

Fairness and Equity in Assessment Task Force

Initial Meeting – January 22, 2021
9am-10am - Virtual meeting - Zoom



Agenda

Greetings and Introductions

Website: <https://assessment.aa.ufl.edu/>

The Charge to the Task Force

The Groups: Guideline Development and Model Development

Getting started: Presentations from the Chairs

Next Steps

The Charge

The Task Force is charged with establishing a set of guidelines for UF faculty, instructors, staff, and administrators to help ensure fairness and equity in assessment in all contexts at the university. The guidelines must address fairness and equity in the entire assessment process, including development, opportunity to learn, administration, scoring, score interpretations, and the evaluation of the measurement properties of the assessment.

Task Force Responsibilities

Attend scheduled Task Force meetings.

Contribute expertise to the proceedings and the development of the guidelines and models.

Communicate Task Force progress to your unit when appropriate.

Definition

Fairness and equity in assessment refer to the process of measurement (e.g., through surveys, batteries, scales, rubrics, tests) such that the interpretations and uses of scores are based on the construct, indicator, or learning outcomes being measured and not the characteristics of the examinee (e.g., race, ethnicity, gender, or disability). Fairness and equity in assessment ensure that no examinees are disadvantaged based on these characteristics.

The Groups

Guideline Development Group

Work with David Miller and Teresa Mutahi to develop the guidelines and review the recommendations from the model development group, and where mutually agreeable, implement the recommendations to modify the guidelines.

Model Development Group

Work with Corinne Huggins-Manley to locate and develop models that operationalize the guidelines in various contexts; provide the guideline development group with modifications to increase their utility.

David Miller, Co-chair

Professor, Research Evaluation and
Methodology
College of Education





Many Resources Available

Professional Standards

Books

Articles


Reports

Experiences

Professional Standards

Standards for Educational and Psychological Testing (2014). American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME).

Code of Fair Testing Practices in Education (2004). Joint Committee on Testing Practices (American Counseling Association, AERA, APA, American Speech-Language-Hearing Association, National Association of School Psychologists, National Association of Test Directors, and NCME).

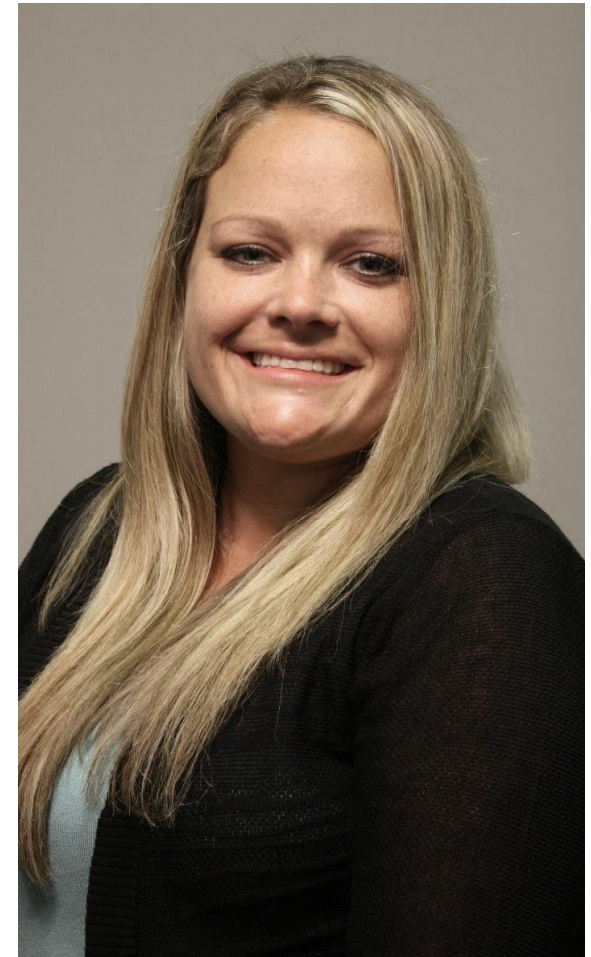



Critical Areas of the Code of Fair Testing Practices in Education

- A. Developing and Selecting Appropriate Tests
- B. Administering and Scoring Tests
- C. Reporting and Interpreting Test Results
- D. Informing Test Takers

Corinne Huggins- Manley, Co-chair

Associate Professor, Research and
Evaluation Methodology
College of Education





Measurement Validity, Fairness, and Equity (In a Nutshell)

Validity as a Unified Concept

- Definition: The degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests.

