Day 2 Welcome

Session 6
Designing and Validating Career Readiness Measures



Maggie Reeves
Directior, K-12 Education
Urban Institute



Chris Domaleski
Executive Director
Center for Assessment



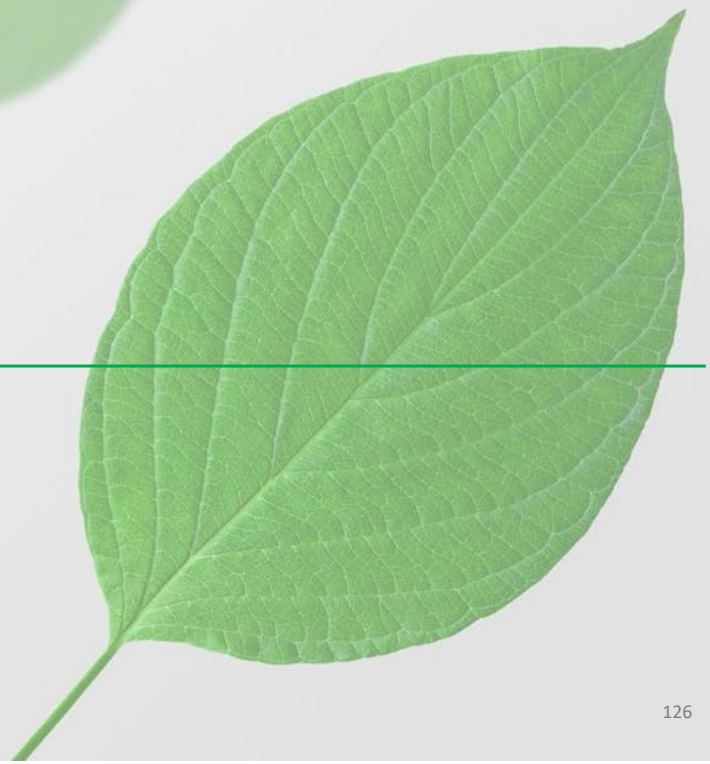
Erika Landl
Senior Associate
Center for Assessment



Laura Hamilton
Senior Associate
Center for Assessment

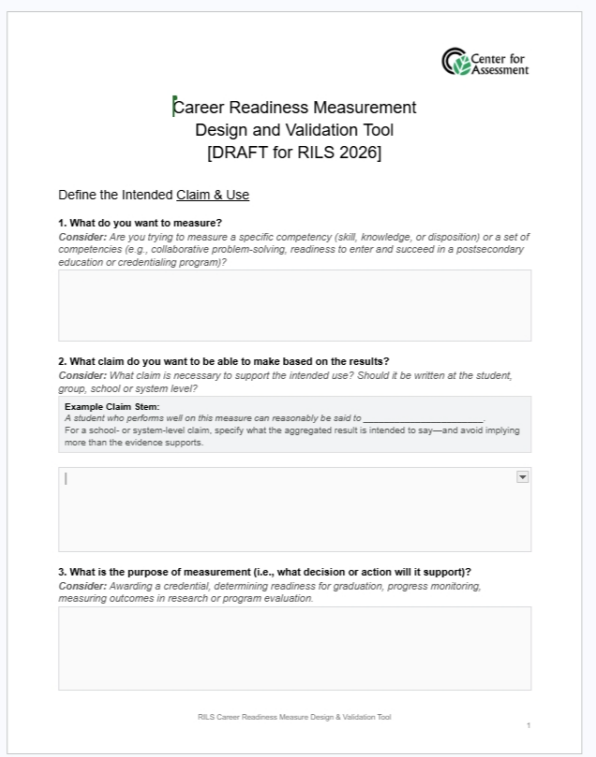
cc

Designing Career Readiness Measures



Career Readiness Measurement Design and Validation Tool

- Developed this tool to help think through the components of a comprehensive validity argument for a career readiness measure.
- Based on thinking done for upcoming chapter in ***Modeling What Matters: The Research and Legacy of Robert J. Mislevy***
- This is a working draft, so we welcome your comments and suggestions.



Career Readiness Measurement
Design and Validation Tool
[DRAFT for RILS 2026]

Define the Intended Claim & Use

1. What do you want to measure?
Consider: *Are you trying to measure a specific competency (skill, knowledge, or disposition) or a set of competencies (e.g., collaborative problem-solving, readiness to enter and succeed in a postsecondary education or credentialing program)?*

2. What claim do you want to be able to make based on the results?
Consider: *What claim is necessary to support the intended use? Should it be written at the student, group, school or system level?*

Example Claim Stem:
A student who performs well on this measure can reasonably be said to _____.
For a school- or system-level claim, specify what the aggregated result is intended to say—and avoid implying more than the evidence supports.

