

STATISTICAL BIAS

What is it?

When evaluating the quality of any screening tool, it is important to determine whether or not the assessment is *biased* against different groups of students. The reason for this is that we want to ensure that students do not receive higher or lower screening scores for reasons other than the primary skill or trait that is being tested. In other words, we want the screener to give us an accurate assessment of risk for all students.

How do I know if a screening tool is biased?

To understand bias, screening test developers may conduct an analysis that examines the degree to which a screening tool is or is not biased against subgroups (e.g., race/ethnicity, gender, socioeconomic status, students with disabilities, English learners). In general, comparisons of group scores are not sufficient for demonstrating bias or the lack thereof because the properties of the items may be conflated with the properties of the students taking the assessment. For example, some items may be more difficult for students who have not been exposed to background knowledge in a certain area. Instead, more sophisticated statistical investigations that examine both properties of items and students at the same time are better suited to provide rigorous examinations of bias.

A joint set of standards set forth by the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education states that, "Fairness is a fundamental validity issue and requires attention throughout all stages of test development and use."

Example

Consider a vocabulary assessment that asks you to define the word "sphygmomanometer" (the technical word for the instrument used to measure your blood pressure). Although you understand what a sphygmomanometer is and what it is used for, you may not have had the exposure to advanced medical training that would make you aware of the more technical term. If this item was included on a medical licensing exam, it would be completely appropriate, but if it is included in an assessment designed to measure general language skills of adults, one might argue this item is biased toward those with formal medical training.

Where do I go from here?

To learn more about statistical bias in screening measures, visit the National Center on Intensive Intervention's (NCII's) [academic](#) and [behavior](#) screening tools charts. NCII publishes these charts to assist educators and families in becoming informed consumers who can select screening tools that best meet their needs.

For more information on literacy screening processes, see resources from the National Center on Improving Literacy: <https://improvingliteracy.org/>.

Academic Screening Tools Chart

Universal screening can be used to identify which children will need the most intensive intervention. In some cases, children with the weakest initial skills may bypass Tier 2 intervention and move directly into intensive intervention. The tools on the academic screening tools chart can be used to identify students at risk for poor academic outcomes, including students who require intensive intervention.

This tools chart has three tabs that include ratings on the technical rigor of the tools: (1) Classification Accuracy, (2) Technical Standards, and (3) Usability Features.

Last updated: July 2019. [Learn more about the content and structural changes to the academic screening tools chart during the most recent update.](#)

Legend

-  Convincing evidence
-  Partially convincing evidence
-  Unconvincing evidence
-  Data unavailable
- ^d Disaggregated data available

[View Chart Resources](#)

FILTER RESULTS



Subject Reading Mathematics

Grade Pre-K Elementary (K-4) Middle School (5-8) High School (9-12)

[Apply](#) [Print Chart](#)

[Hide/Show Advanced Filters](#) [Clear Filters](#)

[Reset Chart](#)
[Compare Tools](#)
[Prev Tab](#)
[Next Tab](#)

All	Title	Area	Grade	Classification Accuracy		Technical Standards		Usability Features
				Reliability	Validity	Sample Representativeness	Bias Analysis Conducted	
<input type="checkbox"/>	Acadience Reading (aka DIBELS Next)	Composite Score	K			Regional without Cross-Validation	Yes	

This handout was produced under the U.S. Department of Education, Office of Special Education Programs, Award No. H326Q160001. Celia Rosenquist serves as the project officer. The views expressed herein do not necessarily represent the positions or policies of the U.S. Department of Education. No official endorsement by the U.S. Department of Education of any product, commodity, service, or enterprise mentioned in this website is intended or should be inferred.