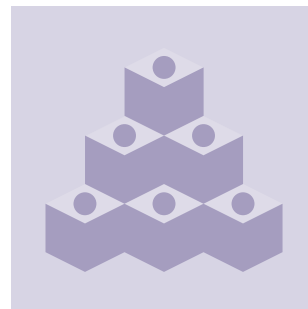


NCRTI Beginning Implementer Training Series: Module 1: Screening Training Manual



April 2011



National Center on Response to Intervention
<http://www.rti4success.org>



U.S. Office of Special
Education Programs

About the National Center on Response to Intervention

Through funding from the U.S. Department of Education's Office of Special Education Programs, the American Institutes for Research and researchers from Vanderbilt University and the University of Kansas have established the National Center on Response to Intervention. The Center provides technical assistance to states and districts and builds the capacity of states to assist districts in implementing proven response to intervention frameworks.



National Center on Response to Intervention

<http://www.rti4success.org>

This document was produced under U.S. Department of Education, Office of Special Education Programs Grant No. H326E070004 to the American Institutes for Research. Grace Zamora Durán and Tina Diamond served as the OSEP project officers. The views expressed herein do not necessarily represent the positions or policies of the Department of Education. No official endorsement by the U.S. Department of Education of any product, commodity, service or enterprise mentioned in this publication is intended or should be inferred. This product is public domain. Authorization to reproduce it in whole or in part is granted. While permission to reprint this publication is not necessary, the citation should be: National Center on Response to Intervention (April 2011). *NCRTI Beginning Implementer Training Series: Module 1 Screening Training Manual*. Washington, DC: U.S. Department of Education, Office of Special Education Programs, National Center on Response to Intervention.



AMERICAN
INSTITUTES
FOR RESEARCH®



**NCRTI Beginning
Implementer Training Series
Module 1: Screening
Training Manual**

Contents

Introduction.....	1
What Is RTI?.....	2
Screening.....	2
Progress Monitoring.....	3
Multi-Level Prevention System.....	3
Data-Based Decision Making.....	4
Understanding Types Of Assessment Within An RTI Framework.....	5
Summative Assessments.....	5
Diagnostic Assessments.....	5
Formative Assessments.....	5
What Is Screening?.....	8
Using Data to Make Decisions.....	9
Identifying At-Risk Students.....	9
Data Analysis and Screening.....	10
Target Identification Rates.....	13
Screening and Specific Learning Disability Eligibility.....	13
Establishing a Screening Process.....	15
Needs, Priorities, and Logistics.....	15
Selecting a Screening Tool.....	17
Frequently Asked Questions.....	18
References.....	21
NCRTI Screening—Glossary of Terms.....	22
Appendix A: Handouts.....	25
Handout 1: Purpose for Screening.....	25
Handout 2: Assessing your Needs, Priorities, and Logistics.....	27
Handout 3: Selecting Screening Tools.....	31

This manual is not designed to replace high-quality, ongoing professional development. It should be used as a supplemental resource to the Module 1: Screening Training PowerPoint Presentation and Handouts. Please contact your state education agency for available training opportunities and technical assistance or contact the National Center on Response to Intervention (www.rti4success.org) for more information.

Introduction

The National Center on Response to Intervention (NCRTI) developed three training modules for beginning implementers of response to intervention (RTI). These modules are intended to provide foundational knowledge about the essential components of RTI and to build an understanding about the importance of RTI implementation. The modules were designed to be delivered over the course of a year and in the following sequence: Screening, Progress Monitoring, and Multi-Level Prevention System. RTI implementers should plan to follow-up with additional intensive training for advanced implementers.

This training is intended for teams in initial planning or implementation of a districtwide RTI framework, as well as for those districts transitioning from traditional prereferral models to a district model of prevention. The training provides district teams an overview of the essential components of RTI, opportunities to analyze school and district RTI data, activities so that they can apply new knowledge, and team planning time.

Module 1: Screening

In this module, participants will become familiar with the essential components of an RTI framework: screening, progress monitoring, multi-level prevention system, and data-based decision making. Participants will gain the necessary skills in order to use screening data to identify students at risk, to conduct basic data analysis using screening data, and to establish an effective screening process.

Module 2: Progress Monitoring

In this module, participants will apply previously learned skills to use screening data to evaluate the general effectiveness of the RTI framework and to establish progress monitoring and intervention schedules. Participants will gain the necessary skills to use progress-monitoring data to evaluate and make decisions about instruction, to establish data decision rules such as goal setting, and to establish an effective progress-monitoring system.

Module 3: Multi-Level Prevention System

In this module, participants will review how screening and progress-monitoring data can assist in decisions at all levels, including school, grade, class, and student. Participants will gain skills to select evidence-based practice, to make decisions about movement between levels of prevention, and to establish a multi-level prevention system.

The Beginning Implementer Series should be delivered by a trained, knowledgeable professional. This training series is designed to be a component of comprehensive professional development that includes supplemental coaching and ongoing support. Each training module includes the following training materials:

- PowerPoint Presentations
- Speaker's Notes
- Handouts
- Videos
- Training Manual

What Is RTI?

NCRTI offers a definition of response to intervention that reflects what is currently known from research and evidence-based practice.

Response to intervention (RTI) integrates assessment and intervention within a school-wide, multi-level prevention system to maximize student achievement and reduce behavior problems. With RTI, schools identify students at-risk for poor learning outcomes, monitor student progress, provide evidence-based interventions, and adjust the intensity and nature of those interventions based on a student's responsiveness. RTI may be used as part of the determination process for identifying students with specific learning disabilities or other disabilities. (National Center on Response to Intervention, 2010)

NCRTI believes that rigorous implementation of RTI includes a combination of high-quality, culturally and linguistically responsive instruction, assessment, and evidence-based intervention. Further, NCRTI believes that comprehensive RTI implementation will contribute to more meaningful identification of learning and behavioral problems, improve instructional quality, provide all students with the best opportunities to succeed in school, and assist in identifying learning disabilities and other disabilities.

This document and training are based on NCRTI's four essential components of RTI:

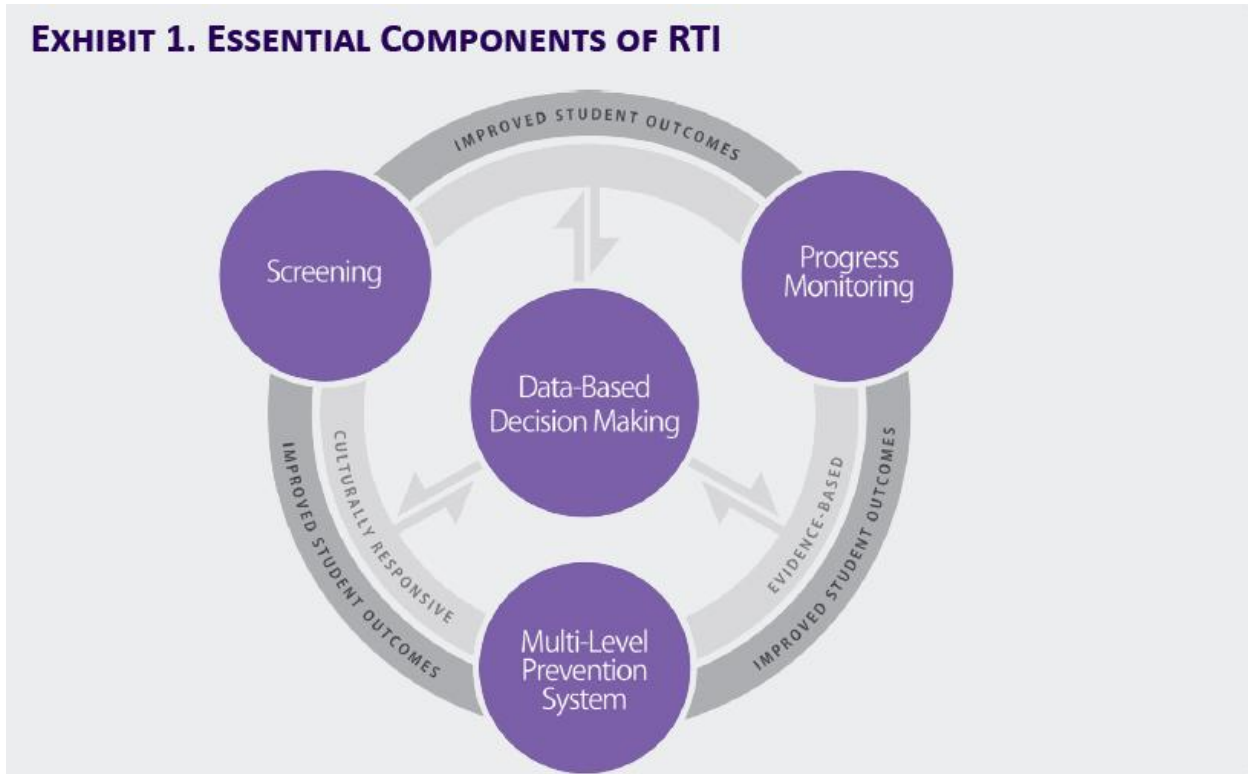
- A schoolwide, multi-level instructional and behavioral system for preventing school failure
- Screening
- Progress monitoring
- Data-based decision making for instruction, movement within the multi-level system, and disability identification (in accordance with state law)

