

Instructor’s Corner

Improving Ethical Clinical Assessment Practice by Improving Knowledge of Measurement Principles

By: Gary L. Canivez, Eastern Illinois University

Numerous ethical principles of psychologists (APA, 2002, 2010 Amendments) apply to the assessment process and use of test scores in research and clinical practice. Examples include:

2.01 Boundaries of Competence: (a) Psychologists provide services, teach, and conduct research with populations and in areas only within the boundaries of their competence, based on their education, training, supervised experience, consultation, study, or professional experience;

2.04 Bases for Scientific and Professional Judgments: Psychologists’ work is based upon established scientific and professional knowledge of the discipline. (See also Standards 2.01e, Boundaries of Competence, and 10.01b, Informed Consent to Therapy.);

9.01 Bases for Assessments: (a) Psychologists base the opinions contained in their recommendations, reports, and diagnostic or evaluative statements, including forensic testimony, on information and techniques sufficient to substantiate their findings;

9.02 Use of Assessments: (a) Psychologists administer, adapt, score, interpret, or use assessment techniques, interviews, tests, or instruments in a manner and for purposes that are appropriate in light of the research on or evidence of the usefulness and proper application of the techniques, (b) Psychologists use assessment instruments whose validity and reliability have been established for use with members of the population tested. When such validity or reliability has not been established, psychologists describe the strengths and limitations of test results and interpretation; and

9.08 Obsolete Tests and Outdated Test Results: (a) Psychologists do not base their assessment or intervention decisions or recommendations on data or test results that are outdated for the current purpose, (b) Psychologists do not base such decisions or recommendations on tests and measures that are obsolete and not useful for the current purpose, (c) Psychologists retain responsibility for the appropriate application, interpretation, and use of assessment instruments, whether they score and interpret such tests themselves or use automated or other services.

In order to employ psychological tests and procedures in an ethical

manner, it is necessary that psychologists understand *how* tests and test scores are assessed to provide evidence regarding various types of reliability, validity, diagnostic utility, and norms. These measurement principles are critical so that psychologists know which scores have acceptable psychometric evidence supporting the variety of possible inferences proffered by test authors, publishers, and trainers. Test scores that lack sufficient evidence of reliability, validity, and utility lead clinicians to make inappropriate and inaccurate inferences about the individual assessed.

Over two decades ago Weiner asserted: Effective clinicians, “(a) know what their tests can do and (b) act accordingly. Knowing what one’s test can do – that is, what psychological functions they describe accurately, what diagnostic conclusions can be inferred from them with what degree of certainty, and what kinds of behavior they can be expected to predict – is the measure of a psychodiagnostician’s competence. Acting accordingly – that is, expressing only opinions that are consonant with the current status of validity data – is the measure of his or her ethicality” (1989, p. 829).

In order to follow Weiner’s sage advice, psychologists must possess fundamental competencies in psychological measurement including test score reliability, validity, utility, and norms. The importance of these competencies cannot be overstated for ethical assessment and clinical practice (Dawes, 2005; McFall, 1991, 2000). However, in commenting on the results of the follow-up to their 1990 survey (Aiken, West, Sechrest, & Reno, 1990), Aiken, West, and Millsap (2008) found there was little change to doctoral training regarding measurement, concluding, “we find it deplorable that a dozen years later, the measurement requirement occupies a median of only 4.5 weeks in the PhD curriculum in psychology” (p. 43). They further lamented the inadequacies of measurement training that resulted in a situation “that most graduates lacked fundamental competency in measurement” (p. 43).

It is within this context that, six years ago, I developed and began presenting a continuing education workshop for professional psychologists entitled *Measurement Matters*. Presentation of this workshop began with independent school districts and expanded to include state and national school psychology association conferences. Recently, this workshop was selected for presentation at the 2011 Annual Convention of the American Psychological Association. An abbreviated and more advanced version focusing on validity and diagnostic utility is scheduled for the 2012 Conference of the International Test Commission, Amsterdam, the Netherlands. While such a workshop cannot be a substitute for graduate level training in psychological measurement, it may help to address the significant shortcomings highlighted by Aiken et al. (2008) and aid professional psychologists in making better-informed decisions in test selection and interpretation.

This workshop begins with a presentation of the professional ethics that should guide professional practice in psychological assessment along with scientific principles and attitudes promoted by McFall (1991, 2000). Other guiding publications serving as foundations include: *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1999), *Adapting Educational and Psychological Tests for Cross-Cultural Assessment* (Hambleton, Merenda, & Spielberger, 2005), *International Test Commission Guidelines for Translating and Adapting Tests* (ITC, 2010), *Guidelines for Test User Qualifications* (Turner, DeMers, Fox, & Reed, 2001), and *Agreement Among Diagnosticians or Observers: Its Importance and Determination* (McDermott, 1988).

