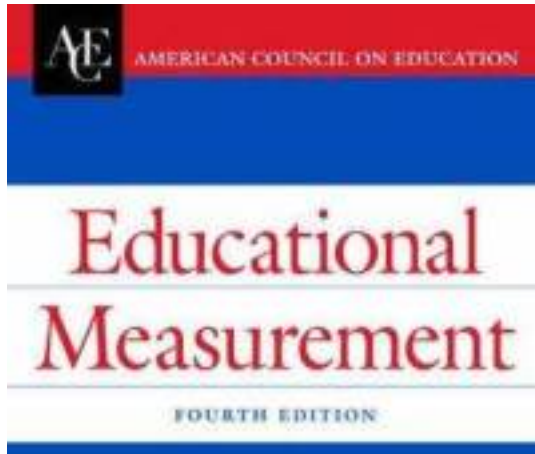# Workplace-based Assessment

Expanding the Validity Argument

Marjan Govaerts
Cees van der Vleuten

Validity inquiry in work-based assessment

**AMERICAN COUNCIL ON EDUCATION**

Educational Measurement

FOURTH EDITION

EDITED BY
ROBERT L. BRENNAN

Sponsored Jointly by
National Council on Measurement in Education
American Council on Education

ACE/PRAEGER SERIES ON HIGHER EDUCATION

3. RELIABILITY AND VALIDITY OF A NEW MULTISOURCE FEEDBACK EVALUATION TOOL (PEDS360) FOR RESIDENTS
Su-Ting T. Li, MD, MPH, University of California Davis, Sacramento, CA,
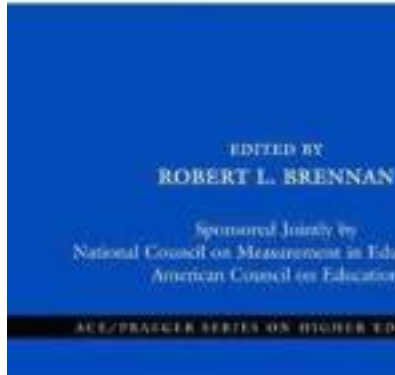
Academic Emergency Medicine
Official Journal of the Society for Academic Emergency Medicine

RESIDENCY ADMINISTRATION

Synchronous Collection of Multisource Feedback Evaluations Does Not Increase Inter-rater Reliability

Gregory Garra, DO, and Henry Thode, PhD

Abstract

Objectives: Most multisource feedback (MSF) evaluations are performed asynchronously, with raters reflecting on the subject's behavior. Numerous studies have demonstrated poor inter-rater reliability of MSF. This may be due to cognitive biases that are inherent in such a process. We sought to determine if

Theoretical assumptions

Generalizable explanations and predictions; true score; context-free criteria (e.g. accuracy, consistency)

Theories of learning, competence, performance interpretation

Assessment design and evaluation (validation)

Theoretical assumptions

▸ Learning (professional development)

  ▸ Deterministic, linear process

  ▸ Common trajectories; to be specified in advance

▸ Competence

  ▸ Fixed, permanent and latent trait / ability

  ▸ Transferable from one context to another

▸ Performance interpretation

  ▸ Objectifiable: single true score

  ▸ Assessors: passive interchangeable measurement instruments

▸ LEARNING
  ▸ Inherently situated, non-linear
    ▸ Emerges through interaction and participation
    ▸ Shaped by unique experiences in a specific context, and the meaning attached to these experiences
  ▸ Transformative, expansive, creative
    ▸ Ability to create *new* knowledge
    ▸ Unpredictable, unanticipated
  ▸ Collaborative
    ▸ No longer located within individual
    ▸ Competence is collective

Challenging assumptions - 1

Predictable and decidable systems of workplace learning?

Predictability and uniformity in learning trajectories?

Competence: individual versus team?

Hager, 2011; Engestrom & Sannino, 2010; Lingard, 2012

► COMPETENCE

- ► Changes over time
- ► Is not portable, transferable from one context to another

Talent won't transfer unless it maps to the new environment

- ► Cannot be defined independently of its context

Challenging assumptions - 2

Performance variability is measurement error?

Competence is trait?

Groysberg, McLean & Nohria, 2006; Groysberg et al., 2008; Wenghofer et al., 2009; Ginsberg et al., 2012; Lingard, 2012

- PERFORMANCE INTERPRETATION
  - Influenced by social structures in assessment task and its context
  - According to the perspectives and values of an individual assessor
  - Influenced by his/her experiences

- Assessors: active information processors, social perceivers

- Multiple true scores

**Challenging assumptions - 3**

True level of performance?

Absolute, objective truth?

Single true score?

Kogan et al., 2011; Govaerts et al., 2013; Lance et al., 2008; Delandshere & Petrosky, 1994

Implications for work-based assessment

▸ Socially situated interpretive act

▸ Assessment as *qualitative inquiry*

  ▸ Addressing outcome as well as process

  ▸ Understanding what, how, when and why

  ▸ Rich *situated* accounts of learning, performance and assessor judgments
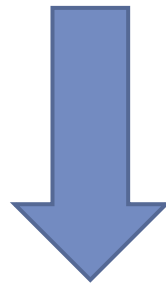
▸ CONSTRUCTIVIST-INTERPRETIVIST ASSESSMENT APPROACHES

Gipps, 1999; Moss et al., 2006; Johnston, 2004

Implications for assessment design

▸ <u>Dialogue</u>; learner engagement

  ▸ <u>Documentation</u> of learning context and process

▸ <u>Elaborate written statements</u> about <u>performance in context</u>, by expert-judges

▸ <u>Broad sampling</u> to gain in-depth understanding of <u>performance repertoire</u>

▸ All stakeholders <u>articulate and document performance interpretations</u>, underlying values and assumptions

▸ Open deliberative and <u>critical dialogue</u> until coherent interpretation can be accounted for

Implications for validity theory/inquiry

Quality principle	Qualitative approach	Techniques / strategies
Truth value of evidence Internal validity	Credibility	Prolonged engagement; multiple sources and methods; triangulation; extended sampling over time/contexts; member checking (trainee engagement) peer examination (critical dialogue); reflexivity
Applicability of evidence External validity	Transferability	Thick, rich (dense) description; explain sampling strategy; time sampling
Consistency of evidence Reliability	Dependability	Stepwise replication; iterative data collection; saturation; thick description of methods and decision trail; triangulation; peer examination; explainable variability
Neutrality of evidence Objectivity	Confirmability	Audit trail; external audit; seek for disconfirming evidence; critical dialogue; reflexivity (articulate role, values,..)

Van der Vleuten et al., 2010; Krefting, 1991

Thank you!

marjan.govaerts@maastrichtuniversity.nl

Marjan Govaerts   Maastricht University