Psychometric Properties of Behavior Screening Tools

Psychometric properties reflect the overall quality of behavioral screening tools. Understanding such concepts helps you select a behavioral screener that best fits your district or school context and provides accurate and valuable information for data-based decision-making. Some psychometric characteristics of screeners, and the questions they address, are listed below. Click the titles below to check detailed information at the National Center on Intensive Intervention (NCII) website.

Content Validity

Is there evidence to suggest the screener measures the behavior of interest?

Select a screener that experts in the field agree meaningfully assesses the behaviors of interest (e.g., internalizing, externalizing). Content Validity provides the evidence needed to make this decision.

Convergent Validity

Is there information to specify whether the results from the screening tool are highly correlated with those from other well-established screening tools?

Newly designed behavior screening tools are compared with other criterion screeners to establish convergent validity. You will have more confidence in the validity of the screener if outcomes suggest strong correlation with another well-established behavior screening tool.

Internal Consistency

Do the screening items appear to measure the same construct by showing strong correlations with each other (are they internally consistent within a construct)?

Choose a screener whose items will provide consistent and useful information about student behavioral needs. High internal consistency indicates the items are highly correlated and measure a similar construct.

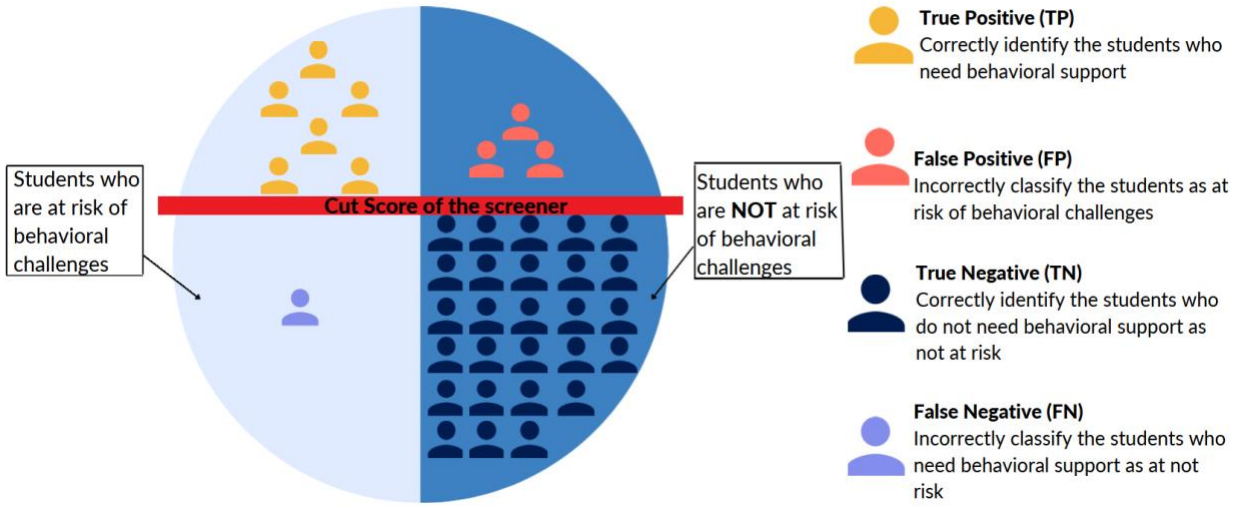
Inter-Rater Reliability

Will different teachers rate the same student similarly using the screening tool?

Choose a screening tool with high inter-rater reliability to increase your confidence that different teachers will provide consistent ratings on the same student.

Classification Accuracy of Behavior Screening Tools

A valid behavioral screening tool should provide accurate classification to distinguish between students who may benefit from additional support and those who may not need it. Four important concepts reflect the level of classification accuracy of the screening tool: positive predictive power, negative Predictive Power, specificity, and sensitivity.



Positive Predictive Power

What proportion of the students who are above the cut score of behavioral risk are actually at behavioral risk?

Students whose screening results are above the cut score of the behavioral screener are classified as at risk. But it is possible only part of the identified students are actually at behavioral risk. Positive Predictive Power is the index that indicates the proportion of the students who are above the cut score that are actually at behavioral risk. It is calculated by dividing True Positive ( ) by True Positive ( ) and False Positive ( ). Higher positive predictive power indicates students who actually have a concern are more likely to be identified by the screener.

Negative Predictive Power

What proportion of the students who are below the cut score of behavioral risk are actually not at behavioral risk?

Negative predicative power indicates the proportion of the students who are below the cut score that are actually not at behavioral risk. Higher negative predictive power means more students who do not need additional behavioral support are identified by the screener as not at risk. Negative Predictive Power is calculated by dividing the True Negative ( ) by the sum of True Negative ( ) and False Negative ( ).

Sensitivity

What proportion of the students who are actually at risk of behavioral challenges will be correctly identified by the screener?

Sensitivity is calculated by dividing True Positive ( ) by the sum of True Positive ( ) and False Negative ( ). Sensitivity indicates how well the screener "senses" the sign of behavioral risk displayed by the students who are at risk. Higher sensitivity means more students who are at risk are correctly identified. Conversely, low sensitivity may indicate under-identification of students who are at risk. In general, it is recommended to choose a screener with high sensitivity.

Specificity

What proportion of the students who are not at risk of behavioral challenges will be identified by the screener as not at risk?

Specificity is calculated by dividing True Negative ( ) by the sum of True Negative ( ) and False Positive ( ). Specificity indicates how well the screener identifies students who are not at risk of behavioral challenges. High specificity means fewer students who are not at risk for behavioral challenges will be identified as being at risk

A behavior screener with strong classification accuracy should ideally maximize rates of true positive and true negative, and minimize rates of false positive and false negative. But in reality, no measure is perfect, and the cost of missing students who are in need of additional support is greater than the cost of inaccurate identification or under-identification. Therefore, we recommend educators to focus more on positive predictive power and sensitivity of the behavior screener when choosing behavior screener if the negative predictive power and specificity is at an acceptable level. A reference criterion given by the National Center on Intensive Interventions (NCII) tools chart is that a screening tool rates highest when it has a sensitivity rate of 70% or higher and a specificity rate of at least 80%.

This document was supported from funds provided by the Center on Positive Behavioral Interventions and Supports cooperative grant supported by the Office of Special Education Programs (OSEP) and Office of Elementary and Secondary Education (OESE) of the U.S. Department of Education (H326S180001). Dr. Renee Bradley 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, or enterprise mentioned in this document is intended or should be inferred.

Suggested Citation for this Publication

Ma, Z., Lane, N., Swinburne Romine, R., Sherod, R. L., Oakes, W. P., Lane, K. L. (July, 2022). Psychometric Properties of Behavior Screening Tools. Center on PBIS, University of Oregon. www.pbis.org