|  |
| --- |
| National Center on Intensive Intervention |
| Standard ProtocolFor Evaluating Academic and Behavior Screening Tools |

July 2018

**NCII Screening Definition**

The National Center on Intensive Intervention defines screening as a process using tools with convincing evidence of classification accuracy, reliability, and validity to identify students who may require intensive intervention efforts to meet their academic, social, emotional, and/or behavioral needs.

## **Please Read and Complete Before You Start**

Are you submitting an academic screening tool (assessing students’ performance in English language arts or math) or a behavior screening tool (assessing students’ social, emotional, or behavior skills)?

Academic

Behavior

To establish whether your tool qualifies for review, please answer the following questions:

|  |  |  |
| --- | --- | --- |
| 1. Can you provide direct evidence\* on the effects of using your tool?   *\*Direct evidence refers to data from a study that has been conducted based on the tool submitted for evaluation. Studies that use data from the use of another tool, even if it is similar, are considered indirect evidence and will not be considered as adequate evidence for the purposes of this review.* | Yes | No |
| 1. Do you have the following classification data for...   an **academic** tool:   * Specificity, sensitivity, positive predictive power, negative predictive power, AND Area Under the Curve (AUC) derived from a Receiver Operating Characteristic (ROC) curve analysis?   a **behavior** tool:   * + Specificity, sensitivity, positive predictive power, and negative predictive power,   OR   * + Area Under the Curve (AUC) derived from a Receiver Operating Characteristic (ROC) curve analysis? | Yes | No |
| 1. Are your classification data analyses conducted using cut points identifying students in need of intensive intervention? For academic tools, this could be students falling below 20th percentile on local or national norm. For behavior tools, this could be students exhibiting the highest level of risk for behavior tools. | Yes | No |
| 1. Does your outcome variable match the construct your tool measures (i.e., reading for a reading screener, math for a math screener, behavior for a behavior screener)? | Yes | No |
| 1. Are there at least three months between the screening and your outcome measure for classification accuracy and predictive validity analysis? | Yes | No |

If you cannot answer YES to all these questions, we will not review your tool. If you can answer YES to all the above questions, then your tool qualifies for review. Please proceed to the following page and begin filling out the protocol.

**Please note:** If your tool assesses more than one component skill within a content area or more than one behavior, you must submit separate protocols for each sub-test that is used to identify students, or provide data on use of the tool as a composite screener.

NCII staff will review all your submitted materials to ensure that they adhere to the qualifications for review, stated above. If it is found that your submission packet needs substantial amount of supplemental information or is missing critical information, the entire packet will be returned to you. A revised protocol packet with additional information may be re-submitted.

Results of the review will be posted on the NCII website, in the Academic Screening Tools Chart. **Once the review has begun, withdrawal of tools from the process will not be permitted**.

NCII staff are available to answer questions or to assist you completing the protocol for submission. Please contact the National Center on Intensive Intervention:

National Center on Intensive Intervention

American Institutes for Research

1000 Thomas Jefferson Street, NW

Washington, DC 20007

Website: [www.intensiveintervention.org](http://www.intensiveintervention.org)

E-mail: [ToolsChartHelp@air.org](mailto:ToolsChartHelp@air.org)

[NCII@air.org](mailto:NCII@air.org)

## Marketing Language Agreement

To be eligible for review, you must read and sign this marketing language agreement.

By signing this agreement, I have indicated my understanding of the intent and purpose of the NCII tools chart, and my agreement to use language that is consistent with this purpose in any marketing materials that will be used to publicize my product’s presence and ratings on the chart.

Specifically, I understand the following:

1. The Technical Review Committee (TRC) rated each submitted tool against established criteria but did not compare it to other tools on the chart. The presence of a particular tool on the chart does not constitute endorsement and should not be viewed as a recommendation from either the TRC or the National Center on Intensive Intervention.
2. All tools submitted for review are posted on the chart, regardless of results. The chart represents all tools that were reviewed, not those that were “approved.”

