIOP Conference

63

63

Hire From Top = Perform Well

Research Need

Improved predictive validity

Wiesen (2026) SIOP Conference

64

64

Hire From Top = Perform Well

Suggestions for Immediate Action

Conduct criterion-related validity studies whenever feasible

Wiesen (2026) SIOP Conference

65

65

Tests Have Sufficient Validity

Assumption/Presumption

Tests are sufficiently valid to use

Validity generalization: all tests are valid

Wiesen (2026) SIOP Conference

66

66

Tests Have Sufficient Validity

Issue:

Validity is higher for training than job
Validity coefficient can change with time
The level of validity varies by job and test
Many false positives and false negatives
- Higher false negative rate for minorities

Wiesen (2026) SIOP Conference

67

67

Tests Have Sufficient Validity

Issue: Test-retest reliability can be low
.23-.30 for one-year retest reliability for three types of performance tests
.48 for M/C job knowledge retest reliability
.66 for 1 wk SJT retest (alternate versions)
.45-.63 for M/C and constructed response
Low reliability means low validity

Wiesen (2026) SIOP Conference

68

68

Tests Have Sufficient Validity

Issue

Criterion and content validity not identical

Project A: verbal and quantitative had negative weights for Vehicle Mechanic
- No content validity study would find that

Wiesen (2026) SIOP Conference

69

69

Tests Have Sufficient Validity

Issue

Some managers and applicants hate tests due to false positives and false negatives

Wiesen (2026) SIOP Conference

70

70

Tests Have Sufficient Validity

Discussion

Do not oversimplify this complex topic

Must improve validity

Wiesen (2026) SIOP Conference

71

71

Tests Have Sufficient Validity

Research Need

Targeted VG studies, e.g., Police Sgt.

Ways to improve reliability of predictors

Wiesen (2026) SIOP Conference

72

72

Tests Have Sufficient Validity

Suggestions for Immediate Action

Temper broad VG claims

Collect test-retest reliability data

Wiesen (2026) SIOP Conference

73

73

Our Tests Are Fair

Assumption/Presumption

Our tests are fair

- perhaps overestimate job performance
of minority test takers

Wiesen (2026) SIOP Conference

74

74

Our Tests Are Fair

Issue

Thorndike's definition is equally logical

Clearly fair tests can be Thorndike unfair

Among qualified candidates, the
**selection rate will be smaller for the
lower scoring group**

Wiesen (2026) SIOP Conference

75

75

Our Tests Are Fair

Issue

Clearly: Are slopes and intercepts equal?

Thorndike: Is the selection rate equal
among equally qualified candidates no
matter what group membership?

Wiesen (2026) SIOP Conference

76

76

Our Tests Are Fair

Issue

B-W mean test score difference is twice
as large as job performance difference

- This seems unfair
- What would a jury say about this?

Wiesen (2026) SIOP Conference

77

77

Our Tests Are Fair

Issue

Clearly definition assumes a fair criterion

Men paid more than women
Tall people paid more than short
Attractive people paid more than homely

Wiesen (2026) SIOP Conference

78

78

Our Tests Are Fair

Issue

Predictive bias exists

Statistical power of differential validity studies is low

- Type 2 error: fail to find differential validity when it exists

Wiesen (2026) SIOP Conference

79

79

Our Tests Are Fair

Discussion

Clear definition favors the testing profession that accepted it

Which approach would a court choose?

Wiesen (2026) SIOP Conference

80

80

Our Tests Are Fair

Research Need

Renewed focus on definitions of test fairness

When each definition is most appropriate

Wiesen (2026) SIOP Conference

81

81

Our Tests Are Fair

Research Need

Machine learning apps are being evaluated in terms of equalize odds

Wiesen (2026) SIOP Conference

82

82

Our Tests Are Fair

Suggestions for Immediate Action

Consider more than one definition

Consider power when testing differences in slopes and intercept

Wiesen (2026) SIOP Conference

83

83

Job Analysis to Outline is Easy

Assumption/Presumption

Task and KSAP analyses easily translate into a clear test blueprint (AKA outline)

Wiesen (2026) SIOP Conference

84

84

Job Analysis to Outline is Easy

Issue

1. Tasks and KSAPs described by short phrases are inadequate for SMEs to rate or to guide test development

Consider: "Knowledge of state and federal law" (what aspects of the law?)

