Improving Literacy Brief
Understanding Screening

STATISTICAL BIAS

When evaluating the quality of any screening tool, it is important to determine whether or not the assessment is statistically biased against different groups of students. We want to ensure that students do not receive higher or lower scores on an assessment for reasons other than the primary skill or trait that is being tested. Said differently, we want to be using a screener that best enables ALL students to demonstrate what they know and can do.

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 statistically biased towards those with formal medical training.

This idea of statistically unbiased or fair assessment is so critical that a joint set of standards set forth by the American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education states, "Fairness is a fundamental validity issue and requires attention throughout all stages of test development and use."

In order to understand statistical bias, screening test developers may conduct an analysis that examines the degree to which a screening tool is or is not statistically biased against subgroups (e.g., race/ethnicity, gender, socioeconomic status, students with disabilities, English learners). In general, comparisons of group means are not sufficient for demonstrating statistical bias or the lack thereof because the properties of the items are conflated with the properties of the students taking the assessment. Instead, more sophisticated investigations that examine both properties of items and students at the same time are better suited to provide rigorous examinations of statistical bias.

SUGGESTED CITATION

National Center on Improving Literacy. (2019). *Understanding Screening: Statistical Bias.* https://www.improvingliteracy.org/resource/understanding-screening-statistical-bias

WHERE DO I GO FROM HERE?

- For more information about identifying high quality screening tools: https://intensiveintervention.org/
- For more information about screening processes: https://improvingliteracy.org/whitepaper/screening-dyslexia

National Center on Improving Literacy

National Center on
INTENSIVE INTERVENTION
at American Institutes for Research

STATISTICAL BIAS

WHAT IS STATISTICAL BIAS?

Statistical bias is an important aspect of high-quality instruments used to measure academic progress. It is determining whether the reading screener is unfair to, or biased against, different groups of students. Unbiased reading screeners should enable all students to demonstrate what they know and can do.

For example, imagine a language assessment that asks you to define the word 'sphygmomanometer' (the technical word for the instrument used to measure your blood pressure). You may understand what a sphygmomanometer is and what it is used for, but you may not have received the advanced medical training to recognize its more technical term. If 'sphygmomanometer' was in a medical exam, it would be completely appropriate. Because it was in an assessment designed to measure general language skills, it may be statistically biased towards those with formal medical training.
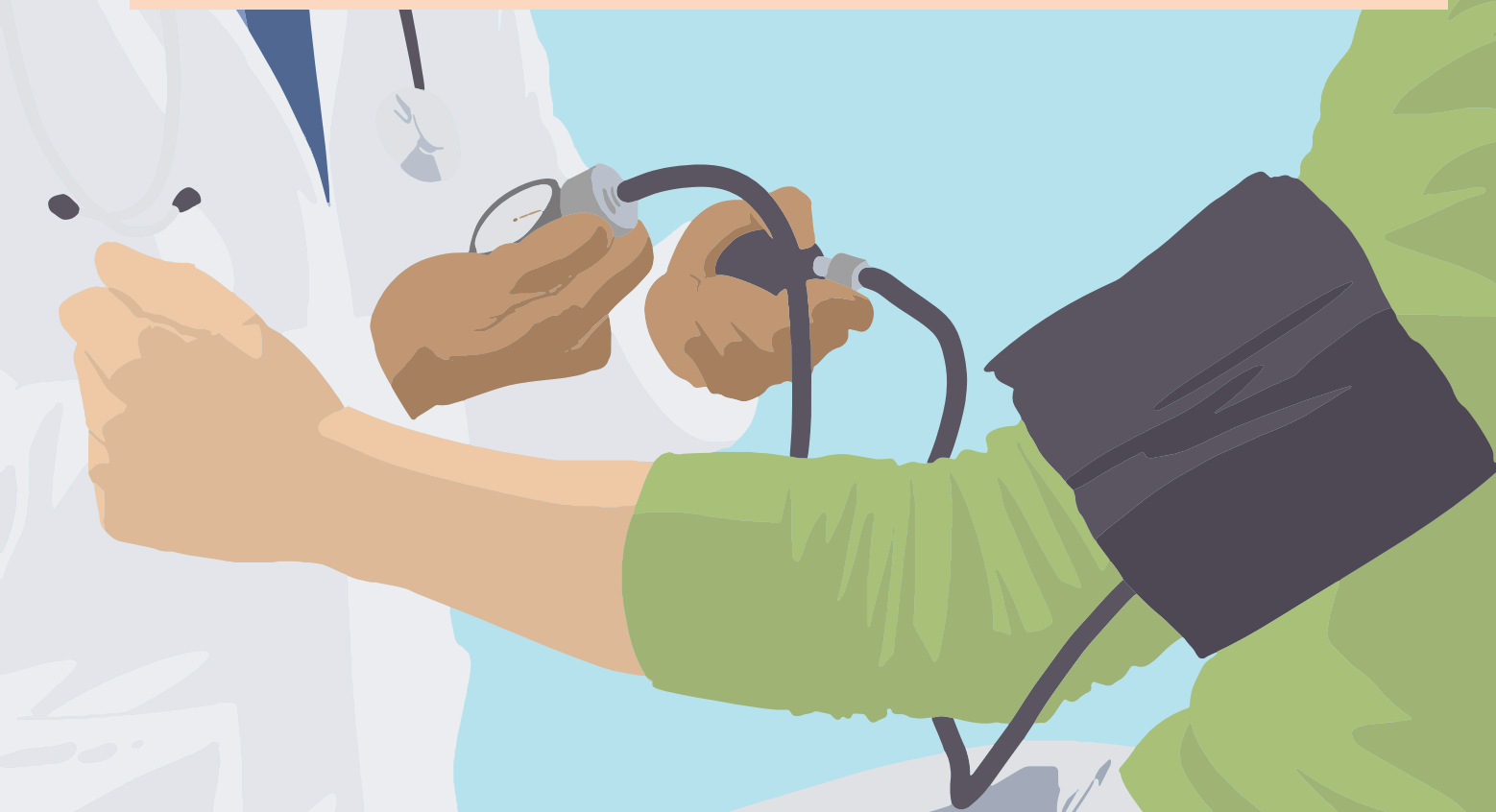
Screening test developers should conduct an analysis that examines the degree to which a screening tool is or is not statistically biased against subgroups (for example, race/ethnicity, gender, socioeconomic status, students with disabilities, English learners).