(Standards for Educational and Psychological Testing, p. 9)

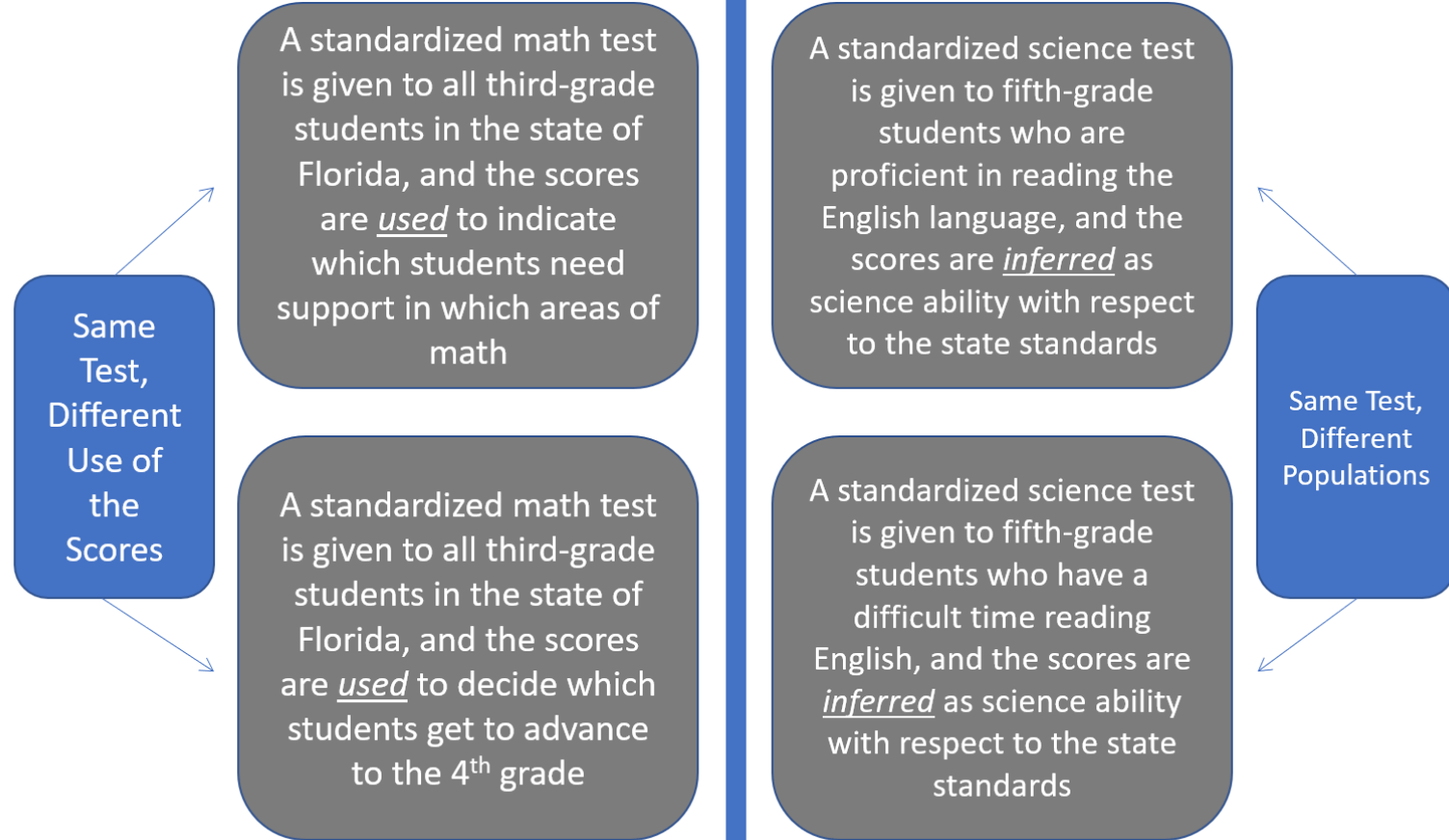
Validity is a
matter of
degree, not all
or none

The test or
instrument is not
to be validated;
the
interpretations
(and uses) are to
be validated

Evidence of validity
comes from multiple
sources, but that
does not mean that
there are multiple
types of validity

The central issue is
appraisal of the
meaning and
consequences of
measurement

It is test score interpretations and uses that can have or lack some degree of validity