Exhibit 1 represents the relationships among the essential components of RTI. Data-based decision making is the essence of good RTI practice; it forms the foundation of the other three components. All components must be implemented using culturally responsive and evidence-based practices.

Screening

Struggling students are identified by implementing a two-stage screening process. The first stage, universal screening, is a brief assessment of all students, conducted at the beginning of the school year, although some schools and districts use it two to three times throughout the school year. For students who score below the cut score on the universal screen, a second stage of screening is conducted to more accurately predict which students are truly at risk for poor learning outcomes. This second stage involves additional, more in-depth testing or short-term progress monitoring to confirm a student's at-risk status. Screening tools must be reliable and valid and demonstrate diagnostic accuracy for predicting which students will develop learning or behavioral difficulties.

EXHIBIT 1. ESSENTIAL COMPONENTS OF RTI



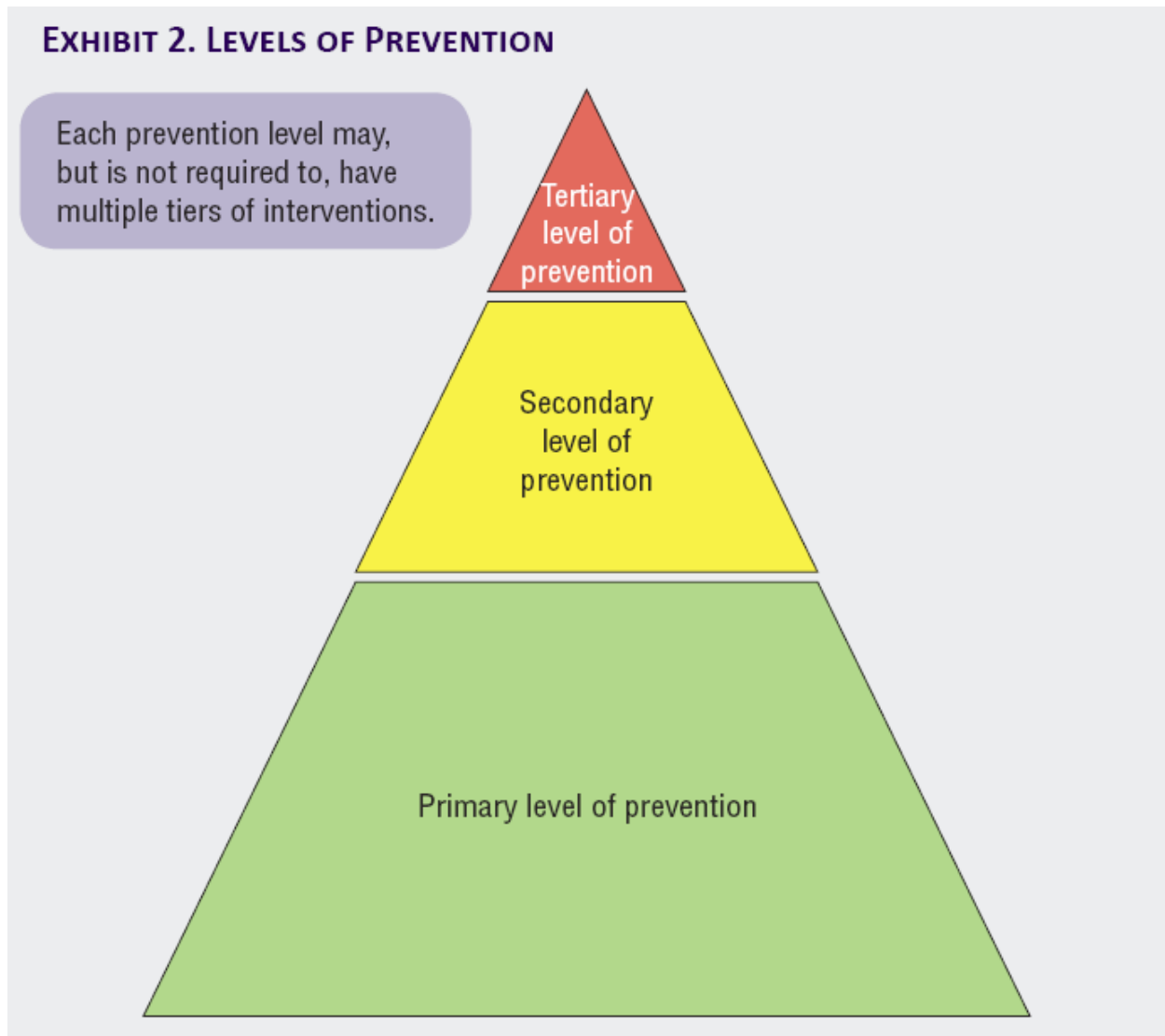
Progress Monitoring

Progress monitoring is used to assess students' performance over time, quantify student rates of improvement or responsiveness to instruction, and evaluate instructional effectiveness. For the students least responsive to effective instruction, progress monitoring is used to formulate effective individualized programs. Progress monitoring tools must accurately represent students' academic development and must be useful for instructional planning and assessing student learning. In addition, in tertiary prevention, educators use progress monitoring to compare a student's expected and actual rates of learning. If a student is not achieving the expected rate of learning, the educator experiments with instructional components in an attempt to improve the rate of learning.

Multi-Level Prevention System

Classroom instructors are encouraged to use research-based curricula in all subjects. When a student is identified via screening as requiring additional intervention, evidence-based interventions of moderate intensity are provided. These interventions, which are in addition to the core primary instruction, typically involve small-group instruction to address specific identified problems. These evidence-based interventions are well defined in terms of duration, frequency, and length of sessions, and the intervention is conducted as it was in the research studies. Students who respond adequately to secondary prevention return to primary prevention (the core curriculum) with ongoing progress monitoring. Students who show minimal response to secondary prevention move to tertiary prevention, where more intensive and individualized supports are provided. All instructional and behavioral interventions

should be selected with attention to their evidence of effectiveness and with sensitivity to culturally and linguistically diverse students. Exhibit 2 shows the three prevention levels.



