

Assessment Centers: Not as (in)Valid as you Think

Kevin R. Murphy

Professor of Psychology and
Information Sciences and Technology

Prototypic AC

- **Purpose:** Promotion into supervisory/management position
- **Dimensions:** Problem Analysis, Decision Making, Leadership, Planning and Organizing
- **Exercises:** In basket, Presentation, Written Case Study, One-on-One Interaction ("Role Play"), Leaderless Group Discussion
- **Assessors:** Middle level managers, HR managers, consultants. Integration discussion to consensus on dimension ratings, then on overall assessment rating.

The Problem

- ACs have many appealing features
 - Careful design
 - Multiple methods designed to elicit relevant behaviors
 - Multiple judges, multiple perspectives
 - Careful integration
- Given all this, assessments of AC validity have not been as positive as they might be

Assessment Center Validity

- Long debate on whether the AC is valid and why
- Numerous studies looking at convergent and discriminant validity, exercise vs. dimension effects, criterion-related validity
 - On the whole, these are as studies are slightly discouraging, especially given the positive features of ACs

Discriminant Validity

- Most ACs use multiple exercises to evaluate multiple dimensions of behavior
- Each dimension is usually measured in more than one exercises
- The usual finding is that there are strong exercise factors and relatively weak dimension factors

Discriminant Validity

- That is, people who do well on exercise 1 but poorly on exercise 2 will tend to be rated highly on all dimensions in exercise 1 and lower on all dimensions in exercise 2
- Like all other humans, assessors seem to show evidence of halo error in their judgments

Why Discriminant Validity is Important

- The rationale of the AC is based in part on the notion that it is possible to assess the same dimensions in several separate exercises
- Feedback about dimensions is likely to be important for many AC applications (e.g., assessing developmental needs)

It's Not Just Discriminant Validity that is Disappointing

- Concerns about the criterion-related validity of Overall Assessment Ratings (OAR) as predictors of performance
 - Several meta-analyses (most recently Hemerlin, Lievens & Robertson, 2007, *Intl. Journal of Selection and Assessment*) suggest that OARs are correlated in the .25-.35 range with performance ratings, depending on the corrections used
 - This is not so impressive in comparison with much simpler and less expensive ability tests (validities into the .50s, depending on the corrections used)

AC Validity: Valid for What ?

- Two broad purposes of tests and assessments that are used in organizations
 - Measurement Instruments
 - Do AC ratings measure what they purport to measure
 - e.g., does a high score on Decision Making reflect real individual differences in Decision Making, or does it reflect something else?
 - Input/Support for decisions and actions
 - Do Acs help organizations and individuals make good decisions?
 - It is useful to look carefully at the purposes and uses of ACs

Valid for What ?

- ACs are used to
 - Direct developmental activities
 - Help make decisions about advancement
 - Help make decisions about placement
- The purpose of AC is typically not to
 - Serve as a replacement for written selection tests
 - Predict future performance evaluations

What is the Ultimate Criterion Here?

- Good decisions
 - Development
 - Advancement
 - placement
- Useful information
 - Correct
 - Credible
 - Accepted

Validity of ACs

- **Measurement Validity**
 - Current methods are reasonable
 - If they were thought of solely as measurement tools, ACs might not be worth the time and expense
- **Validity as a decision support**
 - Stakeholders' strategies for using AC information is part of the definition of validity

Are Acs Valid for Their Intended Uses?

- **Development**
 - Better choice of developmental activities, strategies than would have been made without AC
- **Advancement**
 - Do they help you find the right person for the job
- **Placement**
 - Do they help you find the right job for the person?

Validation Strategies

- **Content**
 - Better than nothing, but not by much
- **Construct**
 - Convergent/discriminant is one piece of a much larger puzzle
- **Criterion-related**
 - Choice of criteria should be tied to uses and purposes of tests
 - Take the fullest advantage possible of existing validation studies

Validation Strategies

- Three groups you need to convince about validity
 - Users
 - Regulators
 - Researchers
- Different strategies likely to appeal to and convince each