Universal Math Screeners in Elementary School

Prepared for Northwest ISD

April 2014



In the following report, Hanover Research reviews best practices and content considerations for selecting and implementing an effective RTI screening process for elementary level math programs.

HANOVER RESEARCH

Table of Contents

EXECUTIVE SUMMARY AND KEY FINDINGS

INTRODUCTION

Response to Intervention (RTI) has become one of the standard approaches to identifying and supporting students who need additional instructional strategies to be successful in school. The first step to an effective RTI framework is the screening process, which determines which students should receive additional help. An effective screening process can allow struggling students to receive the assistance they need with minimal disruption to other students, while a poor screening process can overlook those students at a critical developmental point or incorrectly target students for intervention.

In this report, Hanover Research reviews best practices and considerations for selecting and implementing an effective screening process for elementary level math programs. Section I introduces a typical RTI framework and discusses the components necessary for an effective screening tool, including explanations of validity and reliability. The section reviews research on the math content areas that experts have identified as most useful for screening and identifying struggling students, and also provides guidelines for selecting and implementing a screening tool. Section II of the report reviews the RTI frameworks used by districts similar to Northwest ISD, with particular focus on screening and relevant progress monitoring components.

KEY FINDINGS

- Research has demonstrated that number sense ability is the most critical area to screen in order to predict a student's future math achievement. Number sense involves understanding and the manipulation of numbers, and forms the foundation for more advanced computation. Screening measures at all grades should test rational number concepts and computational proficiency.
- Research on grade-specific math indicators is skewed toward kindergarten and grade 1. At that age, evidence-based tasks that can be incorporated in screens include magnitude comparison/quantity discrimination, strategic counting, retrieval of basic arithmetic facts, word problems, and numeral recognition. While there is less research available on the best content areas to screen in the upper elementary grades, state standardized tests can act as an initial screen to determine which students need additional testing.
- Curriculum-based measurement is an approach to screening that draws on a student's core class curriculum and/or state standards. CBM can be a strong alternative to commercial screening tools, as districts can develop their own probes to fit specific curricula and these tools can be easily used for subsequent progress monitoring.

- The majority of the districts profiled use a combination of assessment methods for mathematics. For example, Campbell County relies on both NWEA MAP tools and customizable CBM probes, while Rockwall ISD incorporates state standardized tests, CBM probes, and district developed screening tools into its RTI framework.
- Reliability and validity measures are the primary methods of assessing a screen's effectiveness. A tool with a high rate of reliability will consistently classify students correctly even if it is administered under different situations or by different individuals. A tool with a high degree of validity will be able to correctly predict which students are at-risk, and which students are not.
- Screening processes that achieve their full potential for reliability and validity often consist of multiple measures and/or multiple stages, or incorporate short-term progress monitoring after an initial screen. Adding these additional components helps to verify the results of an initial screen and correctly categorize students who may have scores close to the cutoff point.
    - Of the districts profiled, there was an equal divide between those that used a single screen to determine a student's intervention need (Campbell, Huntington, and Augusta), and those that incorporated additional progress monitoring measures (Hays and Rockwall).
- Screening efficiency is generally based on how quickly a screen can be administered, scored, and analyzed. Generally, the ideal time for a screen is 10 to 20 minutes, although there are highly regarded screens that take both more and less time.

SECTION I: BEST PRACTICES FOR RTI SCREENING TOOLS

In this section, Hanover reviews available literature on best practices in RTI screening tools, including structural considerations, selection and use of a screening tool, and critical components of mathematics screening tools. This section concludes with a comparison of several commercial screening tools.

