Behavior Progress Monitoring Tools Chart Rating Rubric

*Please note that the following rubrics are applied separately for **each sub-scale, grade level/span, and informant** targeted by the tool.*

Tools Chart Tab 1: Foundational Psychometric Standards

Note: For all standards in Tab 1, it is expected that evidence is drawn from a sample that is representative of students across all performance levels. Tools not intended for use with students across all performance levels may receive a rating of "N/A" in these categories.

1A. Reliability

Rating	Definition
Full Bubble	(a) A model based approach to reliability was reported with at least two sources of variance *or* (b) At least two other types of reliability were reported appropriate for the purpose of the tool, and drawn from at least two samples that are representative of students across all performance levels *and* For each type of reliability reported the lower bound of the confidence interval around the median estimate met or exceeded 0.70.
Half Bubble	A model-based approach to reliability was reported with at least two sources of variance *or* At least two other types of reliability were reported appropriate for the purpose of the tool, drawn from at least one sample that is representative of students across all performance levels *and/or* For each type of reliability reported the lower bound of the confidence interval around the median estimate fell below 0.70 but met or exceeded 0.60.
Empty Bubble	Does not meet full or half bubble.

1B. Validity

Rating	Definition
Full Bubble	There are at least two types of appropriately justified validity analyses* from a sample representative of students across all performance levels *and* The lower bound of the confidence interval around the each estimate met or exceeded 0.60 (or if not, within an acceptable range given the expected relationship with the criterion measure(s)).
Half Bubble	Analyses, measures, and sample were appropriate, but evidence was mixed, with one or more measures either not meeting or exceeding 0.60 or not within an acceptable range given the expected relationship with the criterion measure(s).
Empty Bubble	Does not meet full or half bubble.

* 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.

1C. Bias Analysis Conducted

Rating	Definition
Yes	One or more of the following three types of analyses were conducted: 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). 4. Testing differential classification accuracy across demographic groups
No	Fails "yes."

Tools Chart Tab 2: Progress Monitoring with Intensive Population

Note: For all standards in Tab 2, it is expected that evidence is drawn from a sample of student in need of intensive intervention. Convincing evidence that children were in need of intensive intervention may include one or more of the following: students have ED label; students are placed in an alternative school/classroom; students have demonstrated non-response to moderately intensive intervention (e.g., Tier 2); or students have demonstrated severe problem behaviors (e.g., Tier 3), according to an evidence-based tool (e.g., systematic screening tool or direct observation).

2A. 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 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 documentation of sensitivity to change).
Empty Bubble	Does not meet full or half bubble.
Dash	Data were not provided.

2B. Reliability: Intensive Population

Rating	Definition
Full Bubble	A model based approach to reliability was reported with at least two sources of variance *or* At least two other types of reliability were reported appropriate for the purpose of the tool, and drawn from at least two samples that are representative of students in need of intensive intervention *and* For each type of reliability reported the lower bound of the confidence interval around the median estimate met or exceeded 0.70.

Rating	Definition
Half Bubble	A model-based approach to reliability was reported with at least two sources of variance *or* At least two other types of reliability were reported appropriate for the purpose of the tool, drawn from at least one sample that is representative of students in need of intensive intervention *and/or* For each type of reliability reported the lower bound of the confidence interval around the median estimate fell below 0.70 but met or exceeded 0.60.
Empty Bubble	Does not meet full or half bubble.
Dash	Data were not provided.

2C. Validity: Intensive Population

Rating	Definition
Full Bubble	There are at least two types of appropriately justified validity analyses* from a sample representative of students in need of intensive intervention *and* The lower bound of the confidence interval around the each estimate met or exceeded 0.60 (or if not, within an acceptable range given the expected relationship with the criterion measure(s)).
Half Bubble	Analyses, measures, and sample were appropriate, but evidence was mixed, with one or more measures either not meeting or exceeding 0.60 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.

2D. Data to Support Intervention Change

Rating	Definition
Full Bubble	The data provided to support decisions about intervention change is (1) strong (2) based on analysis of progress monitoring measurement collected at least weekly over the period of time that is deemed necessary for the decision rules; and (3) from a sample of students that is in need of intensive intervention.
Half Bubble	The data provided to support decisions about intervention change has (1) moderate empirical support for targeted behavior; (2) is based on analysis of progress monitoring measurement collected at least weekly over the period of time that is deemed necessary for the decision rules; and (3) is from a sample of students that is in need of intensive intervention.
Empty Bubble	Does not meet full or half bubble.
Dash	Data were not provided.

2E. Decision Rules for Changing Instruction

Rating	Definition
Full Bubble	The data provide guidance on intervention choice (e.g., a class of relevant interventions or a specific intervention), that is (1) strongly evidence-based; (2) based on analysis of progress monitoring measurement collected at least weekly over the period of time that is deemed necessary for the decision rules, and (3) from a sample of students that is in need of intensive intervention.
Half Bubble	The data provide guidance on intervention choice (e.g., a class of relevant interventions or a specific intervention), that is (1) moderately evidence-based; (2) based on analysis of progress monitoring measurement collected at least weekly over the period of time that is deemed necessary for the decision rules, and (3) from a sample of students that is in need of intensive intervention.
Empty Bubble	Does not meet full or half bubble.
Dash	Data were not provided.