When marketing my product, I will not use any language that is inconsistent with the above. Examples of inappropriate marketing language include, but may not be limited to, the following:

* + Reference to a “top-ranked” product in comparison to other products on the chart
  + Reference to “approval” or “endorsement” of the product by the NCII

If the Center becomes aware of any marketing material on my product that violates this agreement, I understand that I risk removal of the product from the chart. I also understand that I may draft language and submit to Center staff for review in advance of releasing it, in order to ensure compliance with this agreement.

**I have read and understand the terms and conditions of this Agreement. By signing below, I signify my agreement to comply with all requirements contained herein.**

Signature Date

Print Name

Organization

## Section I: Basic Information

### A. Tool Information

1. Screening Tool Name:
2. Developer:
3. Publisher:
4. Publication Date:
5. Submission Contacts
   1. Primary Contact:

Title/Organization:

Email address:

Telephone:

* 1. Alternate Contact:

Title/Organization:

Email address:

Telephone:

### B. Descriptive Information

1. Provide a brief (1-2 paragraph) overview and description of the tool:

1. What grade(s) does the tool target? (Check all that apply.)

Pre-K and younger

Kindergarten

1st grade

2nd grade

3rd grade

4th grade

5th grade

6th grade

7th grade

8th grade

9th grade

10th grade

11th grade

12th grade and older

1. What age(s) does the tool target, if applicable? (Check all that apply.)

0-4 years old

5 years old

6 years old

7 years old

8 years old

9 years old

10 years old

11 years old

12 years old

13 years old

14 years old

15 years old

16 years old

17 years old

18+ years old

1. What is the language of administration?

1. **ACADEMIC INTERVENTION**: What skills does the tool screen? (Check all that apply.)

|  |  |
| --- | --- |
| **Reading** | |
| Phonological processing:  RAN  Memory  Awareness  Letter sound correspondence  Phonics  Structural analysis  Word ID  Accuracy  Speed  Nonword  Accuracy  Speed  Spelling  Accuracy  Speed  Passage  Accuracy  Speed | Reading comprehension:  Multiple choice questions  Cloze  Constructed Response  Retell  Maze  Sentence verification  Other (please describe):  Listening comprehension:  Multiple choice questions  Cloze  Constructed Response  Retell  Maze  Sentence verification  Vocabulary  Expressive  Receptive |

|  |
| --- |
| **Mathematics** |
| Global Indicator of Math Competence  Accuracy  Speed  Multiple Choice  Constructed Response  Early Numeracy  Accuracy  Speed  Multiple Choice  Constructed Response  Mathematics Concepts  Accuracy  Speed  Multiple Choice  Constructed Response  Mathematics Computation  Accuracy  Speed  Multiple Choice  Constructed Response  Mathematic Application  Accuracy  Speed  Multiple Choice  Constructed Response  Fractions/Decimals  Accuracy  Speed  Multiple Choice  Constructed Response  Algebra  Accuracy  Speed  Multiple Choice  Constructed Response  Geometry  Accuracy  Speed  Multiple Choice  Constructed Response  Other (please describe): |

|  |
| --- |
| **Other** |
| List specific skills or subtests: |

1. **BEHAVIOR ONLY:** Which category of behaviors does your tool target?

Internalizing

Externalizing

1. **BEHAVIOR ONLY:** Please identify which broad domain(s)/construct(s) are screened by your tool and define each sub-domain or sub-construct. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

### C. Acquisition Information

1. Where can your tool be obtained?

Website:

Address:

Phone number:

Email address:

1. Cost Information:

**Initial cost for implementing program:**

Cost $Enter amount here.

Unit of cost (*e.g., district, school, student*) Enter unit here.

**Replacement cost for subsequent use:**

Cost $Enter amount here.

Unit of cost (*e.g., district, school, student*) Enter unit here.

License Duration (*e.g., year, number of forms*) Enter duration here.

Additional Cost Information: Describe basic pricing plan and/or cost structure of the tool, including, as applicable: cost per student per year, start-up or other one-time costs, reoccurring costs, training cost, and what is included in the published tool. Also, provide information on what is not included but required for implementation (e.g., computer and/or internet access.)

1. Provide information about available accommodations for students with disabilities.

1. Does the tool require technology?

Yes

No

1. If yes, what technology is required to implement your tool? (Select all that apply)

Computer or tablet

Internet connection

Other technology (please specify)

If your program requires additional technology not listed above, please describe the required technology:

## Section II: Usability

### A. Time, Administration, and Frequency

1. **BEHAVIOR ONLY:** What type of informant (administrator) is your tool designed for? (Check all that apply.)