3. What is the purpose of measurement (i.e., what decision or action will it support)?
Consider: *Awarding a credential, determining readiness for graduation, progress monitoring, measuring outcomes in research or program evaluation.*

RILS Career Readiness Measure Design & Validation Tool
1

Career Readiness Measurement Design and Validation Tool

Center for Assessment

- Part 1: Define the Intended Claim & Use
- Part 2: Establish Context & Describe the Evidence
- Part 3: Operationalize the Measure
- Part 4: Evaluate the Measurement Approach

cc

Center for Assessment

A Word on Claims

A clear claim identifies:

- Who or what the statement is about: a student, group of students, school, program or system
- The result or condition that triggers the claim: a score a classification, completed experience, observed performance
- What, specifically, the result allows you to conclude about the knowledge, skill, disposition or construct of interest.

Too vague to guide evidence	Clear enough to guide evidence
The student is an effective communicator.	*A student who meets the standard can organize and present evidence-based ideas clearly for a specified audience and purpose*

The Claim Establishes the Evidentiary Bar

Center for Assessment

Different claims require different kinds of evidence. Evidence can support a narrower claim than originally intended, but **it cannot be stretched to support a stronger claim merely because it is available**.

If the claim says…	Evidence must show…	What it does NOT establish
Student had access to…	The opportunity was available to the students.	The students participated, learned or became competent.
Students participated in or completed…	Enrollment, attendance, completion, or another verified participation record.	The students acquired the intended knowledge or skill.
Students know or can demonstrate…	A direct demonstration or sufficiently strong evidence of the specified knowledge or skill.	Students participated in or completed a particular course or activity OR that the school had a role in supporting attainment of the competency.
A school is effective in supporting…	Results from a study that supports causal conclusions (e.g., school efforts/initiatives were effective in improving performance).	The school is equally effective (or ineffective) for all students or groups.

CR Measurement Design Activity

- At your table select one of the use cases provided on the next slide.
- Go to page 3 of the CR Measurement Tool
- In your tables talk through how you would respond to questions 7-13 based on this use case and the associated claim.
- If you have comments about how the tool may be improved as you engage in this task, please add them to this document for our consideration.

https://tinyurl.com/yny96unj



CC BY

Purpose/User	Claim
1. Evaluate the extent to which students are leaving high school with competencies that prepare them for college, careers, and other postsecondary opportunities. State, district, and school leaders use overall and disaggregated results to identify differences in outcomes across schools and student groups and target resources and support.	Schools with high scores have high proportions of students who can succeed in an entry-level job, postsecondary institution, or military service upon graduation.
2. Develop a measure of responsible decision-making aligned to the district's Portrait of a Graduate. Educators use the results to evaluate students' progress toward demonstrating this competency and identify opportunities for additional instruction and support. Students who demonstrate the competency earn a badge that can be included in their Portrait of a Graduate portfolio.	A student who earns the Responsible Decision Maker badge demonstrates the ability to gather relevant information, evaluate options from multiple perspectives, and consider potential consequences before making a decision.
3. Evaluate the extent to which a district's career-planning program helps students explore career options and make informed choices about their next steps. District and school leaders use the results to identify strengths and gaps in the program and inform improvements to curriculum, instruction, and supplemental activities.	Students who complete the program can identify careers aligned with their interests and goals, evaluate which courses and extracurricular opportunities can help them explore and prepare for different careers, and develop a plan for taking their next steps.

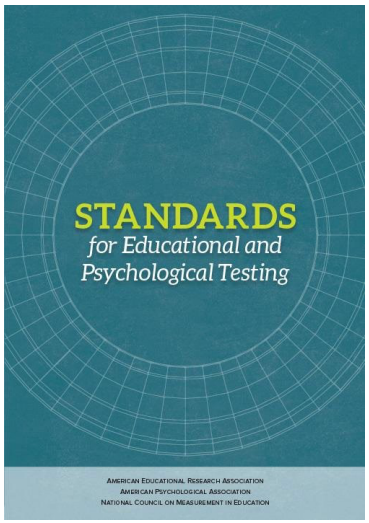


Validity 101

Validity

"*The degree to which evidence and theory support the interpretations of test scores for proposed uses of tests*" (Standards, p.11)

- This unified conceptualization emphasizes ongoing evidence collection
- Validity is inherently tied to interpretations and uses
- References to types of validity, or to measures as "valid," are inconsistent with modern validity theory



https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf

Center for Assessment

Validity: What's the Evidence?