Wiesen (2026) SIOP Conference

85

85

Job Analysis to Outline is Easy

Issue

2. KSAPs vary in scope (e.g., one management knowledge vs split into multiple KSAPs)

Equal item counts for equally important KSAPs may ignore scope differences

Wiesen (2026) SIOP Conference

86

86

Job Analysis to Outline is Easy

Issue

3. Often job analysis rates 2 or more content sources: tasks, KSAPs, and (sometimes) books, SOPs, etc.

A two or three-way test outline is cumbersome to develop and use

Wiesen (2026) SIOP Conference

87

87

Job Analysis to Outline is Easy

Issue

4. Job analysis findings can be illogical

Sgt brings officers up on charges daily

A budgeting/spreadsheet management job requires no math

Wiesen (2026) SIOP Conference

88

88

Job Analysis to Outline is Easy

Issue

5. Sometimes find considerable disagreement between SMEs

e.g., Both highest and lowest ratings

Wiesen (2026) SIOP Conference

89

89

Job Analysis to Outline is Easy

Issue

6. Assumes one ideal set of KSAPs

Can use different sets of KSAPs to do job

Example: Female firefighters lift ladder off truck on shoulder, males lift over head

Wiesen (2026) SIOP Conference

90

90

Job Analysis to Outline is Easy

Discussion

A number of unaddressed or unresolved issues with job analysis

Some are related: imprecise wording and divergent SME ratings

Wiesen (2026) SIOP Conference

91

91

Job Analysis to Outline is Easy

Research Need

1. Develop criteria for a KSAP worded with enough detail to write items

2. What size (scope) KSAP is best?
Develop a metric for scope

Wiesen (2026) SIOP Conference

92

92

Job Analysis to Outline is Easy

Research Need

3. Proper role of each job analysis rating scale in test outline (task frequency/criticality, KSAP importance, source importance)

Wiesen (2026) SIOP Conference

93

93

Job Analysis to Outline is Easy

Research Need

4. How to deal with rater disagreements

5. How to fix illogical job analysis results

Wiesen (2026) SIOP Conference

94

94

Job Analysis to Outline is Easy

Research Need

6. How to measure fungibility of KSAPs

How to determine the degree of diversity of KSAPs needed or desirable

Select by team or individual employee

Wiesen (2026) SIOP Conference

95

95

Job Analysis to Outline is Easy

Suggestions for Immediate Action

Before collecting job analysis data, determine how job analysis data/ratings will be used to develop the test outline

Define KSAPs precisely enough to guide test content development

Wiesen (2026) SIOP Conference

96

96

Job Analysis to Outline is Easy

Suggestions for Immediate Action

Use multiple job analysis methods

Address weaknesses above

Wiesen (2026) SIOP Conference

97

97

SMEs Can Brainstorm KSAPs

Assumption/Presumption

Incumbents make the best SMEs

SMEs will identify all job-relevant KSAPs

SMEs will word KSAPs clearly

Wiesen (2026) SIOP Conference

98

98

SMEs Can Brainstorm KSAPs

Issue

Incumbents may lack KSAP vocabulary

SMEs know tasks better than KSAPs

Wiesen (2026) SIOP Conference

99

99

SMEs Can Brainstorm KSAPs

Issue

SMEs may even ignore the official job description

Wiesen (2026) SIOP Conference

100

100

SMEs Can Brainstorm KSAPs

Issue

Grit

Fleishman areas

Big-Five facets

Achievement motivation

Locus of control

Peripheral vision

Wiesen (2026) SIOP Conference

101

101

SMEs Can Brainstorm KSAPs

Directional hearing

Hearing in noisy environments

Ability to understand speech (speed, accents, impediments)