General education teacher

Special education teacher

Parent

Child

External observer

Other (please specify):

1. **BEHAVIOR ONLY:** Can students be rated concurrently by one administrator?

Yes

No

If yes, how many students can be rated concurrently?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. What is the administration format? (Check all that apply.)

Direct observation

Rating scale

Checklist

Performance measure

Questionnaire

Direct: Computerized

One-to-one

1. What is the administration context? (Check all that apply.)

Large group  Small group

Individual

Other (please specify):

1. How long does it take to administer?

Minutes per student:

Minutes per total group:

1. How long does it take to score?

Minutes per student:

Minutes per total group:

Scoring is automatic:

1. Does your tool provide discontinue rules? (Check all that apply.)

Not provided

Basals

Ceilings

Other (please specify):

1. If relevant, how many alternate forms are available?

[#] alternate forms per [grade/level/unit]

1. Are norms available?

Yes

No

1. Are benchmarks available?

Yes

No

If yes, how many benchmarks per year? \_\_\_\_\_\_\_\_\_\_

If yes, for which months are benchmarks available? \_\_\_\_\_\_\_\_\_\_\_

### B. Training

1. How long is tester training?

Less than 1 hour of training

1-4 hours of training

4-8 hours of training

Training not required

Information not available

1. Are there minimum qualifications of the examiner?

Yes (please describe): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

No

1. Are training manuals and materials available?

Yes

No

1. Are training manuals/materials field-tested?

Yes

No

1. Are training manuals/materials included in cost of the tool?

Yes

No

If not, please describe training costs: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. How can customers obtain ongoing technical support?

### C. Scoring

1. How are scores calculated?

Scores calculated automatically

Scores calculated manually

1. What types of scores are available? (Check all that apply.)

Raw score  Standard score

Percentile score  Grade equivalents

IRT-based score  Age equivalents

Stanines  Normal curve equivalents

Developmental benchmarks  Developmental cut points

Equated  Probability

Lexile score  Error analysis

Composite scores  Subscale/subtest scores

Other (Please specify):

1. Is a scoring key available?

Yes (please describe): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

No

1. What is the basis for calculating standard and percentile scores? (Check all that apply.)

Age norms  Grade norms

Stanine  Normal curve equivalents

1. Specify what comprises cluster/composite scores.

1. Describe the tool’s approach to screening, behavior samples, test format, and/or scoring practices, including steps taken to ensure that it is appropriate for use with culturally and linguistically diverse populations and students with disabilities.

1. What types of decision rules does your tool include?

1. Can you provide evidence in support of multiple decision rules?

## Section III: Technical Standards

### Technical Standard 1. Classification Accuracy

1. Please provide the following classification data for your tool disaggregated by grade level, time of year, informant (if applicable), and outcome measure.

Please note, we will only rate information on two different criterion measures. If you provide information from more than two criterion measures, we will report but not rate that additional information.

The cut points for this analysis should be aligned with students needing intensive intervention. If you provide information on cut points not associated with students needing intensive intervention, we will report but not rate that additional information.

In the tables below, enter the classification accuracy data corresponding to the criterion measure (criterion 1 or 2) used and time of year (fall, winter, or spring) your tool was administered. If you do not have data corresponding with the criterion/time of year, leave that table blank. You must submit classification data for at least one criterion and one time of year for each grade level that you are submitting for review. The criterion measure may vary by grade level. Complete a column for each grade level for which you are submitting data.

To learn more about what is expected for each row in the tables below, consult the Academic Screening or Behavior Screening Frequently Asked Questions guide on our [website](https://intensiveintervention.org/about-charts-review-process). Copy columns and tables as needed to present evidence corresponding to all times of year, informants, and grade levels that apply.

