

Minimizing Adverse Impact: Learning From Some Possible Serious Flaws in the Wonderlic Personnel Test

Joel P. Wiesen, Ph.D, Director
Applied Personnel Research
wiesen@personnelselection.com

29th Annual IPMAAC Conference
Orlando, Florida; June, 2005

Why Discuss the Wonderlic

- Perhaps the most widely used commercial test in our field
- Enjoying success in court

Features

- Short, timed test: 12 minutes
- Aims to test general cognitive ability
- Yields one score
- Many validity studies

One Major Problem

- Adverse Impact (AI)
- B-W mean score difference almost 1 s.d.
- Not unique to the Wonderlic

Possible Causes of AI

- Many possible causes of AI
- Dozens of possible reasons for AI
- Nevertheless, let's not simply assume the test is unbiased.
- We will look closely at the Wonderlic

Indications of Possible Test Bias

- Copyright before AI was discussed: 1942
- cursory review of questions raises red flag

Review of Test Materials

- Test Items
- Test Manual

Wonderlic Instructions

First Sample Question

REAP the opposite of

1 obtain, **2** cheer, **3** continue,

4 exist, **5** sow

Wonderlic Instructions

Third Sample Question

MINER MINOR -- Do these words

1 have similar meanings,

2 have contradictory meanings,

3 mean neither the same or opposite?

Wonderlic Personnel Test

Question omitted from this file for the sake of security.

Confidential. For Conference Only. Do Not Reproduce or Distribute.

By Dr. Joel P. Wiesen; Applied
Personnel Research; (617) 244-
8859; IPMAAC Conference, June

Wonderlic Personnel Test

Question omitted from this file for the sake of security.

Confidential. For Conference Only. Do Not Reproduce or Distribute.

By Dr. Joel P. Wiesen; Applied
Personnel Research; (617) 244-
8859; IPMAAC Conference, June

Wonderlic Personnel Test

Question omitted from this file for the sake of security.

Confidential. For Conference Only. Do Not Reproduce or Distribute.

By Dr. Joel P. Wiesen; Applied
Personnel Research; (617) 244-
8859; IPMAAC Conference, June

Wonderlic Personnel Test

Question omitted from this file for the sake of security.

Confidential. For Conference Only. Do Not Reproduce or Distribute.

By Dr. Joel P. Wiesen; Applied
Personnel Research; (617) 244-
8859; IPMAAC Conference, June

Reasoning or Vocabulary

If you do not know the words used, an item measures vocabulary rather than reasoning.

Wonderlic Personnel Test

Question omitted from this file for the sake of security.

Confidential. For Conference Only. Do Not Reproduce or Distribute.

By Dr. Joel P. Wiesen; Applied
Personnel Research; (617) 244-
8859; IPMAAC Conference, June

Wonderlic Personnel Test

Question omitted from this file for the sake of security.

Possibly Biased Questions

- About 9 such potentially biased questions
- B-W mean score difference is about 7

New Types of Test Items

- “The field of testing suffers from ideational stagnation”

Robert Sternberg

1998 American Psychologist

New Types of Test Items

- Examples of new or improved items
- Examples of new item types

New or Improved Items

- Verbal Reasoning
- Math
- Clerical Speed
- Face Recognition
- Short-Term Memory
- Oral Comprehension
- Mechanical Ability
- Intelligence
- Spatial Ability
- Physical Ability

Old (Wonderlic) Math Item

Question omitted from this file for the sake of security.

New Math Item

A store is having a 20% off sale, and you see a jacket selling for \$100. How much did the jacket cost before the sale?

New Clerical Speed Item

- Names not all Anglo-Saxon
- Different fonts (more common today)

New Clerical Speed Item

How many of the six pairs of names listed below are exactly the same?

Shiva, Vandanna

Shiva, Vandana

Rivierra, Luis

Riviera, Luis

Fajardo, Makeda

Fajardo, Makieda

Rasheed, Fouad

Rashed, Fouad

Abdul, Hakeem

Abdul, Hakeim

Hernandez, Luis

Hernandes, Luis

Problem Solving or Concept Formation

- Seldom used item type, but similar to items in the Wonderlic
- Simple instruction:
 - Select the one that does not belong.

Problem Solving Item Type

- Select the one that does not belong.
 - truck
 - car
 - motorcycle
 - bicycle

Problem Solving Item Type

- Select the one that does not belong.
 - wall
 - floor
 - door
 - chair

Hard to Identify Bias by Content

- Content reviews are generally ineffective in detecting items with large B-W differences
- Many suggest it is futile to try
- The issue of test bias is scientifically dead (Hunter & Schmidt, 2000)

SAT Has Ethnic Bias

- Most B-W DIF with analogies and antonyms (which provide least context)
- Least B-W DIF with reading comprehension items (with maximal verbal context) (Freedle, 2003)

Success in Reducing Adverse Impact

- The WTMA, my test of mechanical aptitude, reduces adverse impact on women as compared with the Bennett
- WTMA uses every day objects
- The test is available from PAR and Ramsay Corporation

Second Major Problem

- Face validity appears low
- Face validity promotes user acceptance

Conclusion to Item Review

- Be creative in writing items
- Consider new item types
- Pursue face validity to maximize user acceptance

Wonderlic Test Manual

- Not clear what the test was designed to measure
- No test outline
- No mention of adverse impact

Why Does the Wonderlic Work?

- Proxy for schooling
- Proxy for knowledge of certain American subculture

Not Just the Wonderlic

- Other published tests have similar content
- Perhaps because published tests tend to be quite old.

Some Closing Comments

- Don't be content with status quo
- Be vigilant in cultural bias review
- Conduct research on fairness in job performance evaluations.
- Consider developing or using tests of non-traditional abilities

Call for Research Partners

- A new General Employment Test is being developed
- Contact me to discuss being a beta or validation site

References

- Freedle, R.O. (2003) Correcting the SAT's ethnic and social-class bias: A method for reestimating SAT scores. *Harvard Educational Review, 73*, 1-43.
- Hunter, J.E. & Schmidt, F.L. (2000) Racial and Gender Bias in Ability and Achievement Tests. *Psychology, Public Policy, and Law, 6*, 151-158.
- Sternberg, R.L. (1998) *American Psychologist, 53*, 576-77.
- Wiesen, J.P. (2005) *Possible Reasons for the Black-White Mean Score Differences Seen With Many Cognitive Ability Tests*
<http://personnelselection.com/adverse.impact.htm>