Data-Based Decision Making

Screening and progress monitoring data can be aggregated and used to compare and contrast the adequacy of the core curriculum as well as the effectiveness of different instructional and behavioral strategies for various groups of students within a school. For example, if 60 percent of the students in a particular grade score below the cut score on a screening test at the beginning of the year, school personnel might consider the appropriateness of the core curriculum or whether differentiated learning activities need to be added to better meet the needs of the students.

To learn more about the essential components of RTI, read the Essential Components of RTI—A Closer Look at Response to Intervention, available through NCRTI (www.rti4success.org).

Understanding Types Of Assessment Within An RTI Framework

The following table describes the three types of assessments used in an RTI framework.

Type	When?	Why?
Summative	After instruction	Assessment of learning
Diagnostic	Before instruction	Identify skill strengths and weaknesses
Formative	During instruction	Assessment for learning

Summative Assessments

Summative assessments measure what students learned over a period of time. They are typically administered after instruction and can help to determine what to teach but not how to teach. Examples of summative assessments include end-of-chapter tests or final exams, high-stakes tests, and the GRE, SAT, and ACT. These assessments are typically used for accountability, resource allocation, and measures of skill mastery. Summative assessments are often time consuming and are not valid for making decisions about individual students.

Diagnostic Assessments

Diagnostic assessments are measures of a student's current knowledge and skills and can be used to identify a suitable program of learning. They are administered before instruction occurs to assist in identifying appropriate instruction and interventions. Examples of diagnostic assessments include the Qualitative Reading Inventory, Diagnostic Reading Assessment, and Key Math. These tests typically require extensive time to administer and are recommended only for some students. Because diagnostic assessments provide detailed information about individual student learning, they are most effective in understanding the needs of specific students rather than all students.

Formative Assessments

Formative assessments are administered during instruction and measure how well students are responding to instruction. They are a form of evaluation used to plan instruction in a recursive way. With formative assessment, student progress is systematically assessed to provide continuous feedback to both the student and the teacher concerning learning successes and failures. Formative assessments can be used to identify students who are not responsive to instruction or interventions (screening) and to understand rates of student improvement (progress monitoring). They can also be used to make curriculum and instructional decisions, evaluate program effectiveness, proactively allocate resources, and compare the efficacy of instruction and interventions. These formal and informal assessments are generally brief measures of direct student performance. Informal assessments are not data driven but rather content and performance driven. Examples of informal assessments are observations or teacher-made assessments. Formal assessments provide data to support the conclusions made from the tests. These types of tests are typically referred to as standardized measures. The data for these formal assessments are mathematically computed and summarized. Scores such as percentiles, stanines, or standard scores are most commonly given from this type of assessment.

Screening and progress monitoring tools in an RTI framework are typically formal, formative assessments. Some examples are AIMSweb–R-CBM, Dynamic Indicators of Basic Early Literacy Skills (DIBELS), and iSTEEP–Oral Reading Fluency. For more examples, visit the NCRTI progress monitoring and screening tools charts (www.rti4success.org).

There are two common types of formative assessments: mastery measures and general outcome measures.

Mastery Measures

Mastery measures determine the mastery of a series of short-term instructional objectives. For example, a student may master multidigit addition and then master multidigit subtraction. To use mastery measures, teachers must determine a sensible instructional sequence and design criterion-referenced testing procedures to match each step in that instructional sequence. Teacher-made tests present concerns given the unknown reliability and validity of these measures. The hierarchy of skills used in mastery measurement is logical, not empirical. This statement means that although it may seem logical to teach addition first and then subtraction second, there is no evidence base for the sequence.

Exhibit 3 provides an example of progress monitoring data from mastery measures in multidigit addition and subtraction. Because mastery measures are based on mastering one skill before moving on to the next skill, the assessment does not reflect maintenance or generalization. It becomes impossible to know whether, after one skill has been taught, the student still remembers how to perform the previously learned skill. In addition, how a student does on a mastery measure assessment does not indicate how he or she will do on standardized tests, because the number of objectives mastered does not relate well to performance on criterion measures.

General Outcome Measures

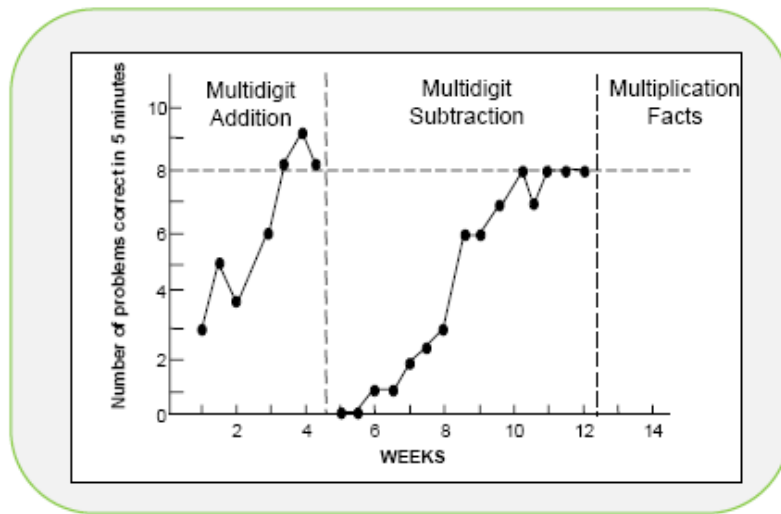
General outcome measures (GOMs) do not have the limitations of mastery measures. They are indicators of general skill success and reflect overall competence in the annual curriculum. They describe students' growth and development over time, or both their "current status" and their "rate of development." Common characteristics of GOMs are that they are simple and efficient, are sensitive to improvement, provide performance data to guide and inform a variety of educational decisions, and provide national or local norms that allow for cross-comparisons of data.

One example of a GOM is curriculum-based measurement (CBM). CBM is an approach to measurement that is used to screen students or to monitor student progress in mathematics, reading, writing, and spelling. With CBM, teachers and schools can assess individual responsiveness to instruction. When a student proves unresponsive to the instructional program, CBM signals the teacher or school to revise that program. Each CBM test is an alternate form of equivalent difficulty. Each test samples the yearlong curriculum in exactly the same way using prescriptive methods for constructing the tests. In fact, CBM is usually conducted with "generic" tests, designed to mirror popular curricula. CBM is highly prescriptive and standardized, which increases the

reliability and validity of scores. It provides teachers with a standardized set of materials that has been researched to produce meaningful and accurate information. CBM makes no assumptions about instructional hierarchy for determining measurement. In other words, CBM fits with any instructional approach. Also, CBM incorporates automatic tests of retention and generalization. Therefore, the teacher is constantly able to assess whether the student is retaining what was taught earlier in the year. Exhibit 4 (on p. 9) provides an example of graphed CBM data. Unlike mastery measures, CBM data allow measurement of growth over time.