Cognitive empathy (what others think)

Emotional empathy (what others feel)

Endurance

Wiesen (2026) SIOP Conference

102

102

SMEs Can Brainstorm KSAPs

Analytic, practical, creative intelligence
Delay of gratification
Striving for competence and mastery
Learning goal orientation
Self-development
Self-control
Vocational interest

Wiesen (2026) SIOP Conference

103

103

SMEs Can Brainstorm KSAPs

Attentional control
Mood stability

Wiesen (2026) SIOP Conference

104

104

SMEs Can Brainstorm KSAPs

Discussion

We are experts in KSAPs, SMEs are not
Supervisors and trainers focus on KSAPs

Wiesen (2026) SIOP Conference

105

105

SMEs Can Brainstorm KSAPs

Suggestions for Immediate Action

Use supervisors/trainers as SMEs also

Compare with official job description, and
other companies or organizations

Search the psychometric literature

Wiesen (2026) SIOP Conference

106

106

SMEs Can Brainstorm KSAPs

Research Need

List of testable, work-related KSAPs

Review non-I/O literature
Cognitive psychology
Differential psychology
Etc.

Wiesen (2026) SIOP Conference

107

107

SMEs Can Brainstorm KSAPs

Suggestions for Immediate Action

Develop job analysis procedures that
include psychologist input on KSAPs

Wiesen (2026) SIOP Conference

108

108

Task/KSAP Statements Suffice for Test Development

Assumption/Presumption

Task/KSAP lists suffice for content validity

Task/KSAP lists suffice for test development

Wiesen (2026) SIOP Conference

109

109

Task/KSAP Statements Suffice for Test Development

Issue

Task/KSAP lists provide little insight into why the job exists, what job tasks are hard or easy, or a typical day's work.

Often such lists are not detailed enough to guide simulation or M/C item writing.

Wiesen (2026) SIOP Conference

110

110

Task/KSAP Statements Suffice for Test Development

Discussion

Knowledge of the State Code

Respond to incidents & assume command

What topics should be tested?

Wiesen (2026) SIOP Conference

111

111

Task/KSAP Statements Suffice for Test Development

Discussion

Small number of test development SMEs

Judgment of a handful of test development SMEs to clarify ratings of hundreds of job analysis SMEs

Wiesen (2026) SIOP Conference

112

112

Task/KSAP Statements Suffice for Test Development

Research Need

Guidelines for an adequate task/KSAP

Wiesen (2026) SIOP Conference

113

113

Task/KSAP Statements Suffice for Test Development

Suggestions for Immediate Action

Collect critical incidents in addition to survey data (preferably 100+ incidents)

Wiesen (2026) SIOP Conference

114

114

Use Task/KSAP Categories to Reduce SME Fatigue

Assumption/Presumption

Category ratings suffice for test development

Wiesen (2026) SIOP Conference

115

115

Use Task/KSAP categories to Reduce SME Fatigue

Issue

Categories are worse than unclear tasks or KSAPs, hide key details

Idiosyncratic SME views of categories make the meaning of ratings unclear

Wiesen (2026) SIOP Conference

116

116

Use Task/KSAP categories to Reduce SME Fatigue

Discussion

Trade-off between rater fatigue and clarity of what is being rated

Wiesen (2026) SIOP Conference

117

117

Use Task/KSAP categories to Reduce SME Fatigue

Research Need

How to measure rater fatigue in job analysis

Other ways to mitigate rater fatigue

Wiesen (2026) SIOP Conference

118

118

Use Task/KSAP categories to Reduce SME Fatigue

Suggestions for Immediate Action

Use clear individual task or KSAP statements when developing test outlines

Mitigate rater fatigue in other ways

Wiesen (2026) SIOP Conference

119

119

Content = Criterion Validity

Assumption/Presumption

Content and criterion-related validity are basically equal in application

Same test content with both approaches

Wiesen (2026) SIOP Conference

120

120

Content = Criterion Validity

Issue

Agreement between criterion and content validity not certain

Project A found negative beta weights for math and verbal abilities for Vehicle Mechanic

Wiesen (2026) SIOP Conference

121

121

Content = Criterion Validity

Issue

Different prediction formulas for different criteria

Project A: Each of 4 job criteria measures had a different prediction equation.