RESPONSE TO INTERVENTION FRAMEWORK

Most commonly accepted models of Response to Intervention in education adhere to a three-tier framework. The three different tiers act as a multi-level prevention system designed to identify students at different degrees of risk for academic failure.[1] The first tier, or primary prevention, consists of what is assumed to be high quality core instruction. Universal screening is used on a regular basis, usually three times a year, to identify struggling students.[2] Students who perform poorly in screenings are moved to the second tier, or secondary prevention. At this tier, students generally receive an intervention of research-based supplemental instruction in a small group setting.[3] Finally, students who show no improvement are moved to tier three, or tertiary prevention. In tier three, students receive more individualized, intensive instruction that targets specific skill deficits.[4] Students who do not show improvement after the tertiary level are often referred for more comprehensive evaluation and consideration for special education services.[5]

A comprehensive RTI system requires that student data be collected for multiple purposes. Data collected through **initial screening** (often universal, as described above) help determine whether a child should be referred for more intensive assessment. Data collected through **diagnostic assessments** are more in-depth and help identify specific strengths and weaknesses of individual students for the purposes of planning interventions. Frequently, effective diagnostic assessment requires the use of multiple assessment tools. **Benchmark progress monitoring** (usually referred to as progress monitoring) is data gathered at predetermined times of the year to determine if students are making progress on established benchmarks or standards. And finally, **summative outcome assessments** produce data gathered at the end of the year to determine the overall effectiveness of instruction and an intervention.[6] This report will focus primarily on tools used for initial

[1] "What is RTI?" RTI Action Network. http://www.rtinetwork.org/learn/what/whatisrti
[2] [1] Ibid.
[2] "Multi-level Prevention System." Center on Response to Intervention. http://www.rti4success.org/essential-components-rti/multi-level-prevention-system
[3] Ibid.
[4] Ibid.
[5] "What is RTI?," Op. Cit.
[6] Wixson, K. and S. Valencia. "Assessment in RTI: What Teachers and Specialists Need to Know." *The Reading Teacher,* 64:6, 2011. pp. 467-468.

screening, although most of the tools and approaches can be used for the other more intensive purposes listed as well.

STRUCTURAL CONSIDERATIONS FOR A UNIVERSAL SCREENING TOOL

Reliability and validity are the primary ways to quantify the effectiveness of a screening tool. **Reliability** refers to how consistently a screening tool can classify students when it is administered under different conditions or by different individuals, while **validity** refers to how accurately a tool measures what it intends to measure.[7] In the context of a screening system, validity is usually meant as **predictive validity,** or the degree to which a tool accurately predicts a student's later performance.[8] Predictive validity is particularly important in terms of adapting instruction and interventions. Validity can be broken down into degree of sensitivity and degree of specificity: **Sensitivity** refers to the degree that a screen can accurately identify at-risk students, while **specificity** indicates the degree to which it accurately identifies students who are *not* at-risk (i.e., students with satisfactory performance are not incorrectly identified as being at-risk).[9]

A fundamental component of a valid and reliable screening tool is an accurate **cut score**. The cut score is the value that divides students considered at-risk from those who are not at-risk.[10] Often, cut-scores align with or have the same value as established learning targets or benchmarks.[11] There is limited consensus on what an ideal cut-score should be.[12] Cut scores in published screening tools are commonly based on national, aggregated norms. Some schools wish to develop cut scores based on local norms if they seem to vary from the national norms. However, doing so requires professional assistance from a statistician and trained test developer and has implications for the overall accuracy of the screening tool.[13]

As illustrated in Figure 1.1, there are four different outcomes for how a screening tool can interpret student performance relative to a specified cut score.[14] False positives and false negatives represent screening errors. Screening assessments with higher validity rates have higher rates of sensitivity and specificity, and thus more true positives and true negatives. A perfect assessment would have validity, sensitivity, and specificity rates of 100 percent

[7] "Screening Tools Chart Rating System." Center on Response to Intervention. http://www.rti4success.org/resources/tools-charts/screening-tools-chart/screening-tools-chart-rating-system

[8] Gersten, R., S. Beckmann, B. Clarke, A. Foegen, L. Marsh, J. Star, and B. Witzel. "Assisting Students Struggling with Mathematics: Response to Intervention (RTI) for Elementary and Middle Schools." Institute of Education Sciences, 2009. p.11. http://ies.ed.gov/ncee/wwc/pdf/practice_guides/rti_math_pg_042109.pdf#page=19

[9] Gersten, R., B. Clarke, N. Jordan, R. Newman-Gonchar, K. Haymond and C. Wilkins. "Universal Screening in Mathematics for the Primary Grades: Beginnings of a Research Base." *Exceptional Children,* 78:4, 2012. p. 437. Retrieved from ProQuest.

