**Fairness EPSY 5221**

From the *Testing Standards* (2014):

* “Fairness is a fundamental validity issue and requires attention throughout all stages of test development and use” (p. 49).
* Measurement bias is a central threat to fairness, but two additional concepts include accessibility (“unobstructed opportunity to demonstrate their standing on the construct(s) being measured,” p. 49) and universal design (approaches that maximize accessibility for all intended test takers, p. 50).
* Fairness is “responsiveness to individual characteristics and testing contexts so that individual test scores will yield valid interpretations for intended uses. …A test that is fair within the meaning of the *Standards* reflects the same construct(s) for all test takers, and scores from it have the same meaning for all individuals in the intended population; a fair test does not advantage or disadvantage some individuals because of characteristics irrelevant to the intended construct.” (p. 50).
* The *Standards* describe general views of fairness
	+ equitable treatment during the testing process
	+ lack of measurement bias (introducing a new range of definitions)
	+ access to the intended measured constructs (a test bias issue)
	+ validity of individual test score interpretation and use
* The core fairness standard (3.0) implies that attention to fairness is intended to “minimize construct-irrelevant variance and to promote valid score interpretations for the intended uses for all examinees in the intended population” (p. 63).
* “The central idea of fairness in testing is to identify and remove construct-irrelevant barriers to maximal performance for any examinee. Removing these barriers allows for the comparable and valid interpretation of test scores for all examinees. Fairness is thus central to the validity and comparability of the interpretation of test scores for intended uses” (p. 63).

**From others:**

Willingham & Cole (1997), fair tests measure the same construct(s) for all test takers and scores have the same meaning for all individuals in the intended population – a concern that covers all stages of test development and use.

And a series of chapters in the edited book, *Fairness in educational assessment and measurement* (Dorans & Cook, 2016):

Zieky (2016), fairness is a matter of validity, and a fair test is one that is valid for different groups of test takers. Construct-relevant sources of score variance allow valid and fair inferences about test takers. CIV leads to unfair inferences if the irrelevant variance is associated with group membership. He also suggests that different views of fairness may vary in usefulness to different audiences.

Wollack & Case (2016), fairness is supported through standardization and securing comparable administration conditions. Fairness relates to equal treatment of all examinees.

Penfield (2016), each scored element should hold the same meaning regarding the proficiency measured and multiple independent scores should be consistent for all test takers regardless of background characteristics.

Sinharay (2016), fairness is concerned with protecting test takers and users in all aspects of testing.

Worrell (2016), when it comes to test fairness, communications and public relations are important vehicles and contexts to promote fair test score interpretation and use.

**References**

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.

Calahan Laitusis, C., & Cook, L.L. (Eds.). (2007). *Large-scale assessment and accommodations: What works*? Alexandria, VA: Council on Exceptional Children.

Cogan Hauck, M., Wolf, M.K., & Mislevy, R. (2013). *Creating a next-generation system of K-12 English learner (EL) language proficiency assessments*. Princeton, NJ: Educational Testing Service.

Darling-Hammond, L., Herman, J., Pellegrino, J., et al. (2013). *Criteria for high-quality assessment*. Stanford, CA: Stanford Center for Opportunity Policy in Education.

Dorans, N.J., & Cook, L.L. (Eds.). (2016). *Fairness in educational assessment and measurement*. New York, NY: Routledge.

Ekstrom, R.B., & Smith, D.K. (Eds.). (2002). *Assessing individuals with disabilities in educational, employment, and counseling settings*. Washington, DC: American Psychological Association.

Elliott, S.N., Kettler, R.J., Beddow, P.A., & Kurz, A. (Eds.). (2018). *Handbook of accessible instruction and testing practices: Issues, innovations, and applications*. New York, NY: Springer.

Geisinger, K.F. (2002). Testing the members of an increasingly diverse society. In J.F. Carlson & B.B. Waterman (Eds.), *Social and personality assessment of school-aged children: Developing interventions for educational and clinical use*. Boston, MA: Allyn & Bacon.

Geisinger, K.F. (2003). Testing and assessment in cross-cultural psychology. In J. R. Graham & J. A. Naglieri (Eds.), *Handbook of psychology* (Volume 10: Assessment Psychology) (95-117). New York, NY: John Wiley.

Geisinger, K.F. (2004). Testing students with limited English proficiency. In J. E. Wall & G. R. Walz (Eds.), *Measuring up: Assessment issues for teachers, counselors, and administrators* (147-159). Greensboro, NC: ERIC Clearinghouse for Counseling and Student Services and the National Board for Certified Counselors.

Geisinger, K.F. (2005). The testing industry, ethnic minorities, and those with disabilities. In R. Phelps (Ed.), *Defending standardized testing* (pp. 187-203). Mahwah, NJ: Erlbaum.