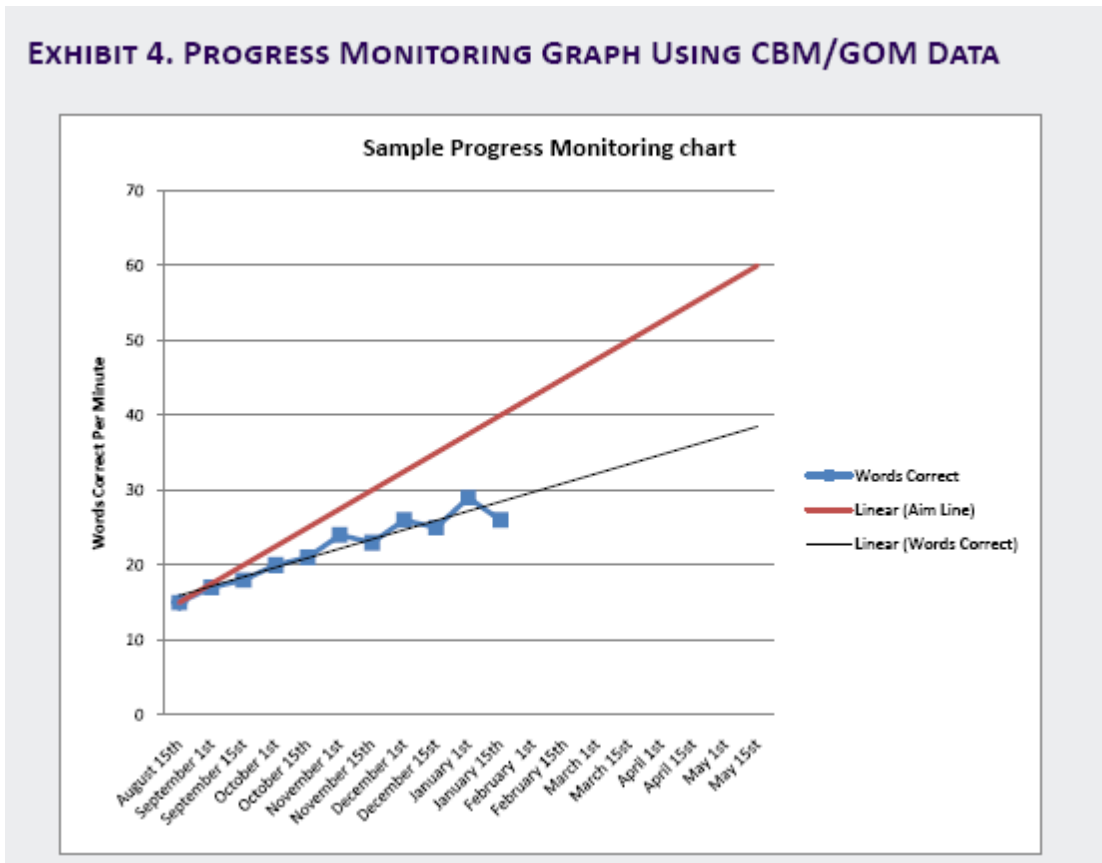
EXHIBIT 3. MASTERY MEASURE MULTIDIGIT ADDITION AND SUBTRACTION

Multidigit Subtraction Mastery Test



What Is Screening?

The purpose of screening is to identify those students who are at risk for poor learning outcomes. The focus of universal screening is on all students, not just those who teachers believe are at risk. Students may still slip through the cracks without an unbiased, systematic process for screening. Remember, screening is not a diagnostic test. It is a brief, reliable, and valid formative assessment used to identify which students may need additional assessments (such as progress monitoring or diagnostic assessments) or additional instructional support. Screening tools should demonstrate diagnostic accuracy for predicting learning or behavioral outcomes. In other words, they should be able to accurately identify at-risk students to the greatest extent possible.



At a minimum, screening should occur at least twice per year (once in the beginning and one in the middle of the school year). However, NCRTI highly recommends conducting screening at least three times a year (fall, winter, and spring) in order to evaluate program effectiveness, establish local norms and cut scores, and provide data for next year’s teachers.

Using Data to Make Decisions

Screening data can assist with data-based decision making at all levels of instruction. Using screening data for all students, not just those who have demonstrated learning difficulties, allows identification of students who might be at risk for poor learning outcomes in the future. Screening data provide an objective measure of a student's skills and can provide evidence of appropriateness of instruction as part of the specific learning disability process. For example, if the majority of the class is successful and the individual student is not, it may suggest that the student is at risk because the overall instruction appears effective for most. If all students in the class are struggling, it may suggest that the general instruction or curriculum might be ineffective.

District-level screening data can provide evidence about whether the core curriculum is effective for most students across schools and grade levels. The data can help to assess the effectiveness of the district's RTI model, assess the effectiveness of the implementation of the model, and inform decisions about innovation and sustainability. Data can be used to ensure that resources are equitably allocated for services and supports across schools. By using screening data to inform decisions, districts can model data-based decision making and increase the buy-in for using data by schools and teachers.

School- and grade-level screening data are essential for instructional decision making at the primary and secondary levels. Data can provide evidence of the effectiveness of instruction and curriculum and the areas of need. School-level screening data can be used to inform and set measurable school improvement goals, and grade-level data can help to identify students who might need additional instruction or assessment.

Identifying At-Risk Students

One of the primary goals of screening is accurately identifying students who are at risk for poor learning outcomes. Unfortunately, the classification of students as at risk or not at risk for academic and behavioral problems is often based on an arbitrary cut score. A cut score is a score on a screening test that separates students who are considered potentially at risk from those considered not at risk. Setting cut scores allows schools to identify an initial pool of students who may require interventions or additional assessment. Using a cut score results in four possible outcomes for identifying at-risk students (Exhibit 5).

- *True positives (TPs)*, or students whom the screening identifies as at risk and who are actually at risk.
- *True negatives (TNs)*, or students whom the screening identifies as not at risk and who are actually not at risk.
- *False positives (FPs)*, or students who are identified as at risk by the screening tool but are actually not at risk.
- *False negatives (FNs)*, or students who are not identified as at risk through the screening tool but are actually at risk.

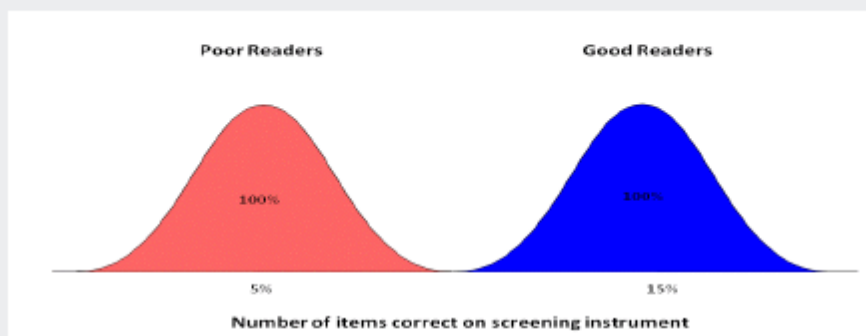
EXHIBIT 5. CLINICAL DECISION MAKING MODEL

		Outcome	
		At-risk	Not at-risk
Screen	At-risk	True Positive	False Positive
	Not at-risk	False Negative	True Negative

From these classifications, a base rate and overall accuracy of the screening instrument and cut score can be calculated. The *base rate* is the proportion of students in the sample who are at risk for poor learning outcomes. *Overall accuracy* is the proportion of true positives and true negatives from the entire sample. Other important pieces of information regarding how well the screener classifies students are sensitivity and specificity. *Sensitivity* ($TP/TP+FN$) is the proportion of students who are at risk and are correctly identified at risk. *Specificity* ($TN/FP+TN$) is the proportion of students not at risk who are correctly identified as not at risk.

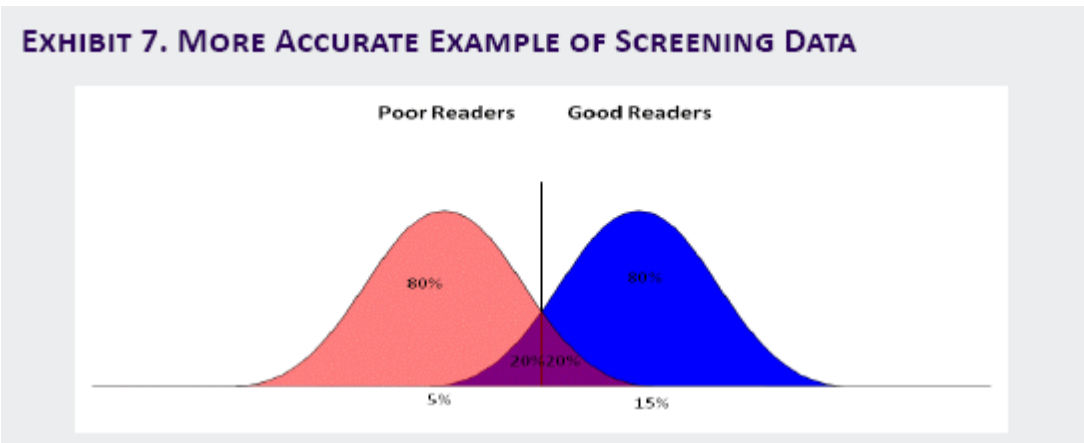
Perfect screening would result in 100 percent accurate identification of students who need additional support (true positives) and those who don't need additional support (true negatives). Exhibit 6 represents the ideal screening data representation. In this case, the screener tool would accurately identify students who did and did not need assistance.