[10] "Screening Briefs Series #2: Cut Scores." National Center on Response to Intervention. p. 1. http://www.rti4success.org/sites/default/files/RTI%20Screening%20Brief2-Cut%20Scores.pdf

[11] "Using Screening Data for Decision Making." National Center on Response to Intervention. p .8. http://www.rti4success.org/sites/default/files/Using_Screening_Data%20Transcript.pdf

[12] Hughes, C. and D. Dexter. "Universal Screening Within a Response-to-Intervention Model." http://www.rtinetwork.org/learn/research/universal-screening-within-a-rti-model

[13] "Screening Briefs Series #2: Cut Scores," Op. Cit. pp. 1-3.

[14] "Screening Briefs Series #1: Classification Accuracy." National Center on Response to Intervention. p.1. http://www.rti4success.org/sites/default/files/RTI%20Screening%20Brief1-Classification%20Accuracy.pdf

each, although identifying a screening tool with a balance between ideal specificity and sensitivity is difficult.[15] A more stringent screening tool would correctly identify most or all at-risk students, but may also result in a number of false positives; meanwhile, a more lenient screening tool that identifies fewer students would not incorrectly identify students as being at-risk, but would likely miss some at-risk students.[16] Expert practitioners suggest that screening tools should have a sensitivity, or true positive rate, of greater than 90 percent, and a specificity, or true negative rate, of greater than 80 percent to be considered effective.[17]

Figure 1.1: Screening Validity Determinants

		Actual Performance (Based on Outcome Assessment)	
		Fail	**Pass**
Predicted Performance (Based on Screening Assessment)	**At Risk**	True Positive	False Positive
	Not at Risk	False Negative	True Negative

National Center on Response to Intervention[18]

Selecting and Using a Screening Tool

Selecting and implementing screening tools is a critical first step in ensuring that teachers have adequate data to make decisions about instructional interventions.[19] Screening tools generally consist of brief assessments focused on skills that are considered highly predictive of future outcomes.[20] A panel assembled by the Institute for Education Sciences (IES), publisher of the What Works Clearinghouse, recommends that the team that selects screening measures should include individuals with expertise in math instruction as well as experts with training in measurement, such as a school psychologist or a member of a district research and evaluation team. It also asserts that measures must be selected on the basis of reliability, predictive validity, and the content addressed, with special consideration for each grade level's instructional objectives.[21]

Research has shown that **engaging in progress monitoring, or using multiple screening stages or multiple screening measures, can increase overall assessment accuracy**.[22] Layering the screening process in this way ensures that screening results are not identifying students based on a one-time performance, but on an actual trend in a student's achievement. IES stresses that because no one screening measure is perfect, schools should monitor the progress of students who score slightly above or below a cutoff score.[23] In two-part screens that result in overall higher sensitivity and specificity, the first stage assesses all

[15] Ibid., p. 2
[16] "Screening Briefs Series #2: Cut Scores," Op. Cit., pp. 3.
[17] "Screening Briefs Series #1: Classification Accuracy," Op. Cit., p. 2
[18] Ibid., p. 1.
[19] Ibid.
[20] Hughes and Dexter, Op. Cit.
[21] "Assisting Students Struggling with Mathematics: Response to Intervention (RTI) for Elementary and Middle Schools," Op. Cit., p.14.
[22] "Screening Briefs Series #1: Classification Accuracy," Op. Cit., p. 3.
[23] "Assisting Students Struggling with Mathematics: Response to Intervention (RTI) for Elementary and Middle Schools," Op. Cit., p. 7.

students to identify low-risk students, or true negatives, and the second stage gives remaining students who might be at risk a more intensive assessment.[24] A multi-part screening process or incorporation of progress monitoring also helps identify exactly where intervention is needed in a student's development.[25]

Once a screening tool has been selected, student scores can be presented and interpreted in several ways. A **norm-referenced score** compares an individual's performance to other students within the same grade level. A norm-referenced score can be quantified by **percentile rank** within a group, or by **standard scores**. Standard scores are expressed in terms of standard deviations from the mean in a group. The following example illustrates how standard scores and percentile ranks are related: "a student achieving a standard score of 60 on an assessment with a mean of 50 and a standard deviation of 10 has scored one standard deviation above the mean. Under normal circumstances, a score that is 1 standard deviation above the mean is ranked at the 84th percentile."[26] Alternatively, a **criterion-referenced score** is based on a set performance standard, and measures students based on how they demonstrate a pre-determined and objective level of knowledge of particular area of content or skill.[27] Criterion scores often separate students into multiple levels, rather than simply at-risk or not at-risk.[28]