Geisinger, K.F. (Ed.). (in preparation). *Psychological testing of Hispanics: Clinical and intellectual issues*. Washington, DC: American Psychological Association.

Guzmán-Orth, D., Lopez, A.A., & Tolentino, F. (2017). *A framework for the dual language assessment of young dual language learners in the United States*. Princeton, NJ: Educational Testing Service.

Ketterlin-Geller, L.R., Johnstone, C.J., & Thurlow, M.L. (2015). Universal design in assessment. In Burgstahler, S.E. (Ed.), *Universal design in higher education: From principles to practice* (2nd ed., pp. 163-175). Cambridge, MA: Harvard Education Press.

Kieffer, M.J., & Thompson, K.D. (2018). Hidden progress of multilingual students on NAEP. *Educational Researcher*, *47*(6), 391-398.

Kopriva, R., Thurlow, M. L., Perie, M., Lazarus, S. S., & Clark, A. (2016). Test takers and the validity of score interpretations. *Educational Psychologist, 5*(1), 108-128*.*

Liu, K.K., Thurlow, M.L., Press, A.M., & Lickteig, O. (2018). *A review of the literature on measuring English language proficiency progress of English learners with disabilities and English learners* (NCEO Report 408). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Liu, K.K., Ward, J.M., Thurlow, M.L., & Christensen, L.L. (2017). Large-scale assessment and English language learners with disabilities. *Educational Policy, 31*(5), 551-583*.*

McConnell, S.R., Wackerle-Hollman, A.K., Roloff, T.A., & Rodriguez, M.C. (2015). Designing a measurement framework for Response to Intervention in early childhood programs. *Journal of Early Intervention, 36*(4), 263-280.

Quenemoen, R. F., & Thurlow, M. L. (2017). Standards-based reform and students with disabilities. In J.M. Kauffman & D.P Hallahan (Eds), *Handbook of special education* (pp. 203-217)*.* New York: Routledge.

Restorff, D., Sharpe, M., Abery, B., Rodriguez, M.C., Kim, N.K. (2012). Teacher perceptions of alternate assessments based on alternate achievement standards: Results from a three-state survey. *Research & Practice for Persons with Severe Disabilities, 37*(3), 185-198.

Rodriguez, M.C. (2009). Psychometric considerations for alternate assessments based on modified academic achievement standards. *Peabody Journal of Education, 84*(4), 595-602.

Ruiz-Primo, M.A., Li, M., Wang, T., Giamellaro, M., Wills, K., Orgeron, M., & Zhao, D.Y. (2013). *Comparing item formats of instructionally sensitive assessments*. DEISA Paper 1. Denver, CO: School of Education and Human Development, University of Colorado. Retrieved at http://source.ucdsehd.net/deisa/

Ruiz-Primo, M.A., Li, M., Wills, K., Giamellaro, M., Lan, M-C., Mason, H., Sands, D. (2012). Developing and evaluating instructionally sensitive assessments in science. *Journal of Research on Science Teaching, 49*(6), 691-712.

Sandoval, J., Frisby, C., Geisinger, K.F., Scheuneman, J. & Ramos-Grenier, J.M. (Eds.). (1998). *Test interpretation and diversity: Achieving equity in psychological assessment*. Washington, DC: American Psychological Association.

Scheerens, J. (Ed.) (2017). *Opportunity to learn, curriculum alignment and test preparation: A research review*. New York, NY: Springer.

Shepard, L.A., Penuel, W.R., & Pellegrino, J.W. (2018). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educational Measurement: Issues and Practice*, *37*(1), 21-34.

Shepard, L.A., Penuel, W.R., & Pellegrino, J.W. (2018). Classroom assessment principles to support learning and avoid the harms of testing. *Educational Measurement: Issues and Practice*, *37*(1), 52-57.

Stone, E., Calahan Laitusis, C., & Cook, L. (2015). Increasing the accessibility of assessments through technology. In F. Drasgow (Ed.), *Technology and testing: Improving educational and psychological measurement*. New York, NY: Routledge.

Thurlow, M.L., & Quenemoen, R.F. (2016). Alternate assessments for students with disabilities: Lessons learned from the National Center and State Collaborative. In C. Wells & M. Faulkner-Bond (Eds.), *Educational measurement: From foundations to future* (pp. 416-432)*.* New York: Guilford.

Thurlow, M.L., & Kopriva, R.J. (2015). Advancing accessibility and accommodations in content assessments for students with disabilities and English learners. *Review of Research in Education, 39,* 331-369

Way, W.D. (2014). *Memorandum on instructional sensitivity considerations for the PARCC assessments*. Partnership for Assessment of Readiness for College and Careers. Retrieved at www.parcconline.org/files/65/.../Instructional\_sensitivity\_memo\_final08\_15\_14.pdf

Willingham, W.W., & Cole, N.S. (Eds.). (1997). *Gender and fair assessment*. New York, NY: Routledge.