



# Emerging Measurement Issues in the WWC 5.0 Standards

---

Allan Porowski, Carter Epstein, and Ginger Stoker

Education Innovation and Research (EIR) Project Directors and Evaluators Meeting

Washington, DC, February 8-9, 2024



# Why Outcome Measures Matter

---

- The measures you choose can affect:
  - Your ability to determine whether an intervention is effective
  - How the WWC reviews your study
- Choose outcome measures that are:
  - Aligned with your logic model
  - Valid and reliable
  - Sensitive to change
  - Collected consistently in the intervention and comparison groups
  - Aligned to a WWC outcome domain



# Agenda

---

- New outcome domains in WWC version 5.0
  - Overview of key changes
  - Implications for AP and dual enrollment outcomes
  - More on the Access to Educational Opportunity domain
- Additional considerations for outcome measures
  - Reliability
  - Independence of outcome measures
  - Over-alignment



# New outcome domains in WWC version 5.0

---



# Overview of Key Changes

# Key Changes to Outcome Domains in WWC v5.0



New Domain in v5.0	Change from v4.1
<b>Academic Readiness, Knowledge, or Skills</b>	
Course Performance	<b>NEW</b> – Includes course grades, <i>formerly part of Academic Achievement</i>
Physical Sciences	<b>REVISED</b> – Combines (1) <i>Chemistry &amp; Physics</i> with (2) <i>Earth/Space Science</i>
Literature	<b>NEW</b> – Includes analysis of literary work
<b>School Leader &amp; Teacher Outcomes</b>	
School Leader Well-Being	<b>NEW</b> – Includes satisfaction, burnout, staff support, and anxiety
Teacher Well-Being	<b>NEW</b> – Includes satisfaction, burnout, staff support, and anxiety

# New and Renamed Outcome Domains in WWC v5.0



New Domain in v5.0	Change from v4.1
<b>Social, Emotional, Behavioral, &amp; Mental Health</b>	
Academic Dispositions	<b>NEW</b> – Student attitudes toward academics or participation in school activities
Civic, Social, & Economic Participation	<b>NEW</b> – Activities and skills beneficial for functioning within society
Mental Health	<b>RENAMED</b> , previously <i>Intrapersonal Competencies</i> – Student emotional & psychological well-being
<b>School Outcomes &amp; Educational Opportunity</b>	
Access to Educational Opportunity	<b>NEW</b> – Includes access and/or participation in opportunities that can influence college readiness
School Accountability Metrics	<b>RENAMED</b> , previously <i>School Academic Performance</i>



# Implications for AP and dual enrollment outcomes

---



# Measures of Performance in AP or Dual Enrollment Fall in Several WWC Domains



## Course Performance – Secondary Domain

- Course grade in an AP or dual enrollment course

## College Readiness Domain

- Passing AP or dual enrollment course
- Number of courses passed
- Number of credits earned
- Earning qualifying score on any AP exam(s)
- Number of AP exams passed

## Other Subject-Specific Domains

- Earning a qualifying score on an AP exam in a subject area identified by WWC
- Score on an AP exam in a subject area identified by WWC  
e.g., Earning a qualifying score on AP Calculus exam → Calculus and Precalculus domain





# What This Means for EIR Evaluations

---

- Select measures that reflect the outcomes targeted by the intervention
- Know how the WWC will classify multiple measures across domains
  - The WWC will average together the effects of the intervention on multiple measures in the same domain
  - Consider whether all outcomes are equally important



# More on the Access to Educational Opportunity Domain

---

# What is the Access to Educational Opportunity Domain?

- **Access to OR participation in educational opportunities**

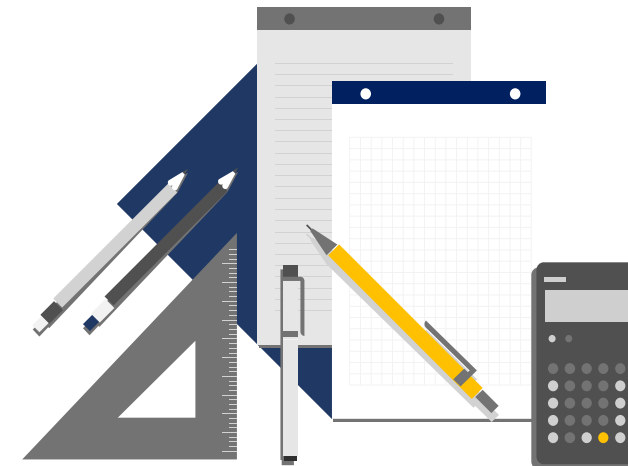
- From pre-k to postsecondary that can influence whether a student remains on track for college readiness from pre-K to postsecondary
- For example, STEM courses (such as *Algebra 1*), advanced courses, Pre-K, and gifted and talented programs