Screening efficiency is generally based on how quickly a screen can be administered, scored, and analyzed. This process can take as little as a few minutes for a single proficiency measure to nearly an hour for multicomponent measures. The most promising research to date suggests that **brief measures that take about 10 minutes and cover a representative sample of grade-level objectives are reasonable and sufficiently reliable**, but the IES notes that research on the ideal length of time that should be given to an assessment is still limited.[29] The IES suggests that a screening measure should take a maximum of 20 minutes to administer, although overall efficiency must be balanced with validity and reliability.[30] Notably, both timed screening measures and untimed screening measures have been shown to be effective.[31]

The greatest challenges to efficient screenings are time and personnel resources. It is common for teachers to collect student data only "when time permits," which tends to draw out the screening and progress monitoring process and delay interventions. Furthermore, screening measures are sensitive to classroom instruction, so differences in assessment frequency between classrooms may result in group performance gaps on the sole basis of exposure to additional instruction. Using teams of screeners to measure all

[24] "Screening Briefs Series #1: Classification Accuracy," Op. Cit., p. 3.
[25] "RTI and Mathematics." RTI Action Network. http://www.rtinetwork.org/professional/rti-talks/transcript/talk/36
[26] "Screening Briefs Series #2: Cut Scores," Op. Cit., p. 2.
[27] [1] Ibid.
[2] "Using Screening Data for Decision Making," Op. Cit., p. 22
[28] "Using Screening Data for Decision Making," Op. Cit., p. 8
[29] "Assisting Students Struggling with Mathematics: Response to Intervention (RTI) for Elementary and Middle Schools," Op. Cit., p. 13
[30] Ibid., p. 14.
[31] Ibid., p. 13.

students within a short period of time is one of the most effective ways to address these challenges.[32]

A final consideration in the implementation of a screening tool is the likelihood of measurement error. Generally, there are three types of measurement error: inconsistent assessment administration, or user error when a tester administers the exam; errors in scoring assessments; and data entry errors. Assessments that have an evaluation component requiring administrator judgment are more susceptible to measurement error.[33]

BEST PRACTICES IN MATH SCREENING TOOLS

As publishers of education resources have recognized the increasing popularity of RTI initiatives, many have designed products to meet demand for commercial RTI tools. In the past, the most popular of these products for both math and reading have included Pearson's AIMSweb, CBT/McGraw Hill's Yearly Progress Pro, Renaissance Learning's STAR Reading and STAR Math, and Edcheckup.[34] Some experts have expressed concern that too much of a reliance on such products can cause schools to fail to understand and integrate fundamental principles of RTI, and thus implement less effective programs.[35] In part because of that concern and because of other legal considerations, discourse on popular and effective screening tools among expert non-profit RTI organizations and practitioners tends to focus on determining what general content should be included in effective tools. Much of the existing research on math screenings at the elementary level has focused primarily on kindergarten and grade 1, so an understanding of the most effective measures for upper grades, particularly using Common Core State Standards, is an area in need of more research.[36] However, existing research from IES and other experts has identified several specific areas that are critical for math screening tools at the earliest elementary grade levels, and provides general direction for the upper grade levels.

Research has shown that the most effective assessments across grade levels measure what is known as number sense.[37] Number sense, or "knowledge of whole numbers,"[38] is the foundation of mathematical knowledge and understanding. Although precise definitions