**Criterion 1 Fall**

Informant (behavior only): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade |  |  |  |  |  |  |
| Criterion |  |  |  |  |  |  |
| Cut points |  |  |  |  |  |  |
| Base rate children requiring intensive intervention |  |  |  |  |  |  |
| False Positive Rate |  |  |  |  |  |  |
| False Negative Rate |  |  |  |  |  |  |
| Sensitivity |  |  |  |  |  |  |
| Specificity |  |  |  |  |  |  |
| Positive Predictive Power |  |  |  |  |  |  |
| Negative Predictive Power |  |  |  |  |  |  |
| Overall Classification Rate |  |  |  |  |  |  |
| Area Under the Curve (AUC) |  |  |  |  |  |  |
| AUC Estimate’s 95% Confidence Interval: Lower Bound |  |  |  |  |  |  |
| AUC Estimate’s 95% Confidence Interval: Upper Bound |  |  |  |  |  |  |
| Specificity Value at 90% Sensitivity |  |  |  |  |  |  |
| Specificity Value at 80% Sensitivity |  |  |  |  |  |  |
| Specificity Value at 70% Sensitivity |  |  |  |  |  |  |

**Criterion 1 Winter**

Informant (behavior only): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade |  |  |  |  |  |  |
| Criterion |  |  |  |  |  |  |
| Cut points |  |  |  |  |  |  |
| Base rate children requiring intensive intervention |  |  |  |  |  |  |
| False Positive Rate |  |  |  |  |  |  |
| False Negative Rate |  |  |  |  |  |  |
| Sensitivity |  |  |  |  |  |  |
| Specificity |  |  |  |  |  |  |
| Positive Predictive Power |  |  |  |  |  |  |
| Negative Predictive Power |  |  |  |  |  |  |
| Overall Classification Rate |  |  |  |  |  |  |
| Area Under the Curve (AUC) |  |  |  |  |  |  |
| AUC Estimate’s 95% Confidence Interval: Lower Bound |  |  |  |  |  |  |
| AUC Estimate’s 95% Confidence Interval: Upper Bound |  |  |  |  |  |  |
| Specificity Value at 90% Sensitivity |  |  |  |  |  |  |
| Specificity Value at 80% Sensitivity |  |  |  |  |  |  |
| Specificity Value at 70% Sensitivity |  |  |  |  |  |  |

**Criterion 1 Spring**

Informant (behavior only): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade |  |  |  |  |  |  |
| Criterion |  |  |  |  |  |  |
| Cut points |  |  |  |  |  |  |
| Base rate children requiring intensive intervention |  |  |  |  |  |  |
| False Positive Rate |  |  |  |  |  |  |
| False Negative Rate |  |  |  |  |  |  |
| Sensitivity |  |  |  |  |  |  |
| Specificity |  |  |  |  |  |  |
| Positive Predictive Power |  |  |  |  |  |  |
| Negative Predictive Power |  |  |  |  |  |  |
| Overall Classification Rate |  |  |  |  |  |  |
| Area Under the Curve (AUC) |  |  |  |  |  |  |
| AUC Estimate’s 95% Confidence Interval: Lower Bound |  |  |  |  |  |  |
| AUC Estimate’s 95% Confidence Interval: Upper Bound |  |  |  |  |  |  |
| Specificity Value at 90% Sensitivity |  |  |  |  |  |  |
| Specificity Value at 80% Sensitivity |  |  |  |  |  |  |
| Specificity Value at 70% Sensitivity |  |  |  |  |  |  |

**Criterion 2 Fall**

Informant (behavior only): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade |  |  |  |  |  |  |
| Criterion |  |  |  |  |  |  |
| Cut points |  |  |  |  |  |  |
| Base rate children requiring intensive intervention |  |  |  |  |  |  |
| False Positive Rate |  |  |  |  |  |  |
| False Negative Rate |  |  |  |  |  |  |
| Sensitivity |  |  |  |  |  |  |
| Specificity |  |  |  |  |  |  |
| Positive Predictive Power |  |  |  |  |  |  |
| Negative Predictive Power |  |  |  |  |  |  |
| Overall Classification Rate |  |  |  |  |  |  |
| Area Under the Curve (AUC) |  |  |  |  |  |  |
| AUC Estimate’s 95% Confidence Interval: Lower Bound |  |  |  |  |  |  |
| AUC Estimate’s 95% Confidence Interval: Upper Bound |  |  |  |  |  |  |
| Specificity Value at 90% Sensitivity |  |  |  |  |  |  |
| Specificity Value at 80% Sensitivity |  |  |  |  |  |  |
| Specificity Value at 70% Sensitivity |  |  |  |  |  |  |

