CLASSIFICATION ACCURACY

What is it?

The main goal of classification accuracy is to understand how well scores on a screening assessment correctly identify students at risk versus those not at risk.

Why is classification accuracy important?

Classifying students' risk status is a key step in universal screening because the aim is to have tools that permit accurate identification of who does and does not need additional support to make progress. Measures with strong classification accuracy have achieved two important goals:

- **Goal 1:** They *maximize* rates of **true positive** (correctly classifying a student as at risk) and **true negative** (correctly classifying a student as not at risk) classifications.
- **Goal 2:** They *minimize* rates of **false positive** (incorrectly classifying a student as at risk when he is not at risk) and **false negative** (failing to classify a student as at risk when he is at risk) classifications.

False positives may raise unnecessary concern and strain school resources by providing additional support to students who do not need it. False negatives may be particularly problematic in academic screening because students may miss out on critical additional support that they need.

How does this apply to the NCII's tools charts?

Sensitivity and **specificity** rates help gauge which tests can achieve a high rate of accurate classifications. Sensitivity is the probability of correctly identifying a problem (i.e., the proportion of true positives that the screener correctly identifies). Specificity is the probability of correctly identifying that there is not a problem (i.e., the proportion of true negatives that the screener correctly identifies). Sensitivity and specificity rates are useful when trying to determine which screening tools can distinguish, with relative accuracy, among at-risk and not at-risk students. Therefore, the National Center on Intensive Intervention's (NCII's) tools charts rate a screening tool highest when it has a sensitivity rate of 70% or higher and a specificity rate of at least 80%. Additional statistical analysis, receiver operating characteristic (ROC) area under the curve (AUC) estimates, provide information about classification accuracy at each potential score of an assessment and indicate which cut score will maximize sensitivity and specificity."

Example

The Transportation Security Administration (TSA) screening at the airport offers a useful illustration of classification accuracy. The security line process is aimed at detecting items that are not allowed on flights early in the travel process. The scanners are set at thresholds, which cause the buzzer to indicate when a certain amount of "unallowable" material has been detected. When a traveler sets off the buzzer, one possibility is that an item that is genuinely "not allowed" (e.g., a knife or nail scissors) has been transported through the scanner—this is a true positive. In this case, the scanner has done its job and steps are taken to resolve the issue. However, there also are instances in which the scanner "detects" something that may not be there or may not be problematic. Perhaps a harmless item that is "allowed" (e.g., some forgotten pocket change, a hip replacement) sets off the detector because it shares some properties (e.g., substance, shape) with "not allowed" items—this is a false positive.

On the other hand, consider a scenario in which the buzzer did not indicate any "not allowed" items. In most cases, the buzzer does not go off because there is a genuine lack of prohibited items (i.e., there was nothing there to detect in the first place—a true negative). However, it also is possible that some prohibited items (e.g., nail scissors) were there and the scanner was not set at a threshold, or sensitive enough, to prompt detection—a false negative. These four scenarios, summarized in the following graphic, illustrate a balance that TSA attempts to strike: setting a threshold on the scanner appropriately to reliably detect items when they are actually present.

		True Condition: Prohibited Item Present	True Condition: No Prohibited Items
Predicted Condition	Prohibited Item Detected	**True Positive** *Buzz for nail scissors*	**False Positive** *Buzz for hip replacement*
	No Prohibited Items Detected	**False Negative** *No buzz despite nail scissors*	**True Negative** *No buzz, no prohibited items*

Where do I go from here?

For more information about the classification accuracy of screening measures, visit 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/.

National Center on INTENSIVE INTERVENTION at American Institutes for Research

National Center on Improving Literacy

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.

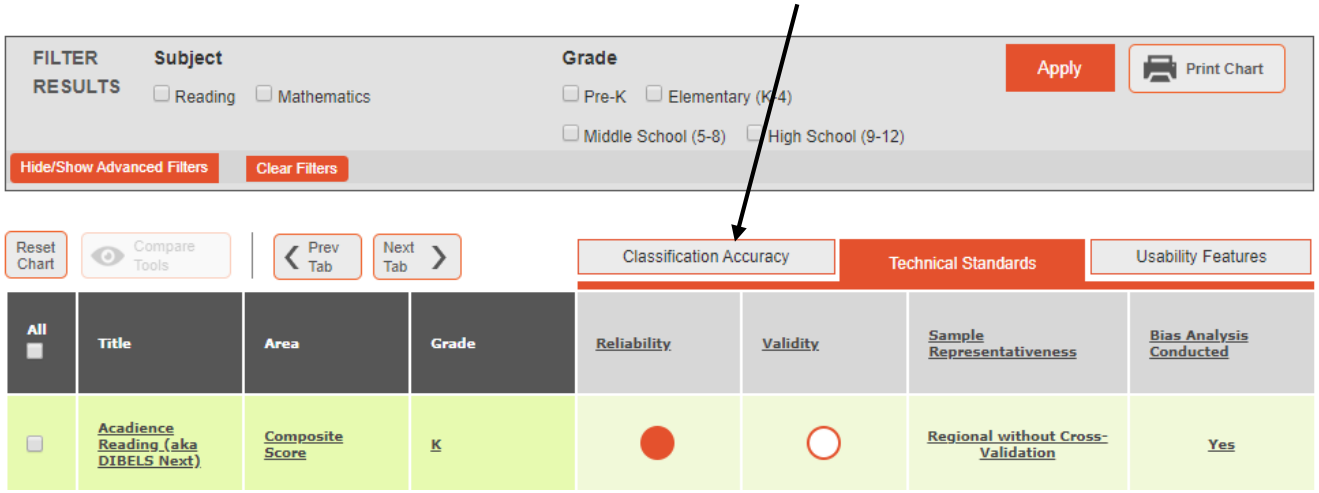
This tools chart has three tabs that include ratings on the technical rigor of the tools: (1) Classification Accuracy, (2) Technical Standards, and (3) Usability Features.

Last updated: July 2019. **Learn more about the content and structural changes to the academic screening tools chart during the most rececent update**.

Legend

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

View Chart Resources



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.

National Center on INTENSIVE INTERVENTION at American Institutes for Research

National Center on Improving Literacy