

Behavior Progress Monitoring Tools Chart Rating Rubric

The following rubrics are applied separately for each subscale, grade level/span, and informant targeted by the progress monitoring tool.

Tools Chart Tab 1: Performance Level Standards

For all standards in Tab 1, the expectation is that evidence comes from a representative sample of students across all performance levels. Evidence not meeting this criterion will receive a rating of “—” to indicate “data not available.”

1. Reliability

Rating	Definition
Full Bubble	At least two types of reliability were reported that are appropriate ¹ for the purpose of the tool. and The analyses are drawn from at least two samples that are representative of students across all performance levels. and The median of the estimates for each type met or exceeded 0.70.
Half Bubble	At least two types of reliability are reported that are appropriate ¹ for the purpose of the tool. and The analyses are drawn from at least one sample representative of students across all performance levels. or The median of the estimates for each type met or exceeded 0.60.
Empty Bubble	Does not meet full or half bubble.

¹ Tests that require human judgment must report interrater reliability to be eligible for a full or half bubble rating. Other types of reliability must include justification of appropriateness given the purpose of the tool.



2. Validity

Rating	Definition
Full Bubble	At least two types of appropriately justified ² validity analyses are reported. and The analyses are drawn from at least one sample representative of students across all performance levels. and The median of the estimates for each met or exceeded 0.60 (or was within an acceptable range given the expected relationship with the criterion measure(s)).
Half Bubble	One type of appropriately justified ² validity analysis is reported. and The analysis is drawn from a sample representative of students across all performance levels. and The median of the estimates met or exceeded 0.60 (or was within an acceptable range given the expected relationship with the criterion measure(s)).
Empty Bubble	Does not meet full or half bubble.

² Appropriately justified analyses must include at least one criterion measure that is external to the progress monitoring system and theoretically linked to the underlying construct measured by the tool.

3. Bias Analysis

Bias analysis refers to an analysis that examines the degree to which a tool is or is not biased against subgroups (e.g., race/ethnicity, gender, socioeconomic status, students with disabilities, English learners).

Rating	Definition
Provided	One or more of the following types of analyses were conducted: <ul style="list-style-type: none"> ▪ Multiple-group confirmatory factor models for categorical item responses ▪ Explanatory group models, such as multiple-indicators, multiple-causes or explanatory item response theory (IRT) with group predictors ▪ Differential Item Functioning from Item Response Theory (DIF in IRT)
Not Provided	Does not meet "Provided"



Tools Chart Tab 2: Growth Standards

For all standards in Tab 2, the expectation is that evidence comes from a sample of students in need of behavioral intervention. Convincing evidence that children need behavioral intervention may include one or more of the following: students have an emotional disturbance label; students are in an alternative school/classroom; students demonstrate nonresponse to moderately intensive intervention (e.g., Tier 2); or students demonstrate severe problem behaviors (e.g., Tier 3), according to an evidence-based tool (e.g., systematic screening tool or direct observation).

4. Sensitive to Behavior Change

Rating	Definition
Full Bubble	The basis for assuming that the data are sensitive to incremental change is strong (e.g., the range of possible scores is sufficient to detect small changes, and documentation of sensitivity to change is consistent with another criterion).
Half Bubble	The basis for assuming that the data are sensitive to incremental change is moderate (e.g., the range of possible scores is sufficient to detect a change, and there is documentation of sensitivity to change).
Empty Bubble	Does not meet full or half bubble.
Dash	Data not provided.

5. Reliability and Validity: Intensive Population

Rating	Definition
Full Bubble	<p>At least two types of reliability are reported that meet the following criteria:</p> <ul style="list-style-type: none"> Are appropriate³ for the purpose of the tool. Are drawn from at least two samples representative of students in need of intensive intervention. The median of the estimates for each type met or exceeded 0.70. <p>and</p> <p>At least two types of validity are reported that meet the following criteria:</p> <ul style="list-style-type: none"> Are appropriately justified⁴. Are drawn from at least one sample representative of students in need of intensive intervention. The median of the estimates for both types of validity met or exceeded 0.60 (or was within an acceptable range given the expected relationship with the criterion measure(s)).



Rating	Definition
Half Bubble	<p>One or more of the following were met:</p> <ol style="list-style-type: none"> 1. At least two types of reliability are reported that meet the following criteria: are appropriate for the purpose of the tool, are drawn from at least two samples representative of students in need of intensive intervention, and the median of the estimates for each type met or exceeded 0.60. However, validity was not reported. 2. At least two types of validity are reported that meet the following criteria: are appropriately justified⁴, are drawn from a sample representative of students in need of intensive intervention, and the median of the estimates for both types of validity met or exceeded 0.60 (or was within an acceptable range given the expected relationship with the criterion measure(s)). However, reliability was not reported. <p>One type of reliability was reported that met the criteria in 1. and One type of validity was reported that met the criteria in 2.</p>
Empty Bubble	Does not meet full or half bubble.
Dash	Data not provided.

³ Tests that require human judgment must report interrater reliability to be eligible for a full or half bubble rating. Other types of reliability must include justification of appropriateness given the purpose of the tool.

⁴ 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.

6. Decision Rules: Changing Intervention

Rating	Definition
Full Bubble	<p>The data provided to support decisions about intervention change are strong. and The data are based on analysis of progress monitoring measurement collected at least weekly during the period deemed necessary for the decision rules. and The data come from a sample of students in need of intensive intervention.</p>
Half Bubble	<p>The data provided to support decisions about intervention change are moderate. and The data are based on analysis of progress monitoring measurement collected at least weekly during the period deemed necessary for the decision rules. and The data come from a sample of students in need of intensive intervention.</p>
Empty Bubble	Does not meet full or half bubble.
Dash	Data not provided.



7. Decision Rules: Choosing Intervention

Rating	Definition
Full Bubble	<p>The data to support guidance on intervention choice (e.g., a class of relevant interventions or a specific intervention) are strongly evidence based.</p> <p>and</p> <p>The data are based on analysis of progress monitoring measurement collected at least weekly during the period deemed necessary for the decision rules.</p> <p>and</p> <p>The data come from a sample of students in need of intensive intervention.</p>
Half Bubble	<p>The data to support guidance on intervention choice (e.g., a class of relevant interventions or a specific intervention) are moderately evidence based.</p> <p>and</p> <p>The data are based on analysis of progress monitoring measurement collected at least weekly during the period deemed necessary for the decision rules.</p> <p>and</p> <p>The data come from a sample of students in need of intensive intervention.</p>
Empty Bubble	Does not meet full or half bubble.
Dash	Data not provided.

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.