**Criterion 2 Winter**

Informant (behavior only): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade |  |  |  |  |  |  |
| Criterion |  |  |  |  |  |  |
| Cut points |  |  |  |  |  |  |
| Base rate children requiring intensive intervention |  |  |  |  |  |  |
| False Positive Rate |  |  |  |  |  |  |
| False Negative Rate |  |  |  |  |  |  |
| Sensitivity |  |  |  |  |  |  |
| Specificity |  |  |  |  |  |  |
| Positive Predictive Power |  |  |  |  |  |  |
| Negative Predictive Power |  |  |  |  |  |  |
| Overall Classification Rate |  |  |  |  |  |  |
| Area Under the Curve (AUC) |  |  |  |  |  |  |
| AUC Estimate’s 95% Confidence Interval: Lower Bound |  |  |  |  |  |  |
| AUC Estimate’s 95% Confidence Interval: Upper Bound |  |  |  |  |  |  |
| Specificity Value at 90% Sensitivity |  |  |  |  |  |  |
| Specificity Value at 80% Sensitivity |  |  |  |  |  |  |
| Specificity Value at 70% Sensitivity |  |  |  |  |  |  |

**Criterion 2 Spring**

Informant (behavior only): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade |  |  |  |  |  |  |
| Criterion |  |  |  |  |  |  |
| Cut points |  |  |  |  |  |  |
| Base rate children requiring intensive intervention |  |  |  |  |  |  |
| False Positive Rate |  |  |  |  |  |  |
| False Negative Rate |  |  |  |  |  |  |
| Sensitivity |  |  |  |  |  |  |
| Specificity |  |  |  |  |  |  |
| Positive Predictive Power |  |  |  |  |  |  |
| Negative Predictive Power |  |  |  |  |  |  |
| Overall Classification Rate |  |  |  |  |  |  |
| Area Under the Curve (AUC) |  |  |  |  |  |  |
| AUC Estimate’s 95% Confidence Interval: Lower Bound |  |  |  |  |  |  |
| AUC Estimate’s 95% Confidence Interval: Upper Bound |  |  |  |  |  |  |
| Specificity Value at 90% Sensitivity |  |  |  |  |  |  |
| Specificity Value at 80% Sensitivity |  |  |  |  |  |  |
| Specificity Value at 70% Sensitivity |  |  |  |  |  |  |

1. If you include, in your user’s manual, classification data that are disaggregated by race, ethnicity, or language proficiency, provide these data below. Copy columns and tables as needed.

Subgroup: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Time of year: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Informant (behavior only): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade |  |  |  |  |  |  |
| Criterion |  |  |  |  |  |  |
| Cut points |  |  |  |  |  |  |
| Base rate children requiring intensive intervention |  |  |  |  |  |  |
| False Positive Rate |  |  |  |  |  |  |
| False Negative Rate |  |  |  |  |  |  |
| Sensitivity |  |  |  |  |  |  |
| Specificity |  |  |  |  |  |  |
| Positive Predictive Power |  |  |  |  |  |  |
| Negative Predictive Power |  |  |  |  |  |  |
| Overall Classification Rate |  |  |  |  |  |  |
| Area Under the Curve (AUC) |  |  |  |  |  |  |
| AUC Estimate’s 95% Confidence Interval: Lower Bound |  |  |  |  |  |  |
| AUC Estimate’s 95% Confidence Interval: Upper Bound |  |  |  |  |  |  |
| Specificity Value at 90% Sensitivity |  |  |  |  |  |  |
| Specificity Value at 80% Sensitivity |  |  |  |  |  |  |
| Specificity Value at 70% Sensitivity |  |  |  |  |  |  |

1. Please provide descriptions for the study samples corresponding to each set of classification data reported above (for each criterion and time of year). Copy columns and tables as needed. Cells should indicate the characteristics of the study sample as described in the left-most column. Where applicable, please provide the percentage of the study sample (e.g., demographic characteristics). Complete a column for each grade level for which you are submitting data.