[32] Ibid., pp. 15-16.
[33] "Screening Briefs Series #4: Ensuring Fidelity of Assessment and Data Entry Procedures." National Center on Response to Intervention. p. 1. http://www.rti4success.org/sites/default/files/RTI%20Screening%20Brief4-Ensuring%20Fidelity.pdf
[34] Rebora, A. "Tools of the Trade." *Education Week*, April 12, 2010. http://www.edweek.org/tsb/articles/2010/04/12/02tools.h03.html
[35] [1] Ibid.
[2] Samuels, C. "Curriculum Developers Seek to Capitalize on RTI's Growth." Education Week, March 2, 2011. http://search.proquest.com/docview/856809502/fulltext/F0BCC525E79C4FAAPQ/1?accountid=132487#
[36] [1] Ibid.
[2] "Universal Screening in Mathematics for the Primary Grades: Beginnings of a Research Base," Op. Cit., p. 442.
[37] "Assisting Students Struggling with Mathematics: Response to Intervention (RTI) for Elementary and Middle Schools," Op. Cit., p. 13.
[38] Ibid.

can vary, number sense is defined broadly an "organizing schemata"[39] that is typically represented by the ability to use mental number lines and manipulate numerical properties.[40] For example, number sense allows individuals to easily estimate and judge the magnitude of numbers, recognize unreasonable results, and be flexible in mental calculations.[41] Mastery of number sense can be tested by single proficiency measures that only assess one skill, such as magnitude comparisons, or by multiple-proficiency measures that assess several skills, such as comparisons, counting, and computation to produce a composite score.[42] Past examinations of screening measures have commonly focused on single proficiency measures, but new research is showing that multiple proficiency measures may have equal or higher predictive validity rates.[43]

IES recommends that **screening measures at all elementary grades should assess student understanding of whole and rational number concepts, as well as computational proficiency**.[44] In grades 4 through 8, IES further recommends that schools use the prior year's state testing results as a part of preliminary screening, so that students who score near the cutoff point for reaching state benchmarks can be considered for subsequent testing. Structuring an RTI screening protocol this way allows districts to "combine a broader measure that covers more content with a screening measure that is narrower but more focused;" however, the IES notes that because there are fewer established screening measures at upper grade levels, districts may need to rely more heavily on commercial assessments or to develop their own screening tools.[45]

Math assessments conducted when a child is in kindergarten or grade 1 can correctly predict which students will struggle with math in later grades, even holding constant for intelligence quotient and socio-economic status.[46] Research reviews have established several skill areas related to number sense or memory capability that have especially high predictive validity when tested in those grades, including:

- **Magnitude Comparison/Quantity Discrimination**: Ability to identify bigger and smaller numbers or amounts. For example, showing a student a picture of a worm and five birds and asking if each bird can have a worm requires the student to make

[39] Gersten, R., B. Clarke, K. Haymond, and N. Jordan. "Screening for Mathematics Difficulties in K-3 Students: Second Edition." Center on Instruction, 2011. pp. 4-5. http://www.centeroninstruction.org/files/Screening%20for%20Math%202nd%20Ed%202-3-12.pdf

[40] [1] "Universal Screening in Mathematics for the Primary Grades: Beginnings of a Research Base," Op. Cit., p. 424. [2] "Screening for Mathematics Difficulties in K-3 Students: Second Edition," Op. Cit., p. 5.

[41] "Screening for Mathematics Difficulties in K-3 Students: Second Edition," Op. Cit., pp. 4-5.

[42] Ibid., p. 9.

[43] [1] "Universal Screening in Mathematics for the Primary Grades: Beginnings of a Research Base," Op. Cit., p. 434. [2] "Assisting Students Struggling with Mathematics: Response to Intervention (RTI) for Elementary and Middle Schools." Op. Cit., p. 11.

[44] "Assisting Students Struggling with Mathematics: Response to Intervention (RTI) for Elementary and Middle Schools." Op. Cit., p. 15

[45] Ibid.

[46] "Screening for Mathematics Difficulties in K-3 Students: Second Edition," Op. Cit., p. 3.

a "gross magnitude judgment."[47] Almost all screening tools use some measure of magnitude comparison.[48] IES has found magnitude comparisons to be one of the most predictive screening measures.[49]