- **Access to teachers**

- Who are qualified, experienced, racially and ethnically representative or diverse, and/or effective

- **Access to resources**

- Needed for learning, such as Wi-Fi





# What This Means for Your Outcome Measures

---

- Ensure the Access to Educational Opportunity domain focuses on school outcomes
- Measure all eligible students in the school for participation in educational opportunities
- Consider whether measures of access are over-aligned with the intervention
  - Is the outcome measuring whether implementation occurred?
  - Example:
    - Intervention = Schools offer more AP courses
    - Outcome = Number of AP courses offered

**OVER-ALIGNED**



# What This Means for Your Baseline Measures

If there is no natural pretest → not acceptable baseline measure

- For example, for *participation in a new STEM course*, can't measure participation for a course that was not offered yet 0% of students in the school participated at baseline

Other acceptable baseline measures

- A broad, approximately continuous & standardized measure of student academic readiness, knowledge or skills **AND**
- At least two of the following: socioeconomic status, race/ethnicity, English learner/dual language status, disability status, disciplinary measures, and grade level/age



# Additional Considerations for Outcome Measures



# Reliability





---



# Measures Must Meet 4 Standards

---

- To be eligible for review by the WWC, an outcome must...

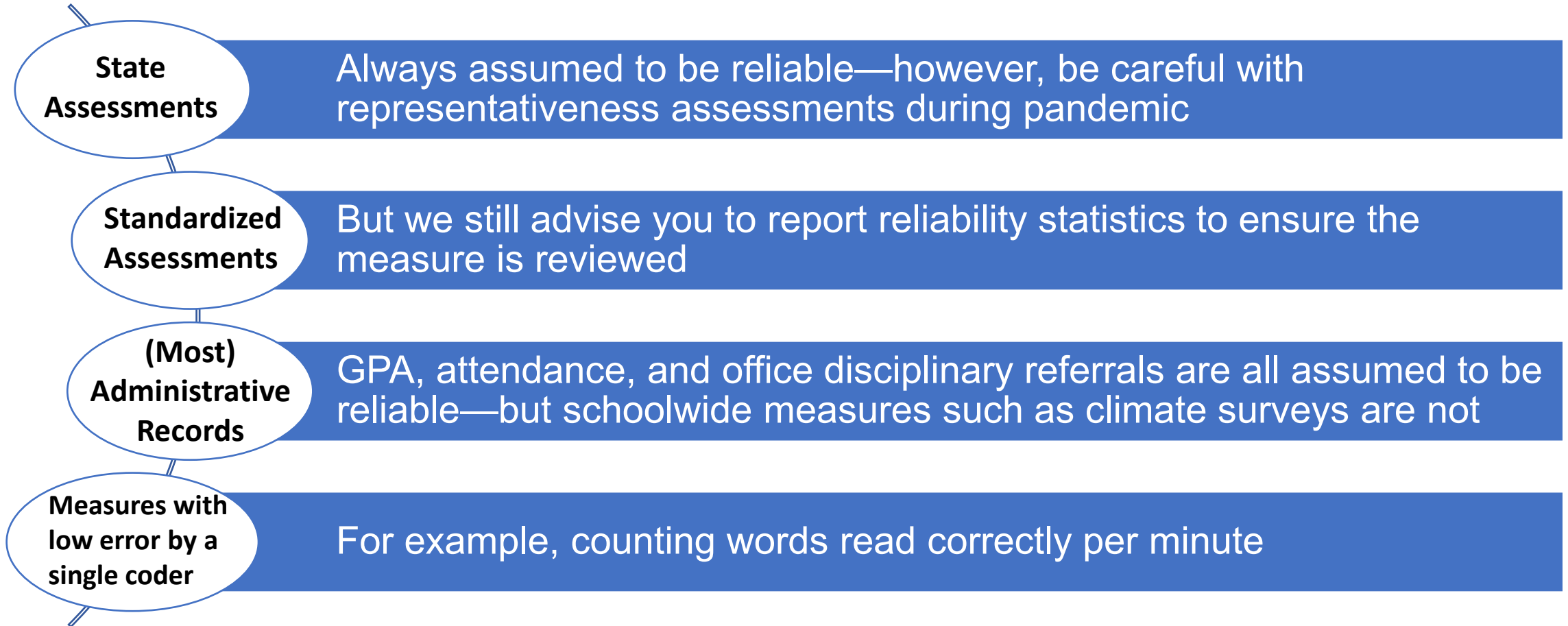
	<b>Have face validity</b>	Measure appears to measure what it's supposed to measure
	<b>Demonstrate reliability</b>	Measure produces consistent findings
	<b>Not be over-aligned</b>	Doesn't provide unfair advantage to one group over another
	<b>Be measured consistently</b>	Same data collection procedures