EXHIBIT 6. IDEAL SCREENING DATA REPRESENTATION

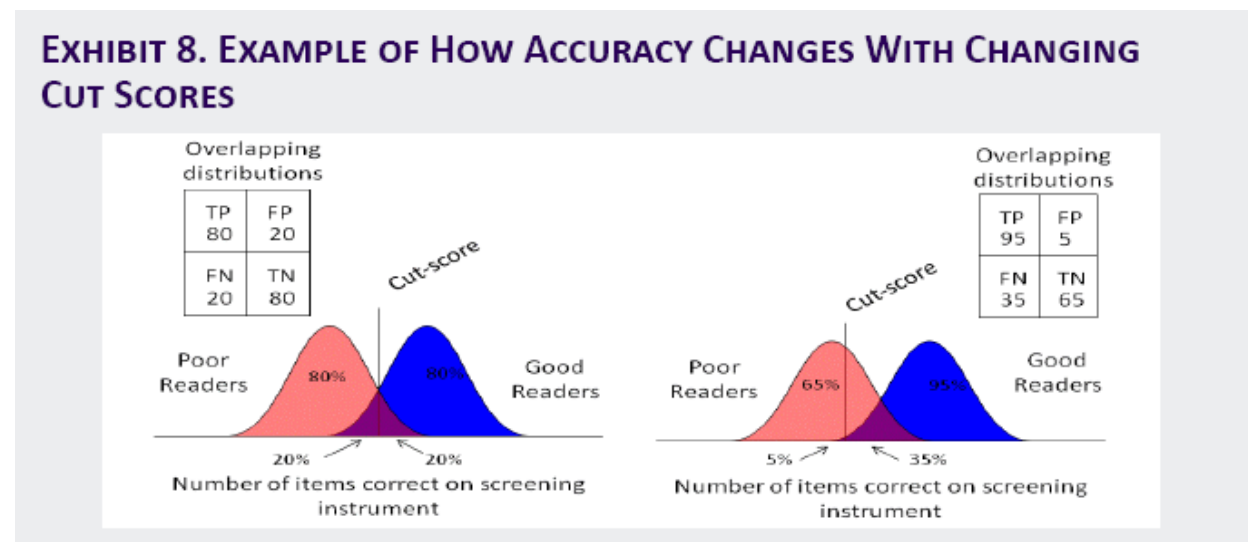


Unfortunately, no screening tool is ideal because all screening tools produce overlapping distributions of good and poor readers. Exhibit 7 shows how some poor readers may score well,

and some good readers may score poorly. Other variables, including the test itself, may impact the accuracy of the results.



Overlapping distributions result in false positive and false negative classifications. Regardless of the type of cut score, if the cut score is changed, the number of students accurately identified or inaccurately identified will also change. Exhibit 8 shows two different classification outcomes for two different cut scores. In both cases, at-risk students were underidentified and overidentified, but the proportion of each differed. Screening tools tend to overidentify students and thus should be followed with progress monitoring to verify the results.



Particular attention is given to the accuracy of screening instruments because errors in identification (overidentification and underidentification) can be costly. In the health care field, overidentification could result in the expense of additional testing plus unnecessary worry. Conversely, underidentification could result in missing serious health problems. In education, overidentification could result in the expense of additional testing and early intervention services. Underidentification is costly to the extent that students miss opportunities for prevention and early intervention.

Establishing Benchmarks and Cut Scores

A benchmark is a predetermined level of performance on a screening test that is considered representative of proficiency or mastery of a certain set of skills. Benchmarks or growth rates indicate when particular skills should be achieved and help to classify students as low, moderate, or high risk. On the basis of the benchmarks that have been set, specific cut scores should be established to separate students who are likely to reach proficiency (not at risk) from those who will need additional support in order to reach proficiency (at risk).

Using consistent benchmarks across schools within a district or state allows for comparisons across schools. When schools develop individualized cut scores, it is difficult to make comparisons across sites. This can complicate resource allocation, data reporting, and making accurate data-based decisions.

Data Analysis and Screening

Data analysis and the subsequent use of that data to inform decisions are important to the entire RTI process. Establishing routines for conducting data analysis and reviewing data at logical and predetermined intervals can improve overall school performance. Clear procedures for analysis and decisions should be established for all levels of instruction, beginning with district-level decisions and working through school-, grade-, and class-level analysis. Explicit decision rules should be set for assessing student progress and classifying students in need of additional support. By using established decision rules and data at all levels, teams can identify trends (positive and negative) and brainstorm why certain trends might be apparent.

Districts and schools must establish a process for examining screening data. This process includes analyzing causes for nonresponse to primary instruction, developing supplemental interventions, and assessing whether students are responding to those interventions. The process of decision making is the same regardless of whether one is examining groups of students or an individual student. More efficient use of time and resources is found when the process is used to benefit groups of children. The RTI team members will have various roles in this process. This collaborative learning cycle results in effective curriculum decisions, scheduling of instruction, student grouping, and allocation of resources.

Norm-Referenced Assessment

Norm-referenced assessment compares a student's performance with that of an appropriate peer group. When using a norm-referenced measure, a student is measured against those undertaking the test, not against any defined criteria. This measurement permits a fixed proportion of students to pass and fail. Because there are differences in the students taking the test from year to year, the standards that are set vary.

Criterion-Referenced Assessment

Criterion-referenced assessment measures what a student understands, knows, or can accomplish in relation to a specific performance objective or criterion. It is typically used to identify a student's specific strengths and weaknesses in relation to an age- or grade-level standard. It does not compare students with other students. Because the criteria

typically do not vary from year to year, the standards do not change. There are multiple ways to determine the criteria that are used.

Target Identification Rates

Target identification rates assist districts and schools in identifying how resources and services can be allocated to address the needs of their at-risk population. It establishes target scores that identify the proportion of students who may need secondary and tertiary instruction. This proportion may be dependent on the program's objectives and resources and may not reflect the total at-risk population. For example, if the majority of the students are below the cut score, it may not be financially feasible to serve all of the students needing secondary or tertiary prevention. In Exhibit 9, School 1 has resources available to serve 20 percent of the students in secondary or tertiary instruction. In contrast, School 2 only has enough resources available to serve 15 percent of students in secondary and tertiary instruction. Regardless, if more than 20 percent of the student population is identified as at-risk, the focus should be on improving core curriculum and instruction.