Several sources of evidence can contribute to a validity argument:

- Test content
- Response processes
- Internal structure
- Relations to other variables
- Consequences of testing



Other Foundational Standards

Reliability:

"*The degree to which test scores for a group of test takers are consistent over repeated applications of a measurement procedure and hence are inferred to be dependable and consistent for an individual test taker*"

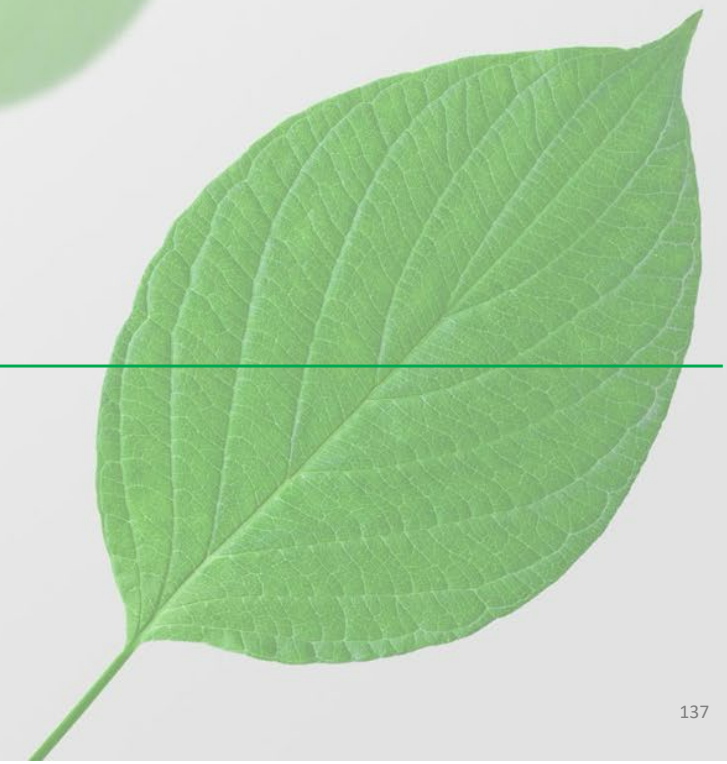
(*Standards*, pp.222-223)

Fairness:

"*The validity of test score interpretations for intended use(s) for individuals from all relevant subgroups… [minimize effects of] individual characteristics and testing contexts that otherwise would compromise the validity of scores for some individuals*."

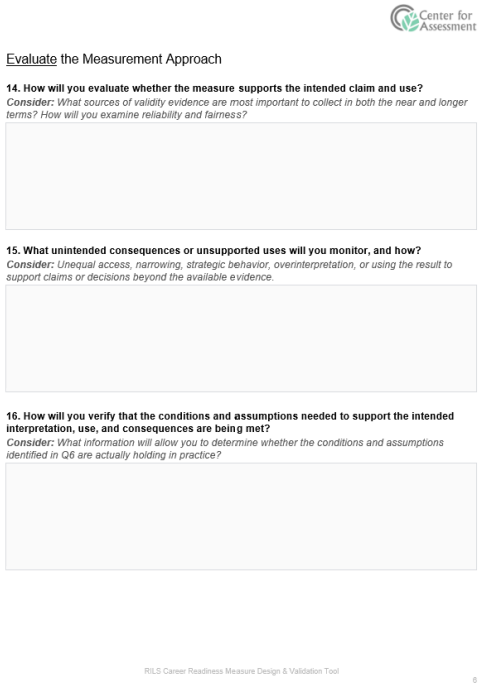
(*Standards*, p.219)

Let's Talk Validity



In Your Tables

- Go to page 6 of the tool.
- Discuss the three evaluation questions as they pertain to your proposed measure and in light of what you just heard from the panelists.
- Voluntell one of your group members to share out with the full group at 10:15.



Evaluate the Measurement Approach

14. How will you evaluate whether the measure supports the intended claim and use?
Consider: *What sources of validity evidence are most important to collect in both the near and longer terms? How will you examine reliability and fairness?*

15. What unintended consequences or unsupported uses will you monitor, and how?
Consider: *Unequal access, narrowing, strategic behavior, overinterpretation, or using the result to support claims or decisions beyond the available evidence.*

16. How will you verify that the conditions and assumptions needed to support the intended interpretation, use, and consequences are being met?
Consider: *What information will allow you to determine whether the conditions and assumptions identified in Q6 are actually holding in practice?*

RILS Career Readiness Measure Design & Validation Tool 6