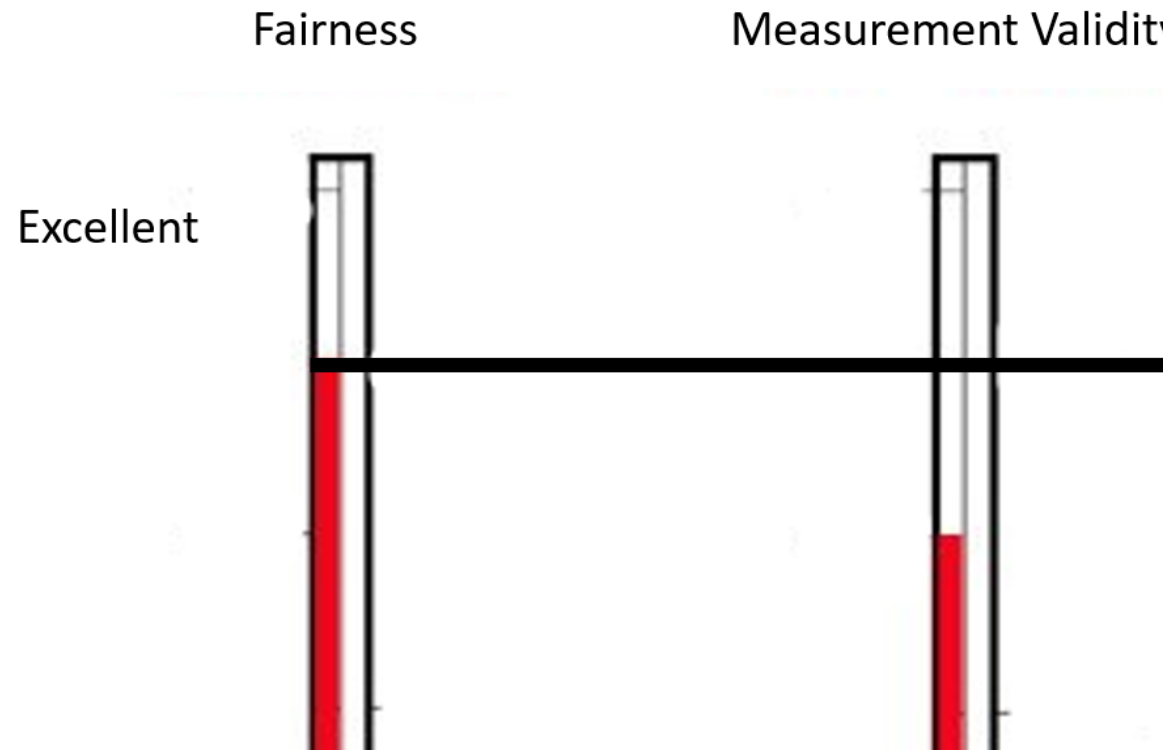


Fairness and Equity in Assessment

- Definition: The process of measurement (e.g., through surveys, batteries, scales, rubrics, tests) such that the interpretations and uses of scores are based on the construct, indicator, or learning outcomes being measured and not the characteristics of the examinee (e.g., race, ethnicity, gender, or disability). Fairness and equity in assessment ensure that no examinees are disadvantaged based on these characteristics.

(UF Fairness and Equity in Assessment Task Force, 2021)

Fairness as a Prerequisite to Validity



There are Many Examples of Evaluating Measurement Fairness

Test-Based Grade Retention: Does It Stand Up to Professional Standards for Fair and Appropriate Test Use?

Randall D. Penfield

First Published March 1, 2010 | Research Article

<https://doi.org/10.3102/0013189X10363007>

[Article information](#) ▼

Educational Researcher

Abstract

A growing body of research showing that grade retention serves as an educationally low-quality placement has raised increasing concerns about whether the use of standardized tests in making decisions concerning grade retention conforms to current standards for appropriate and nondiscriminatory test use. This article examines the extent to which test-based grade retention policies comply with standards for fair and appropriate test use based on norms established by the professional testing community. The results of the investigation indicate that test-based retention policies potentially violate several of the professional standards. The points of violation are presented, and the implications for test-based retention policies are discussed.