Wiesen (2026) SIOP Conference

122

122

Content = Criterion Validity

Discussion

Consultants propose content studies even with large N

Wiesen (2026) SIOP Conference

123

123

Content = Criterion Validity

Research Need

Compare test outlines based on content and criterion-related validation

For what jobs are content test outlines similar to criterion-related test outlines

Wiesen (2026) SIOP Conference

124

124

Content = Criterion Validity

Suggestions for Immediate Action

Use criterion validity whenever possible

More cross organization validity studies

Wiesen (2026) SIOP Conference

125

125

Recruitment Boosts Diversity

Assumption/Presumption

Recruiting more minority applicants will result in more minority hires

Hiring will tend to reflect recruitment demographics

Wiesen (2026) SIOP Conference

126

126

Recruitment Boosts Diversity

Issue

More recruitment with the same % minority will exacerbate adverse impact

Lower selection ratios yield more severe adverse impact

Wiesen (2026) SIOP Conference

127

127

Recruitment Boosts Diversity

Discussion

Fewer applicants increase diversity hires

Higher proportion of minority applicants

Higher quality minority applicants

Wiesen (2026) SIOP Conference

128

128

Recruitment Boosts Diversity

Discussion

Recruit better minority applicants

Effective recruitment is the most practical way to improve job performance

Improving validity is very hard

Wiesen (2026) SIOP Conference

129

129

Recruitment Boosts Diversity

Research Need

Can we predict which recruitment sources will yield the best applicants?

Best ways to attract better applicants?

Wiesen (2026) SIOP Conference

130

130

Recruitment Boosts Diversity

Suggestions for Immediate Action

Focus on quality of applicants, not just quantity

Focus on the ethnic mix of applicants while maintaining applicant quality

Wiesen (2026) SIOP Conference

131

131

Focus on Validity, Not Utility

Assumption/Presumption

We should design tests based on validity

Wiesen (2026) SIOP Conference

132

132

Focus on Validity, Not Utility

Issue

Managers are interested in utility only

Utility and validity are related, not identical

Wiesen (2026) SIOP Conference

133

133

Focus on Validity, Not Utility

Discussion

Utility is a mathematical function of:

- Validity
- Quality of applicants
- Selection ratio

Validity is the hardest to improve

Wiesen (2026) SIOP Conference

134

134

Focus on Validity, Not Utility

Research Need

How to measure applicant quality

How to recruit high quality applicants

How to design a test based on utility

Wiesen (2026) SIOP Conference

135

135

Focus on Validity, Not Utility

Suggestions for Immediate Action

I/Os get involved in recruitment

Predict utility numerically

Probe skill gaps with management

Wiesen (2026) SIOP Conference

136

136

Mini Wrap-Up

Improve & develop new testing methods

Note: today's topics are not exhaustive

Wiesen (2026) SIOP Conference

137

137

Presentation Website

- Copy of slides
- Audio of presentation
- Miscellaneous other presentations

<http://jpwphd.com/siop2026>

Wiesen (2026) SIOP Conference

138

138

Presentation Citation

Wiesen, J. P. (2026). Ways to Address Flawed Assumptions in Testing [Master Tutorial]. *Society for Industrial and Organizational Psychology Annual Conference*, New Orleans, LA, United States.

Wiesen (2026) SIOP Conference

139

139

References

All in pdfs on my website

<http://jpwphd.com/siop2026>

Wiesen (2026) SIOP Conference

140

140

Q&A

Wiesen (2026) SIOP Conference

141

141