Below is an outline of the topics and methods for estimating reliability, validity, and diagnostic utility of test scores addressed in the workshop and each is described in terms of its particular importance in test selection and score interpretations and illustrated by at least one published research study to provide exemplars of application.

Measurement Matters: Applying Psychological Measurement Principles in Clinical Assessment

Outline

- Introduction
 - Ethical Principles, Standards, Test Use, and Measurement Principles
 - Scientific Thinking and Practice
 - Theories of Measurement
- Methods for Estimating Reliability
 - Sources of Measurement Error, Standard Error of Measurement, & Confidence Intervals
 - Internal Consistency (Item/Scale Homogeneity)
 - Test – Retest (Stability)
 - Interrater Agreement (Interobserver Agreement)
 - Alternate Forms (Equivalence)
- Methods for Estimating Validity
 - Content Validity or Evidence based on Test Content
 - Concurrent Validity or Relationships with other Variables
 - Predictive Validity or Relationships with other Variables
 - Age/Developmental Changes
 - Distinct Group Differences
 - Theory Consistent Intervention Effects
 - Convergent & Divergent/Discriminant Validity
 - Multitrait – Multimethod Matrix
 - Factorial/Structural Validity: Exploratory & Confirmatory
 - Incremental Validity
- Methods for Estimating Diagnostic Utility
 - Diagnostic Efficiency/Utility Statistics
 - Receiver Operator Characteristic (ROC) Analysis
 - Cluster Analysis
- Norms and their Importance

It is hoped that continued offering of this workshop at the international, national, state, and local levels will both aid and facilitate the improvement of professional psychologists’ knowledge of measurement principles so they can more competently read and critically consider information in test technical manuals and professional literature, and become more proficient in recognizing what information is presented, what might be missing, and what requires further examination to provide evidence for the many and varied clinical inferences proffered. In consideration of the

relationship between these measurement principles and ethical test use, I have amended the title to *Measurement Matters: Applying Psychological Measurement Principles for Ethical Clinical Assessment*. Ultimately, graduate training must answer the call of Aiken et al. (2008) by increasing and elevating measurement training to a standard that promotes better test use among professional psychologists. It is this competency in measurement that allows psychologists to follow Weiner’s (1989) advice to, “(a) know what their tests can do and (b) act accordingly” (p. 829).

References

Aiken, L. S., West, S. G., Sechrest, L. B., & Reno, R. R. (1990). Graduate training in statistics, methodology, and measurement in psychology: An international survey. *American Psychologist*, 45, 721–734.

Aiken, L. S., West, S. G., & Millsap, R. E. (2008). Doctoral training in statistics, measurement, and methodology in psychology: Replication and extension of Aiken, West, Sechrest, and Reno’s (1990) survey of PhD programs in North America. *American Psychologist*, 63, 32-50.

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME) (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

American Psychological Association. (2002, 2010 Amendments). *Ethical principles of psychologists and code of conduct*. Washington, DC: APA.

Dawes, R. M. (2005). The ethical implications of Paul Meehl’s work on comparing clinical versus actuarial prediction methods. *Journal of Clinical Psychology*, 61, 1245-1255.

International Test Commission (ITC) (2010). *International Test Commission Guidelines for translating and adapting tests*. Downloaded electronically on [10 July 2011] from www.intestcom.org/guidelines/index.php

Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (Eds.) (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Mahwah, NJ: Lawrence Erlbaum Associates.

McDermott, P. A. (1988). Agreement among diagnosticians or observers: Its importance and determination. *Professional School Psychology*, 3, 225-240.

McFall, R. M. (1991). Manifesto for a science of clinical psychology. *The Clinical Psychologist*, 44, 75–88.

McFall, R. M. (2000). Elaborate reflections on a simple manifesto. *Applied and Preventive Psychology*, 9, 5–21.

Turner, S. M., DeMers, S. T., Fox, H. R., & Reed, G. M. (2001). APA’s guidelines for test user qualifications: An executive summary. *American Psychologist*, 56, 1099-1113.

Weiner, I. B. (1989). On competence and ethicality in psychodiagnostic assessment. *Journal of Personality Assessment*, 53, 827–831.

For more information about Dr. Canivez or his workshops, please email him at glcanivez@eiu.edu or go to his website at <http://www.ux1.eiu.edu/~glcanivez>.

Coming to the *Instructor’s Corner* in January, Dr. Marsha Lovett writes about evaluating the effectiveness of teaching online statistics courses.

Look for Your APA
Apportionment Ballots
Coming This
November!
Remember to Vote!