TESTING TRIAGE: GREATER ASSESSMENT EFFECTIVENESS AT LOWER COST

Sharon Shrock and William Coscarelli

Testing Triage:
Greater Assessment Effectiveness at Lower Cost

Sharon Shrock, SAShrock@siu.edu and William Coscarelli, coscarel@siu.edu, Southern Illinois University, Carbondale, USA

Objectives

This session will enable participants to do the following:

- Increase validity and efficiency in testing by using hierarchical analysis
- Test beyond the memorization level using knowledge tests that match jobs
- Determine when to calculate reliability and establish inter-rater reliability
- Establish content validity
- Establish a legally defensible cut-off/mastery score

Session Goal

To convince you that testing in an organization can be vastly improved with relatively little additional investment – if you know how.

Importance of testing

- Provide feedback to persons regarding their competence ⇒ informs learner
- Determine whether persons can do a job or task ⇒ informs organization
- Check effectiveness of instruction or other PT intervention ⇒ informs training or HRD department
- Establish return on investment attributable to training or other PT intervention ⇒ informs CEO, et. al.

Why isn't testing done?

Most of the time and money organizations spend on testing is wasted because the tests are invalid, i.e., don't measure what they are supposed to measure.

OR

Why done so poorly?

Staff:

- Not trained in how to create good tests
- Don't want to be trained in testing technology
- Think testing is difficult to understand, involves math
- Think testing is time consuming and therefore expensive, i.e., not doable

Conclusion

- Organizations that think sophisticated testing is too expensive should at least consider that resources they currently spend on testing may very well be wasted.
- In most instances testing correctly is no more expensive than testing incorrectly - increasing the knowledge of the test designer is the only required investment.

Five Stages in the Creation of a Test

Analysis - Determine what to test

Advice #1: Create a test that looks like the job, not like other tests you have taken or seen.

- Use more performance tests:
  - more valid than paper and pencil tests
  - require creation of a rating scale for scoring, but involves no statistics
- Test at the top of the skills hierarchy:
  - provides greater validity
  - more cost effective than testing at lower levels

Construction - Write test items

Advice #2: Choose the most powerful item format, i.e., multiple choice. (Exception is paper and pencil items that simulate a job, e.g., form completion.)

Advice #3: Test beyond recall/memorization level. Write questions that require test-takers to work with previously unseen examples of concepts or applications of principles, not with definitions or formal statements of principles or facts.

Reliability - Determine if test yields consistent results

This process requires statistics. Concern yourself with validity not reliability.

Advice #4: For paper and pencil tests, skip reliability calculation.

Advice #5: For performance tests, calculate a simple percentage of agreement between raters.

Validity - Determine if test measures what it purports to measure

Advice #6: Establish content validity, rather than concurrent validity.

- Concurrent validity requires statistics.
- Content validity is the most important validity hurdle in lawsuits.
- Establishing content validity a matter of documenting that test measures job related tasks.
- Typically done via expert review of test.

Standard Setting - Process of deciding what cut-off scores for mastery should be

Advice #7 : Use Angoff-Nedelsky method of establishing cut-off score. Does not require large amounts of test-taker data or statistics.