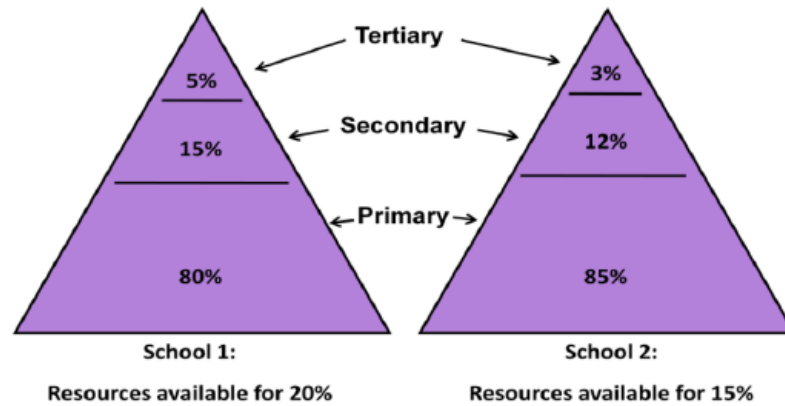
Unique target identification rates may be specified for different skill areas. For example, a school may have a larger target identification rate for word reading than for comprehension because of resource availability.

Screening and Specific Learning Disability Eligibility

To ensure that underachievement in a child suspected of having a specific learning disability (SLD) is not due to lack of appropriate instruction in reading or math, the group must consider the following, as part of the evaluation described in the Individuals With Disabilities Education Improvement Act of 2004:

- Data that demonstrate that prior to, or as a part of, the referral process, the child was provided appropriate instruction in regular education settings, delivered by qualified personnel.
- Data-based documentation of repeated assessments of achievement at reasonable intervals, reflecting formal assessment of student progress during instruction, which was provided to the child's parents.

EXHIBIT 9. COMPARISON OF DIFFERENT TARGET IDENTIFICATION RATES IN TWO SCHOOLS



Screening data that portray the growth rate of all students can provide data “that demonstrate that prior to, or as a part of, the referral process, the child was provided appropriate instruction in regular education settings.” Together with documentation of the duration and nature of the instruction, screening results can demonstrate the effectiveness, or “appropriateness,” of the instruction for this student in comparison with his or her peers. Appropriate instruction is often viewed as instruction that provides benefit to the majority of students. Progress monitoring that tracks student progress on a regular basis and is shared with the student’s family can help to support the second point, “data-based documentation of repeated assessments of achievement at reasonable intervals.”

Establishing a Screening Process

Establishing a screening process begins with identifying the needs and resources of the district or school and then selecting a screening tool that matches those needs and resources. Before tool selection, teams must consider why screening is being conducted, what they hope to learn from the screening data, and how the results will be used. Conducting an assessment of needs, priorities, and logistics is a logical first step. The NCRTI screening tools chart (www.rti4success.org) provides practitioners with publisher-created summaries that may assist districts and schools in identifying tools that match their needs and resources.

Needs, Priorities, and Logistics

Districts and schools should consider the following when establishing a screening system: the desired outcome, the timing and schedule of screening, and the role of staff members. Schools and districts also must consider the logistics necessary for implementing screening, such as what is needed for administration and scoring, how much training is needed to implement screening with fidelity, and what resources are available to support screening implementation. Schools and districts should accurately identify their needs but might be unable to address all of them because of lack of resources.

Outcome Measures

Districts and schools should identify what outcome measures(s) are the focus of the prevention model. Screening tools are selected on the basis of their ability to predict success on these outcome measures. Outcomes are not limited to reading and math and may include measures of mental and physical health, speech and language, behavior, graduation, or postschool outcomes. Schools and districts may want to measure multiple outcomes for their students. In this case, it is necessary to identify different screeners to assess different outcomes. In selecting outcome measures, districts and schools should consider how the outcome of interest maps to the curriculum and state standards. Schools must choose age-appropriate outcome measures that capture student ability.

Timing

The timing of screening is critical as children are developing the very skills schools and districts are interested in measuring. In effect, schools are trying to measure a “moving target” (Speece, 2005). Therefore, how the screening is timed with this development can make a big difference in its accuracy. For example, related to reading, “To have good classification accuracy, screens must target reading or reading-related skills that are pertinent to the grade and time the screen is administered” (Jenkins, Hudson, & Johnson, 2007, p. 585). In kindergarten, relevant skills could include phonemic awareness, letter and sound knowledge, and vocabulary. In first grade, phonemic spelling, decoding, word identification, and text reading are important skills to assess (Compton, Fuchs, Fuchs, & Bryant, 2006). In second and third grades, measures should assess number and type of words students can read and comprehend and the fluency of those skills. In higher grades, comprehension of more difficult texts is an important relevant reading measure.

Schools and districts must also consider how frequently they will screen students. To ensure that screening data can provide an accurate representation of a student's knowledge level, NCRTI recommends that screening be conducted at least three times during the year (fall, winter, spring). This frequency provides sufficient data for evaluating program effectiveness, establishing local norms and cut scores, and providing data to the following year's teacher.

Although screening data are informative, time spent taking and scoring assessments displaces time available for instruction for both teachers and students. To limit time wasted during screening, schools and districts must consider the most effective and efficient manner to conduct screening. The time demanded for screening can vary by type of screener. Classwide screeners may take 3–60 minutes to administer, whereas individual screeners typically take 1–2 minutes per student. The length of the screening will depend on type of assessment and instructional domain. Schools and districts should set aside sufficient time for test administration, data analysis, and professional development.

Staff Roles

Trained staff are essential to an effective screening process. Staff administer and score screening assessments, analyze screening data, and make decisions based on the data. Schools and districts must identify who will be involved in each stage of the screening process. This process might include considering whether the teacher, paraeducator, or an assessment team will conduct the screening and who will be involved with the data team. In considering staff, it is also important to consider their knowledge and abilities. For example, are the people participating in the data team knowledgeable about using data to make decisions?

Administration

Different types of screening assessments may demand different types of materials. In making decisions about tool selection, schools and districts must consider how the tool is administered. Some assessments are paper-and-pencil assessments, whereas others are computer based. Paper-and-pencil assessments often require printing or the purchasing of new materials each year. Schools and districts must decide whether it is feasible to select a computer-based program, given their current level of access to computers. It might not be wise to purchase a computer-based screening tool if the computers are on loan for a short time. Regardless of the decision to use paper-and-pencil or computers, districts and schools should consider the long-term feasibility of supporting the implementation of the tool. Teams should also consider the data management needs in addition to the tool administration. Some screening tools include data analysis and reporting features, whereas others may demand additional statistical programs and data warehouses to track and analyze the data.

Training

Training is required to help ensure the fidelity of implementation. Before selecting a screening tool and screening process, one must consider what training resources are necessary to build the capacity of relevant staff. A number of forms of training can occur, such as use of field-tested training manuals (typically provided by the tool

developers), professional development activities conducted in person or over the Web, and ongoing technical assistance support. Publishers often provide a recommended training schedule. Administrators should ensure that the publisher-recommended professional development matches the resources of the district or school before purchasing any tool.

Funding

A number of costs are associated with screening, including the cost of the tool and any additional materials, training, and instruction for students identified by the screening assessment. The costs of screening tools vary, but they typically are \$1–5 per student. Some screening measures also have additional system costs, especially computer-based tools. Another significant cost related to screening is the cost of training staff to administer screening tools and to analyze and use the data appropriately.

Selecting a Screening Tool

NCRTI has developed a screening tools chart that provides relevant information for selecting tools. Each year NCRTI has a call for tool developers to submit their tools for review. A technical review committee made up of experts in the field reviews the tools for technical rigor. The NCRTI screening tools chart is not an exhaustive list of all available screening measures as vendors or tool developers must submit their tools in order for it to be reviewed. One can learn more about the tools available on the screening tools chart by visiting www.rti4success.org. The tools chart provides information on a measure's technical rigor, efficiency of use, implementation requirements, and supporting data. One can learn about the different information that the tools chart provides and the suggested steps for review by viewing the user guide.

Once a tool is selected, districts and schools need to continuously evaluate whether the screening tool matches their needs and resources and provides the data needed to inform their decisions.

Frequently Asked Questions

What is at the heart of RTI?