- **Strategic Counting**: This refers to a) Knowledge of basic counting principles, such as understanding that changing the order that one counts items does not mean that the overall quantity will change, and b) Counting skills such as correctly using manipulatives (like fingers or objects) to count to a certain number.[50] Asking students to identify missing numbers from a sequence is a common way to measure this skill.[51] IES has found strategic counting to be one of the most predictive screening measures.[52]
- **Retrieval of Basic Arithmetic Facts**: Research has found that students who cannot retrieve basic addition and subtraction number combinations without using manipulatives may be more likely to have a math-related learning disability. This indicates a problem with semantic memory, or the ability to store and retrieve abstract information.[53] The relationship between working memory and number sense is not fully understood, and at this time research shows fact retrieval is most viable when used as an additional variable to a set of other predictors.[54]
- **Word Problems**: These are recent additions to many early math screens, based on evidence that young children usually find solving addition and subtraction word problems easier than the same problems written only in terms of numbers.[55]
- **Numeral Recognition**: This is an understanding of the naming system used for numbers and math functions and "serves as a gateway skill to formal mathematics."[56]

[47] "Screening Briefs Series #3: Predicting Students at Risk for Reading and Mathematics Difficulties." National Center on Response to Intervention. p. 3. http://www.rti4success.org/sites/default/files/RTI%20Screening%20Brief3-Predicting%20Students.pdf
[48] "Universal Screening in Mathematics for the Primary Grades: Beginnings of a Research Base," Op. Cit., p. 432
[49] "Assisting Students Struggling with Mathematics: Response to Intervention (RTI) for Elementary and Middle Schools." Op. Cit., p. 11
[50] [1] "Screening for Mathematics Difficulties in K-3 Students: Second Edition," Op. Cit., p. 6.
[2] "Screening Briefs Series #3: Predicting Students at Risk for Reading and Mathematics Difficulties," Op. Cit., p. 3.
[51] [1] "Screening for Mathematics Difficulties in K-3 Students: Second Edition," Op. Cit., p. 10.
[2] "Universal Screening in Mathematics for the Primary Grades: Beginnings of a Research Base," Op. Cit., p. 423
[52] "Assisting Students Struggling with Mathematics: Response to Intervention (RTI) for Elementary and Middle Schools." Op. Cit., p. 11.
[53] [1] "Screening for Mathematics Difficulties in K-3 Students: Second Edition," Op. Cit., p. 7.
[2] "Universal Screening in Mathematics for the Primary Grades: Beginnings of a Research Base," Op. Cit., p. 433
[54] [1] "Universal Screening in Mathematics for the Primary Grades: Beginnings of a Research Base," Op. Cit., p. 435
[2] "Assisting Students Struggling with Mathematics: Response to Intervention (RTI) for Elementary and Middle Schools." Op. Cit., p. 11.
[55] [1] "Universal Screening in Mathematics for the Primary Grades: Beginnings of a Research Base," Op. Cit., p. 433
[2] "Screening Briefs Series #3: Predicting Students at Risk for Reading and Mathematics Difficulties," Op. Cit., p. 3
[3] "Assisting Students Struggling with Mathematics: Response to Intervention (RTI) for Elementary and Middle Schools." Op. Cit., p. 11.
[4] "Screening for Mathematics Difficulties in K-3 Students: Second Edition," Op. Cit., p. 8.
[56] "Screening for Mathematics Difficulties in K-3 Students: Second Edition," Op. Cit., p. 8.

MATH SCREENS AND CURRICULUM-BASED MEASUREMENT

One of the most popular frameworks for combining screening and progress monitoring is **curriculum-based measurement** (CBM).[57] CBM functions through short, timed tasks known as "probes" that are derived from the curriculum. Probes are administered using standardized conditions, and are easy score.[58] Teachers can develop their own probes or draw on probes from external sources derived from standard grade level learning objectives.[59] CBM probes can measure both fluency (speed) and accuracy: in CBM, a student is given credit for computation problems based on the number of correct digits, rather than the overall accuracy of the final response.[60] CBM requires that a teacher index a student's performance with grade level curriculum benchmarks.[61] In other words, it processes assessment results as criterion-referenced scores. CBM systems with probes derived from state standards have been found to yield reliable scores that have moderate correlation with more comprehensive math measures.[62] This research base, combined with the cost-effective nature of these systems in comparison to commercial RTI screens, has made state-aligned CBMs an appealing option for implementation for many districts.[63]

Dr. Amanda VanDerHeyden, RTI expert and a member of the Advisory Panel for the National Center for Learning Disabilities, recommends starting with CBM screening by identifying all computation and procedure skills in a set of state standards, and then incorporating them into both a screening probe and a progress monitoring probe. The progress monitoring probe represents a range of skills a student is expected to master throughout the year. Even if a child has not mastered a particular skill yet, the monitoring probe establishes a baseline so that student development can be measured using the same probe at different points during the year.[64] Figure 1.2 illustrates one example of this approach offered by Dr. VanDerHeyden.