**Criterion 1**

Time of year: \_\_\_\_\_\_\_\_\_\_

Informant (behavior only): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade |  |  |  |  |  |  |
| Criterion |  |  |  |  |  |  |
| National/Local Representation1 |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |
| Sample Size |  |  |  |  |  |  |
| Male |  |  |  |  |  |  |
| Female |  |  |  |  |  |  |
| Gender Unknown |  |  |  |  |  |  |
| Free or Reduced-price Lunch Eligible |  |  |  |  |  |  |
| Other SES Indicator |  |  |  |  |  |  |
| White, Non-Hispanic |  |  |  |  |  |  |
| Black, Non-Hispanic |  |  |  |  |  |  |
| Hispanic |  |  |  |  |  |  |
| American Indian/Alaska Native |  |  |  |  |  |  |
| Other |  |  |  |  |  |  |
| Race/Ethnicity Unknown |  |  |  |  |  |  |
| Disability Classification |  |  |  |  |  |  |
| First Language |  |  |  |  |  |  |
| Language Proficiency Status |  |  |  |  |  |  |

1**National/Local Representation** should be used to report what geographic regions and/or states were represented in the study sample. If participants were drawn from multiple states, indicate whether the population represents a Northeastern region (New England or Middle Atlantic), Midwestern Region (East North Central or West North Central), Southern region (South Atlantic, East South Central, or West South Central), or Western region (Mountain or Pacific). Also indicate which states data were gathered from. For guidance on how to present sample representativeness data, see the Academic Screening or Behavior Screening Frequently Asked Questions guide on our [website](https://intensiveintervention.org/about-charts-review-process).

**Criterion 2**

Time of year: \_\_\_\_\_\_\_\_\_\_

Informant (behavior only): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade |  |  |  |  |  |  |
| Criterion |  |  |  |  |  |  |
| National/Local Representation1 |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |
| Sample Size |  |  |  |  |  |  |
| Male |  |  |  |  |  |  |
| Female |  |  |  |  |  |  |
| Gender Unknown |  |  |  |  |  |  |
| Free or Reduced-price Lunch Eligible |  |  |  |  |  |  |
| Other SES Indicator |  |  |  |  |  |  |
| White, Non-Hispanic |  |  |  |  |  |  |
| Black, Non-Hispanic |  |  |  |  |  |  |
| Hispanic |  |  |  |  |  |  |
| American Indian/Alaska Native |  |  |  |  |  |  |
| Other |  |  |  |  |  |  |
| Race/Ethnicity Unknown |  |  |  |  |  |  |
| Disability Classification |  |  |  |  |  |  |
| First Language |  |  |  |  |  |  |
| Language Proficiency Status |  |  |  |  |  |  |

1**National/Local Representation** should be used to report what geographic regions and/or states were represented in the study sample. If participants were drawn from multiple states, indicate whether the population represents a Northeastern region (New England or Middle Atlantic), Midwestern Region (East North Central or West North Central), Southern region (South Atlantic, East South Central, or West South Central), or Western region (Mountain or Pacific). Also indicate which states data were gathered from. For guidance on how to present sample representativeness data, see the Academic Screening or Behavior Screening Frequently Asked Questions guide on our [website](https://intensiveintervention.org/about-charts-review-process).

1. Describe the criterion (outcome) measure(s) including the degree to which it/they is/are independent from the screening measure.

1. Describe how the classification analyses were performed and cut-points determined. Describe how the cut points align with intensive need.

1. Were the children in the study/studies involved in an intervention in addition to typical classroom instruction between the screening measure and outcome assessment?

Yes

No

If yes, please describe the intervention, what children received the intervention, and how they were chosen.

### Technical Standard 2. Reliability

In the section below, describe the reliability analyses conducted, and provide results. You may report more than one type of reliability (e.g., model-based, internal consistency, inter-rater reliability); however, you must also justify the appropriateness of the method used given the type and purpose of the tool (e.g., inter-rater reliability is provided for tools that require human judgment).

Please ensure that you submit evidence for each informant (for behavior screeners) and grade level targeted by the tool. If you fail to submit data for a targeted grade level, that grade will receive a “dash” rating for this standard.