The purpose of RTI is to provide all students with the best opportunities to succeed in school, identify students with learning or behavioral problems, and ensure that they receive appropriate instruction and related supports. The goals of RTI are as follows:

- Integrate all the resources to minimize risk for the long-term negative consequences associated with poor learning or behavioral outcomes.
- Strengthen the process of appropriate disability identification.

What impact does RTI have on students who are not struggling?

An important component of an effective RTI framework is the quality of the primary prevention level (i.e., the core curriculum), where all students receive high-quality instruction that is culturally and linguistically responsive and aligned with a state's achievement standards. This framework allows teachers and parents to be confident that a student's need for more intensive intervention or referral for special education evaluation is not due to ineffective classroom instruction. In a well-designed RTI system, primary prevention should be effective and sufficient for about 80 percent of the student population.

Will the RTI process delay identification?

The RTI process takes longer than a traditional one-step comprehensive evaluation. However, beginning at the secondary level of prevention, students are receiving services designed to remediate their learning problems. The hope is that the prevention built into RTI will reduce false-positive identification (i.e., students incorrectly identified as having disabilities because they have not received strong instruction) and help many students get on a trajectory toward successful academic outcomes. Also, RTI facilitates prevention and identification early in the primary grades (in contrast to the traditional IQ-achievement discrepancy, which often requires years of schooling before a sizeable discrepancy can accrue).

Does each child have to go through RTI, or can a child receive a traditional assessment?

All students are screened in the RTI model. However, schools honor parent requests for a traditional one-step comprehensive evaluation in lieu of the RTI process.

Who initiates the RTI process?

Typically, students are identified to participate in the secondary level of prevention on the basis of their universal screening scores. Many times, such universal screening is supplemented with short-term progress monitoring (e.g., 6-10 weeks) to determine the student's response to general education.

What proportion of students is likely to be identified as at risk?

The proportion of students identified for different steps in the RTI process depends largely on the quality of general education and available funds. When general education instruction is of questionable quality, research suggests that 20–25 percent of a school population is likely to be identified as at risk and demonstrate unresponsiveness to the core curriculum. Of course, providing the secondary level of prevention to 25 percent of a school population creates resource challenges. On the other hand, research also suggests that with high-quality general education, only 9–10 percent of students will be identified as at risk and respond inadequately to the core curriculum, with approximately half those students responding to high-quality secondary interventions. Clearly, it is important to ensure high-quality general education. In a similar way, integrity of the RTI process requires a strong secondary level of prevention.

What is the difference between screening and diagnostic assessments?

Screening tools are administered to all students at least twice during the school year, with the goal of identifying at-risk students, whereas a diagnostic is generally administered to some students once, with the goal of identifying specific deficits in student learning and planning an intervention. Screening is a type of assessment that is characterized by providing quick, low-cost, repeatable testing of age-appropriate critical skills (e.g., identifying letters of the alphabet or reading a list of high-frequency words) or behaviors (e.g., tardiness, aggression, or hyperactivity). In the RTI model, screening is used to designate students who might be in need of closer monitoring in their general education curriculum or of a more intense intervention. Information on how to select a screening tool can be found on NCRTI’s screening tools chart (www.rti4success.org).

How does one pick a good screening tool?

To select a tool, the leadership team should discuss the needs of the school or district and evaluate available options. When selecting a screening tool, the team should select one that targets skills pertinent to the grade and time the screen is administered. It is also important to consider the tool’s accuracy, validity, cost, and the technology needed to support the tool. NCRTI created a screening tools chart (www.rti4success.org) to assist the leadership team in evaluating tools and recommends a six-step process for using it: (1) gather a team, (2) determine your needs, (3) determine your priorities, (4) familiarize yourself with the content and language of the chart, (5) review the ratings and implementation data, and (6) ask for more information.

How does one know whom to progress monitor and screen?

All students should be screened in an RTI framework to identify who may be at risk for poor learning outcomes. It is impossible for screening tools to predict with 100 percent accuracy which students will need additional support. Thus, screening tools tend to overidentify so as not to underidentify. Because of this overidentification, schools may consider conducting additional assessments, such as progress monitoring, to determine whether students were inappropriately screened for additional support. Progress monitoring should also be conducted for all students receiving additional interventions. Screening tools and progress monitoring tools depend on cut scores to determine who needs additional assessment and support. NCRTI defines a cut score as a score on the scale of a screening tool or a progress-monitoring tool. For universal screeners, educators use the cut score to determine whether to provide additional intervention. For progress-monitoring tools, educators use the cut score to determine whether

the student has demonstrated adequate response, whether to make an instructional change, and whether to move the student to more or less intensive services.

How do screening tools align with state assessments?

Although they may be in the same content area, screening tools and state assessments assess different skills and knowledge. Screening tools often assess access skills, or those skills needed to access the content assessed on the state test. For example, a screening tool may assess a student's ability to read connected text, whereas a state assessment assesses a student's ability to use that skill to comprehend a novel passage. Screening tools assess indicators of reading through brief assessments. Many screening tools have been correlated to outcomes on state tests. Schools and districts can contact the publisher of a screener to find out whether it has been correlated with their state test and whether cut scores have been established.

References

- Burns, M. K., Appleton, J. J., & Stehouwer, J. D. (2005). Meta-analysis of response-to-intervention research: Examining field-based and research-implemented models. *Journal of Psychoeducational Assessment, 23*, 381–394.
- Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant, J. D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology, 98*, 394–409.
- Dexter, D. D., Hughes, C. A., & Farmer, T. W. (2008). Responsiveness to Intervention: A Review of Field Studies and Implications for Rural Special Education. *Rural Special Education Quarterly, 27*(4), 3–9.
- Individuals With Disabilities Education Improvement Act, 34 C.F.R. §§ 300.307, 300.309, 300.311 (2004).
- Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for at-risk readers in a response to intervention framework. *School Psychology Review, 36*, 582–600.
- National Center on Response to Intervention (March 2010). Essential Components of RTI – A Closer Look at Response to Intervention. Washington, DC: U.S. Department of Education, Office of Special Education Programs, National Center on Response to Intervention.
- Simmons, D. C., Coyne, M. D., Kwok, O., McDonagh, S., Harn, B. A., & Kame'enui, E. J. (2008). Indexing Response to Intervention: A Longitudinal Study of Reading Risk From Kindergarten Through Third Grade. *Journal of Learning Disabilities, 41*, 158–173.
- Speece, D. (2005). Hitting the moving target known as reading development: Some thoughts on screening children and secondary interventions. *Journal of Learning Disability, 38*, 487–493.

NCRTI Screening—Glossary of Terms

Area Under the Curve (AUC)

AUC is an overall indication of the diagnostic accuracy of a receiver operating characteristic (ROC) curve (see definition that follows). AUC values closer to 1 indicate the screening measure reliably distinguishes among students with satisfactory and unsatisfactory reading performance, whereas values at .50 indicate the predictor is no better than chance.

Benchmark

A benchmark is a predetermined level of performance on a screening test that is considered representative of proficiency or mastery of a certain set of skills.

Classification Accuracy

The classification accuracy indicates the extent to which a screening tool is able to accurately classify students into “at risk for poor learning outcomes” and “not at risk for poor learning outcomes” categories.

Coefficient Alpha

The coefficient alpha is a measure of the internal reliability of items in an index. Values of alpha coefficients can range from 0 to 1.0. Alpha coefficients closer to 1.0 indicate the items are more likely to be measuring the same thing.

Construct Validity

Construct validity is a type of validity that assesses how well one measure correlates with another measure purported to represent a similar underlying construct.

Content Validity

Content validity is a type of validity that uses expert judgment to assess how well items measure the universe they are intended to measure.

Criterion Measure

A criterion measure is a dependent variable or outcome measure in a study.

