

Academic Progress Monitoring Tools Chart: English tools that include English Learners

Tools Chart Tab 1: Performance Level Standards

NOTE: For all standards in Tab 1, evidence must be drawn from a sample that is representative of students across all performance levels.

1A. Reliability of Performance Level Score

Rating	Definition
Full Bubble	<p>(a) A model-based approach to reliability was reported with at least two sources of variance.</p> <p>or</p> <p>(b) At least two types of reliability were reported appropriate for the purpose of the tool (e.g., inter-rater reliability is provided for tools that require human judgment), and evidence was drawn from <u>at least two samples</u> that are representative of students across all performance levels.</p> <p>and</p> <p>For each type of reliability reported the lower bound of the confidence interval around the median coefficient met or exceeded 0.70.</p>
Half Bubble	<p>(a) A model-based approach to reliability was reported with at least two sources of variance.</p> <p>or</p> <p>(b) At least two types of reliability were reported appropriate for the purpose of the tool (e.g., inter-rater reliability is provided for tools that require human judgment), and evidence was drawn from <u>at least one sample</u> that is representative of students across all performance levels.</p> <p>and/or</p> <p>For each type of reliability, the lower bound of the confidence interval around the median coefficient met or exceeded 0.60.</p>
Empty Bubble	Does not meet full or half bubble.



1B. Validity of Performance Level Score

Rating	Definition
Full Bubble	<p>At least two types of validity were reported that were appropriately justified¹ for the tool and evidence was drawn from a sample representative of students across all performance levels.</p> <p>and</p> <p>For each type of validity reported the lower bound of the confidence interval around the median coefficient met or exceeded 0.60 (or was within an acceptable range given the expected relationship with the criterion measure(s)).</p>
Half Bubble	<p>At least two types of validity were reported that were appropriately justified¹ for the tool and evidence was drawn from a sample representative of students across all performance levels.</p> <p>and</p> <p>One type of validity met the following criteria and the other did not: the lower bound of the confidence interval around the median coefficient met or exceeded 0.60 (or was within an acceptable range given the expected relationship with the criterion measure(s)).</p>
Empty Bubble	Does not meet full or half bubble.

¹ Appropriately justified analysis must include criterion measures that are external to the progress monitoring system and theoretically linked to the underlying construct measured by the tool.

1C. Bias Analysis Conducted

Rating	Definition
Yes	<p>One or more of the following types of analyses were conducted:</p> <ol style="list-style-type: none"> 1. Multiple-group confirmatory factor models for categorical item responses 2. Explanatory group models such as multiple-indicators, multiple-causes (MIMIC) or explanatory IRT with group predictors 3. Differential Item Functioning from Item Response Theory (DIF in IRT)
No	Does not meet "yes"



Tools Chart Tab 2: Growth Standards

NOTE: For all standards in Tab 2, evidence must be drawn from a sample that includes children in need of intensive intervention.

2A. Sensitivity: Reliability of Slope

Rating	Definition
Full Bubble	The analysis conducted was appropriate with sufficient number and spacing of data points ² from a sample that includes children in need of intensive intervention. <i>and</i> The lower bound of the confidence interval around the median coefficient met or exceeded 0.50.
Half Bubble	The analysis conducted was appropriate with sufficient number and spacing of data points ² from a sample of children in need of intensive intervention. <i>and</i> The lower bound of the confidence interval around the median coefficient met or exceeded 0.40.
Empty Bubble	Does not meet full or half bubble.
Dash	Data were not provided.

² Sufficient number and spacing of data points is defined as at least 10 regularly collected measurements over a period of at least 20 weeks.

2B. Sensitivity: Validity of Slope

Rating	Definition
Full Bubble	There is at least one appropriately justified validity analysis ³ , with sufficient number and spacing of data points ⁴ , from a sample that includes children in need of intensive intervention, <i>and</i> The lower bound of the confidence interval around each coefficient met or exceeded 0.40 (or if not, within an acceptable range given the expected relationship with the criterion measure(s)).
Half Bubble	Analyses, measures, number and spacing of data points, and sample were appropriate, but evidence was mixed, with one or more measure either not meeting or exceeding 0.40 or not within an acceptable range given the expected relationship with the criterion measure(s).



Empty Bubble	Does not meet full or half bubble.
Dash	Data were not provided.

³ Appropriately justified analyses must include criterion measures that are external to the progress monitoring system and theoretically linked to the underlying construct measured by the tool.

⁴ Sufficient number and spacing of data points is defined as at least 10 regularly collected measurements over a period of at least 20 weeks.

2C. Alternate Forms

Rating	Definition
Full Bubble	There are at least 20 alternate forms and evidence is strong for comparability of alternate forms.
Half Bubble	There are at least 10 alternate forms and evidence for comparability is at least moderate for the alternate forms.
Empty Bubble	Does not meet full or half bubble.
Dash	Data were not provided.

2D. Decision Rules for Setting and Revising Goals

Rating	Definition
Yes	At least one of the following is present. The basis for establishing decision rules for setting and revising goals is: (1) evidence-based for English learners; (2) based on analysis of progress monitoring measurement collected at least monthly over the period of time that is deemed necessary for the decision rules, and (3) from a sample of EL students that includes children in need of intensive intervention.
No	Does not meet "yes"



2E. Decision Rules for Changing Instruction

Rating	Definition
Yes	The basis for establishing decision rules for when changes to instruction need to be made is (1) strongly evidence-based; (2) based on analysis of progress monitoring measurement collected at least monthly over the period of time that is deemed necessary for the decision rules, and (3) from a sample that includes children in need of intensive intervention.
No	Does not meet “yes”

This resource was produced under U.S. Department of Education, Office of Special Education Programs, Award No. H326Q210001. 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 document is intended or should be inferred.