There are Many Examples of Evaluating Measurement Fairness

Assessment of Differential Item Functioning for Performance Tasks

Rebecca Zwick, John R. Donoghue, Angela Grima

First published: September 1993 | <https://doi.org/10.1111/j.1745-3984.1993.tb00425.x> | Citations: 119

Abstract

Although the belief has been expressed that performance assessments are intrinsically more fair than multiple-choice measures, some forms of performance assessment may in fact be more likely than conventional tests to tap construct-irrelevant factors. The assessment of differential item functioning (DIF) can be helpful in investigating the effect on subpopulations of the introduction of performance tasks. In this study, two extensions of the Mantel-Haenszel (MH; 1959) procedure that may be useful in assessing DIP in performance measures were explored. The test of conditional association proposed by Mantel (1963) seems promising as a test of DIF for polytomous items when the primary interest is in the between-group difference in item means, conditional on some measure of ability. The generalized statistic proposed by Mantel and Haenszel may be more useful than Mantel's test when the entire response distributions of the groups are of interest. Simulation results showed that, for both inferential procedures, the studied item should be included in the matching variable, as in the dichotomous case. Descriptive statistics that index the magnitude of DIP were also investigated.



ERIC

[Collection](#)[Thesaurus](#)[Search](#)[Advanced Search Tips](#)☐ Peer reviewed only ☐ Full text available on ERIC

Accommodations and Universal Design: Supporting Access to Assessments in Education

Ketterlin-Geller, Leanne R.; Johnstone, Christopher

Journal of Postsecondary Education and Disability, v19 n2 p163-172 Fall 2006

Because of recent legislative mandates, students with disabilities have unprecedented opportunities to attend education. Access to instruction and assessment is provided through the use of reasonable accommodations. accommodations are legally and procedurally complex. This article addresses the legal and procedural evidence to receive testing accommodations. In addition, we discuss procedures for supporting student needs by applying universal design to assessments. By changing assessment practices to include support structures for all students, education can be promoted.

There are Many Examples of Evaluating
Measurement Fairness

Teresa Mutahi, Co-chair

Senior Lecturer & Biology Major
Undergraduate Coordinator



- Where are we?
 - What guidelines/policies exist?
 - Are there fairness and equity issues in assessments that we should address?
 - Are there strengths indicative of fairness and equity in assessments that could be replicated across UF?
-
- Developing Assessment tools
 - Administering Assessments
 - Informing examinees
 - Reporting and Interpreting Scores
 - Use and Evaluation of Assessments

Status of Fairness and Equity in assessments at UF

Examination of UF's Status on Fairness & Equity in Assessments

- *Guidelines Development Group*
- *Model Development Group*
- Faculty/College/Unit representatives
- Student Representatives
- Administrators/UFHR
- Library
- Graduate School
- Division of Student Affairs Assessment and Research

Scrutinize UF's Fairness & Equity Gaps/ Issues/Strengths

- Assessment tools
- Systems
- Are our assessment systems inclusive, fair and equitable for all faculty, staff and student populations?
- Processes
- Methods
- “Methods that reflect inherent biases are also much less likely to be detected by assessment teams unless they include diverse faculty and staff” (Huston C., 2020)

Scrutinize UF's Fairness and Equity Gaps/ Issues/Strengths

- Guidelines
- Policies- individuals, programs & institution
- Practices
 - Are our current assessment practices open and inclusive of faculty, staff or student populations? (Huston C., 2020).
 - Non-science faculty added the practice of student-centered assessment (peer assessment, competency-based grading, opportunity to submit multiple drafts) compared to exams while science faculty did not (Yanowitz & Hahs-Vaughn (2007).



Status of Fairness and Equity in Assessments at UF

- Practices
- Experiences
- Literature
- Case Studies
- Models
- Identify Gaps/issues and strengths

Fairness & Equity Gaps/Issues & Strengths Survey

- https://ufl.qualtrics.com/jfe/form/SV_gNBYvAYioRqqDSC

References

- Huston, C. (2020). Strategies for Change: Equity in Assessment Practices. Urbana, IL: National Institute for Learning Outcomes Assessment, Council for the Advancement of Standards in Higher Education, and Campus Labs. Retrieved January 20th 2020 from <http://www.learningoutcomesassessment.org/>
- Karen L. Yanowitz & Debbie L. Hahs-Vaughn (2007). Changes in student-centered assessment by postsecondary science and non-science faculty, Teaching in Higher Education, 12:2, 171-184, DOI: 10.1080/13562510701191927

Next Steps

- You will receive the survey via email, please complete it and submit it as soon as possible
- You will receive a doodle poll soon requesting your availability for bi-weekly meetings by group – we will ask you to identify times you are open
- The chairs will contact you about the first meeting time and date as soon as all responses have been received and the day/time is set



Thank You!