Cross-Validation

Cross-validation is the process of validating the results of one study by performing the same analysis with another sample. In the cross-validation study, cut scores derived from the first study are applied to the administration of the same test and criterion measure with a different sample of students.

Cut Score

A cut score is a score on a screening test that separates students who are considered potentially at risk from those considered not at risk.

Disaggregated Data

Data are disaggregated when they are calculated and reported separately for specific subpopulations (e.g., race, economic status, academic performance).

Generalizability

Generalizability is the extent to which results generated from one population can be applied to another population. A tool is considered more generalizable if studies have been conducted on larger, more representative samples.

Interrater Reliability

Interrater reliability is the extent to which raters judge items in the same way.

Kappa

Kappa is an index that compares the agreement against that which might be expected by chance. Kappa can be thought of as the chance-corrected proportional agreement. Possible values range from +1 (perfect agreement) via 0 (no agreement above that expected by chance) to -1 (complete disagreement).

Norm

A norm is a standard of performance on a test that is derived by administering the test to a large sample of students. Results from subsequent administrations of the test are then compared to the established norms.

Predictive Validity

Predictive validity is a type of validity that assesses how well a measure predicts performance on some future similar measure.

Receiver Operating Characteristic (ROC) Curve

An ROC curve is a generalization of the set of potential combinations of sensitivity and specificity possible for predictors. It is a plot of the true positive rate (sensitivity) against the false positive rate (1-specificity) for the different possible cut points of a diagnostic test. The area under the curve (AUC) represents an overall indication of the diagnostic accuracy of an ROC curve. AUC values closer to 1 indicate the screening measure reliably distinguishes between students with satisfactory and unsatisfactory reading performance, whereas values at .50 indicate the predictor is no better than chance.

Reliability

Reliability is the consistency with which a tool classifies students from one administration to the next. A tool is considered reliable if it produces the same results when the test is administered under different conditions, at different times, or when using different forms of the test.

Response to Intervention (RTI)

RTI integrates assessment and intervention within a multi-level prevention system to maximize student achievement and reduce behavior problems. With RTI, schools identify students at risk for poor learning outcomes, monitor student progress, provide evidence-based interventions, adjust the intensity and nature of those interventions depending on a student's responsiveness, and identify students with learning disabilities.

Screening

Screening involves brief assessments that are valid, reliable, and evidence based. The assessments are conducted with all students or targeted groups of students to identify students

who are at risk of academic failure and therefore likely to need additional or alternative forms of instruction to supplement the conventional general education approach.

Sensitivity

Sensitivity is the extent to which a screening measure accurately identifies students at-risk for the outcome of interest.

Specificity

Specificity is the extent to which a screening measure accurately identifies students not at risk for the outcome of interest.

Split-Half Reliability

Split-half reliability is a method of assessing internal reliability by correlating scores from one half of the items on an index or test with scores on the other half of the items.

Test-Retest Reliability

Test-retest reliability is a correlation of scores on a test given at one time to scores on the test given at another time to the same subjects.

Validity

Validity is the extent to which a tool accurately measures the underlying construct it is intended to measure.

Appendix A: Handouts

Handout 1: Purpose for Screening

Individually, consider your purpose for screening, and identify your top three reasons for screening from the list that follows. If you do not see your purpose listed, use the space provided to add your own. Once you have completed your individual selections, share your selections with your team and discuss why you chose them.

What is our purpose for implementing screening?

- To evaluate core curriculum, core instruction, and intervention effectiveness.
- To evaluate program effectiveness for target groups (special education, Title 1, racial/ethnic groups).
- To identify students who need additional assessment and intervention.
- To identify number/percentage of students at risk.
- To assist in decisions about resource allocation.
- To maximize performance on end-of-year tests .
- To contribute to school improvement efforts.
- To assist in reporting district, school, or student performance to parents.
- To identify students at risk for poor learning outcomes.
- To reduce inappropriate referrals to special education or overidentification of students as students with disabilities.
- Other:
- Other:
- Other:

Top Three Reasons for Screening

Once everyone has shared their top three selections, work as a team to come to a consensus about the top three reasons for screening in your context, and justify your answers. List the top three reasons that your team identified for screening in your context, and justify your answers.

1. _____

2. _____

3. _____

Handout 2: Assessing your Needs, Priorities, and Logistics

In order to select a screening tool and establish a process for screening, you must first identify the needs and priorities of your district and determine the logistics for implementing screening. As a team, you have already identified your top three purposes for screening in *Handout 1*. Keep those purposes in mind while answering the questions that follow. If you begin by filling out the questions individually, make sure to take time to share responses and come to a consensus as a team.

1. What are the outcomes we are focused on?

What outcomes are you focused on? (e.g., behavioral, academic, postschool)	
What specific outcome measures are you focused on? (e.g., word reading fluency, computation)	
How are they aligned with the current curriculum?	

2. What is our scope?

Is it a district- or school-level initiative? Are efforts aligned?	
Are there other efforts in place that can be aligned with screening?	

3. Who is the target population?¹

What school(s)?	
What grade level(s)?	
Are there specific subgroup(s)?	
What are the demographics/characteristics of the population we need to consider (e.g., English language learners (ELLs), race/ethnicity)?	

4. What do we need to know about timing?

How frequently will screening occur?	
When during the year will screening occur?	
Will screening measures be administered individually or as a group?	
How will time be scheduled for the administration and scoring of the assessments? When will the assessments be scored?	
How will time be scheduled for reviewing results and incorporating them into instruction? When will this occur?	

¹ In thinking about the target population, it may be necessary to look at district-level data in order to understand the demographics of your population.

5. What are the staff roles related to screening?

Who will conduct the screening?	
Who will score the screening tools?	
Who will interpret the results?	
Does staff have adequate skills and knowledge?	

6. What do we need for administration of screening?

What materials are required to administer the screening tools?	
Are there accommodations for students with special needs and/or ELLs?	
What, if any, statistical software will be used for analysis, reporting, and monitoring?	

7. What funds are available?

What funds are available to purchase screening tools and materials?	
What costs are associated with training?	
Are there additional cost demands for hiring staff?	

8. What do we need for training?

How much time is available for training on screening tool administration and data analysis?	
Who will provide the training and technical assistance (internal or external)?	
How much training is needed to implement with fidelity? How much training is needed to use assessment results for decision-making?	

Handout 3: Selecting Screening Tools

In *Handout 2*, your team determined your needs, priorities, and logistics for implementing screening in your district, based on the purposes your team identified in *Activity Handout 1*. You will now use that information to help select a tool or determine whether the screening tool that is currently being used in your district is appropriate. It is important to recognize that the “best tool” may not be the same at every grade level or school.

The National Center on Response to Intervention (NCRTI) has developed a screening tools chart (www.rti4success.org) which lists and provides ratings for commercially available screening tools. Each tool has been reviewed by NCRTI’s Technical Review Committee (TRC). Information and data are provided about the technical rigor of the tool, the efficiency of the tool, and implementation requirements.

To answer the questions that follow, please first select a tool from the screening tools chart or think of the tool that your district is currently using to screen students. If you completed the following information individually, be sure to discuss your answers with your team and come to a team consensus.

	Yes	No	Comment
Does the tool match your priorities and provide the data needed to make the decisions you hope to make?			
Has the screening tool been shown to be valid?			
Has the screening tool been shown to be reliable?			
Has the screening tool been shown to be generalizable?			

Has your screening tool been shown to accurately classify students?			
Is the tool appropriate for your student population?			
Is the tool appropriate for the resources you have available?			
Is the tool appropriate for the materials that you have?			

National Center on Response to Intervention

1000 Thomas Jefferson Street, NW

Washington, DC 20007

Phone: 877-784-4255

Fax: 202-403-6844

Web: <http://www.rti4success.org>