1. Offer a justification for each type of reliability reported, given the type and purpose of the tool.

1. Describe the sample(s), including size and characteristics, for each reliability analysis conducted.

1. Describe the analysis procedures for each reported type of reliability.

1. In the chart below, report the reliability of performance level score (e.g., internal consistency, inter-rater reliability, model-based).

| Type of Reliability | Informant (Behavior Only) | Grade | n | Coefficient | 95% Confidence Interval\*: Lower Bound | 95% Confidence Interval\*: Upper Bound |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

\*If model-based evidence is being submitted for reliability, note that providing Test Information Function (TIF) / Standard Error (SE) plots to judge the relative precision of the model-based estimate(s) is acceptable in place of providing confidence intervals. Such plots may be provided for aggregate and disaggregated data.

Results for other forms of reliability not conducive to the table format:

If your manual cites other published studies on reliability, provide these citations.

1. If you include, in your user’s manual, reliability data that are disaggregated by subgroup (e.g., race-ethnicity, ELL students), provide these data below.

| Type of Reliability | Informant (Behavior Only) | Subgroup | Grade | n | Coefficient | 95% Confidence Interval: Lower Bound | 95% Confidence Interval: Upper Bound |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

Results for other forms of disaggregated reliability not conducive to the table format:

If your manual cites other published reliability studies, provide these citations.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

### Technical Standard 3. Validity

In the section below, describe the validity analyses conducted, and provide results. You may report more than one type of validity (e.g., concurrent, predictive, evidence based on response processes, evidence based on internal structure, evidence based on relations to other variables, and/or evidence based on consequences of testing), and more than one criterion measure. However, you must justify the choice of analysis and criterion measures given the theoretical assumptions about the relationship between your tool and other, similar constructs.

Please ensure that you submit evidence for each informant (for behavior screeners) and grade level targeted by the tool. If you fail to submit data for a targeted grade level, that grade will receive a “dash” rating for this standard.

1. Describe each criterion measure used and explain why each measure is appropriate, given the type and purpose of the tool. (NOTE: To support validity and generalizability, the TRC expects criterion measures that are ***external to the progress monitoring system***.

1. Describe the sample(s), including size and characteristics, for each validity analysis conducted.

1. Describe the analysis procedures for each reported type of validity.

1. In the chart below, report validity information for the performance level score (e.g., concurrent, predictive, evidence based on response processes, evidence based on internal structure, evidence based on relations to other variables, and/or evidence based on consequences of testing), and the criterion measures.

| Type of Validity | Informant  (Behavior Only) | Grade | Criterion | n | Coefficient | 95% Confidence Interval\*: Lower Bound | 95% Confidence Interval\*: Upper Bound |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

\*If model-based evidence is being submitted for validity, note that providing Test Information Function (TIF) / Standard Error (SE) plots to judge the relative precision of the model-based estimate(s) is acceptable in place of providing confidence intervals. Such plots may be provided for aggregate and disaggregated data.

Results for other forms of validity (e.g., factor analysis) not conducive to the table format:

If your manual cites other published validity studies, provide these citations.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Describe the degree to which the provided data support the validity of the tool.

1. If you include, in your user’s manual, validity data that are disaggregated by diverse populations (e.g., race-ethnicity, ELL students), provide these data below.

| Type of Validity | Informant  (Behavior Only) | Subgroup | Grade | Criterion | n | Coefficient | 95% Confidence Interval: Lower Bound | 95% Confidence Interval: Upper Bound |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

Results for other forms of disaggregated validity (e.g., factor analysis) not conducive to the table format:

If your manual cites other published validity studies, provide these citations.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

### Technical Standard 4. Cross-Validation

1. Has a cross validation study been conducted? *In a cross-validation study, previously identified classification rules are used to predict risk and non-risk in the new sample.*

Yes

No

If yes, in the tables below, enter the cross-validation study data corresponding to the criterion measure (criterion 1 or 2) used and time of year (fall, winter, or spring) your tool was administered. Note that the outcome/criterion measure used for the cross-validation study should be one that was used to supply classification accuracy data. That is, the TRC expects the outcome measure reported in this section to match one of the criterion measures named in the classification accuracy section of the protocol.