# Report at Least One Form of Reliability ...for Most Measures



Measure	Definition	WWC Standard
Internal Consistency	Do the measure's items provide consistent information about individual respondents?	$\geq .60$
Test-Retest Reliability	Do individuals get similar scores in repeated assessments?	$\geq .40$
Inter-Rater Reliability	Do raters score individuals consistently?	$\geq .50$ (correlation) $\geq .80$ (percentage agreement) $\geq .60$ (kappa)

# Some Measures Are Assumed to Be Reliable





# What This Means For You

---

- Select measures demonstrated to meet reliability standards based on findings from other studies
- Pilot measures developed for your evaluation to ensure they meet WWC reliability standards
- Include reliability information in your report for any measures not assumed to be reliable



# Independence of Outcome Measures

# Measures Developed for a Specific Intervention May Overestimate Effects

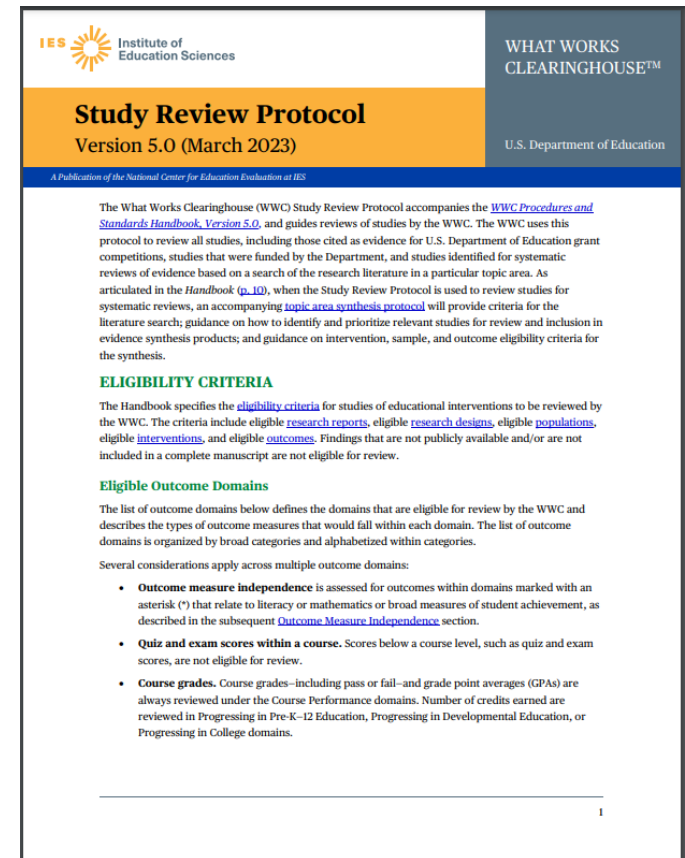
*“A measure will be considered non-independent if either it was developed by the study authors and is not in broader use, or if it was developed by the intervention’s developers”*

- Measures developed to be used with a particular intervention will always be non-independent if used to estimate impacts of that same intervention
- These criteria apply to literacy and mathematics domains only

# WWC Study Review Protocol Lists Known Independent Measures



- For literacy and math domains, WWC relies on independent measures to determine effectiveness
- If a literacy or math measure is determined non-independent:
  - Effect not reviewed as a main finding
  - Finding not factored into WWC effectiveness rating (only main findings)
  - Is eligible as a supplemental finding



<https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/WWC-SRP50-508.pdf>



# What This Means for You<sub>2</sub>

---

- Use independent measures if you're measuring literacy or math outcomes

**BUT**

- The list of independent measures is not exhaustive
- Consult with your TA provider if you have questions





# Over-alignment

---



# Measures Should Not Provide an Unfair Advantage to One Group Over Another

This type of unfair advantage is called “**overalignment.**”

The WWC considers a measure to be overaligned if:

1. It contains content or materials provided to one study group but not the other, **OR**
2. It is based on learning objectives known to implementers (e.g., teachers) in one study group but not the other.



# What This Means for You<sub>3</sub>

---

- Use measures validated in other studies, if possible, to avoid over-alignment between measures developed specifically for the study and the intervention.
- Talk to the grantee to learn whether individuals in the comparison condition will be exposed to similar content.
- Be sure to document what materials are provided in study conditions.
- Describe the extent to which the learning objectives and content of instruction is similar in both conditions, if possible.



Questions?



# Contact

---

EIR Evaluation TA Team

[EIREvalTA@AbtAssoc.com](mailto:EIREvalTA@AbtAssoc.com)



BOLD  
THINKERS  
DRIVING  
REAL-WORLD  
IMPACT

[www.abtassociates.com](http://www.abtassociates.com)

## Suggested citation:

Porowski, A., Epstein, C., & Stoker, G. (February 8-9, 2024). *Emerging Measurement Issues in the WWC 5.0 Standards*. Presentation at the Education Innovation and Research (EIR) Project Directors and Evaluators Meeting, Washington, D.C.