To learn more about what is expected for each row in the tables below, consult the Academic Screening or Behavior Screening Frequently Asked Questions guide on our [website](https://intensiveintervention.org/about-charts-review-process). Copy columns and tables as needed to present evidence corresponding to all times of year, informants, and grade levels that apply.

Time of year: \_\_\_\_\_\_\_\_\_\_

Informant (behavior only): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade |  |  |  |  |  |  |
| Criterion |  |  |  |  |  |  |
| Cut points |  |  |  |  |  |  |
| Base rate children requiring intensive intervention |  |  |  |  |  |  |
| False Positive Rate |  |  |  |  |  |  |
| False Negative Rate |  |  |  |  |  |  |
| Sensitivity |  |  |  |  |  |  |
| Specificity |  |  |  |  |  |  |
| Positive Predictive Power |  |  |  |  |  |  |
| Negative Predictive Power |  |  |  |  |  |  |
| Overall Classification Rate |  |  |  |  |  |  |
| Area Under the Curve (AUC) |  |  |  |  |  |  |
| AUC Estimate’s 95% Confidence Interval: Lower Bound |  |  |  |  |  |  |
| AUC Estimate’s 95% Confidence Interval: Upper Bound |  |  |  |  |  |  |
| Specificity Value at 90% Sensitivity |  |  |  |  |  |  |
| Specificity Value at 80% Sensitivity |  |  |  |  |  |  |
| Specificity Value at 70% Sensitivity |  |  |  |  |  |  |

1. Please provide a description of the cross-validation study sample on which the preceding data are based.

Time of year: \_\_\_\_\_\_\_\_\_\_

Informant (behavior only): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade |  |  |  |  |  |  |
| Criterion |  |  |  |  |  |  |
| National/Local Representation1 |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |
| Sample Size |  |  |  |  |  |  |
| Male |  |  |  |  |  |  |
| Female |  |  |  |  |  |  |
| Gender Unknown |  |  |  |  |  |  |
| Free or Reduced-price Lunch Eligible |  |  |  |  |  |  |
| Other SES Indicator |  |  |  |  |  |  |
| White, Non-Hispanic |  |  |  |  |  |  |
| Black, Non-Hispanic |  |  |  |  |  |  |
| Hispanic |  |  |  |  |  |  |
| American Indian/Alaska Native |  |  |  |  |  |  |
| Other |  |  |  |  |  |  |
| Race/Ethnicity Unknown |  |  |  |  |  |  |
| Disability Classification |  |  |  |  |  |  |
| First Language |  |  |  |  |  |  |
| Language Proficiency Status |  |  |  |  |  |  |

1**National/Local Representation** should be used to report what geographic regions and/or states were represented in the study sample. If participants were drawn from multiple states, indicate whether the population represents a Northeastern region (New England or Middle Atlantic), Midwestern Region (East North Central or West North Central), Southern region (South Atlantic, East South Central, or West South Central), or Western region (Mountain or Pacific). Also indicate which states data were gathered from. For guidance on how to present sample representativeness data, see the Academic Screening or Behavior Screening Frequently Asked Questions guide on our [website](https://intensiveintervention.org/about-charts-review-process).

1. Describe the criterion (outcome) measure(s) including the degree to which it or they are independent from the screening measure.

1. Describe how the cross-validation analyses were performed and cut-points determined. Describe how the cut points align with intensive need.

1. Were the children in the cross-validation study involved in an intervention in addition to typical classroom instruction between the screening measure and outcome assessment?

Yes

No

If yes, please describe the intervention, what children received the intervention, and how they were chosen.

### Technical Standard 5. Bias Analysis

1. Have you conducted additional analyses related to the extent to which your tool is or is not biased against subgroups (e.g., race/ethnicity, gender, socioeconomic status, students with disabilities, English language learners)? Examples might include Differential Item Functioning (DIF), or invariance testing in multiple-group confirmatory factor models.

Yes

No

If yes, please provide the following information.

* 1. Describe the method used to determine the presence or absence of bias.

* 1. Describe the subgroups for which bias analyses were conducted.

* 1. Describe the results of the bias analyses conducted, including data and interpretative statements. Include magnitude of effect (if available) if bias is identified.