Figure 1.2: Probe Guidelines

	FALL SCREENING	WINTER SCREENING	SPRING SCREENING
1st Grade	Sums to 5	Sums to 18 or 20	Addition and Subtraction 0-20
2nd Grade	Addition and Subtraction 0-20	Multi-digit addition or subtraction without regrouping	Fact Families and Addition/Subtraction 0-20

[57] Hughes and Dexter, Op. Cit.
[58] Wright, J."Curriculum-Based Measurement: A Manual for Teachers." Intervention Central. http://www.jimwrightonline.com/pdfdocs/cbaManual.pdf
[59] "Screening for Mathematics Difficulties in K-3 Students: Second Edition," Op. Cit., p. 18
[60] Ibid., pp. 3, 6
[61] Fuchs, D. "Validated Forms of Progress Monitoring in Reading and Mathematics." RTI Action Network. http://www.rtinetwork.org/essential/assessment/progress/validated-forms-progressmonitoring
[62] VanDerHeyden, A. "RTI and Math Instruction." RTI Action Network. http://www.rtinetwork.org/learn/what/rtiandmath
[63] "RTI and Mathematics." Op. Cit.
[64]" Implementing RTI in Mathematics." RTI Action Network. http://www.rtinetwork.org/mpdf_print.php?htc=YToxOntzOjEzOiJyZXBvcnRfaWRjaGF0IjtzOjE6IjgiO30=

	Fall Screening	Winter Screening	Spring Screening
3rd Grade	Fact Families and Addition/ Subtraction 0-20 or 3-digit addition and subtraction with and without regrouping (this is hard for most third graders but reflects a skill that most are expected to be able to do)	Multiplication 0-9 or 0-12	Multiplication and Division 0-12
4th Grade	Fact Families and Multiply/Divide 0-12	Multi-digit multiplication without or with regrouping	Multi-digit division with and without remainders
5th Grade	Multi-digit multiplication without or with regrouping	1 digit into 2-3 digit dividend with remainders	Reduce fractions
6th Grade	Decimals multiplication	Find least common denominator	Substitution of whole number to solve equations

Source: RTI Action Network[65]

Commercial Tool Comparisons

While the National Center on Response to Intervention cannot make specific recommendations on which tools to use, it can assess underlying publisher claims. In 2011, the Center's Technical Review Committee (TRC) reviewed data behind current commercial RTI screening tools to compare effectiveness and efficiency of implementation. These tools were voluntarily submitted for review by their developers, and were assessed against a pre-established set of criteria.[66]

Figure 1.3 provides an overview of the quality indicators that the TRC examined. In addition to examining evidence regarding reliability and validity, the TRC also measured generalizability, or the extent to which results from one group could be applied to another group; quality of disaggregated data for subgroups such as economic status; and classification accuracy, or the extent to which a tool can accurately divide students into two categories.[67] Figure 2.4 collects information on implementation parameters and overall efficiency. This includes information on cost, training and administration time, and technology requirements for the test-taking phase. All of the tests are scored by computer, and a large number require internet access.

While Hanover sought to find examples of specific types of tasks measured by each tool to compare to the best practices in screening content described earlier in this section, we found that such information was generally not available to the public. As a result, the information included in the figures on the following pages focuses on general effectiveness

[65] Ibid., p. 3.
[66] "Screening Tools Chart Rating System," Op. Cit.
[67] "User's Guide to Universal Screening Tools Chart." National Center on Response to Intervention, May 2010. p. 4. http://www.rti4success.org/sites/default/files/UniversalScreeningUsersGuide.pdf

and implementation parameters rather than a specific breakdown of the structure of each tool. Our analysis yielded the following overall findings:

- The tool with the **highest rankings across all quality indicators** was Star Math.
- The tools with the **highest rates of validity and reliability** were the single proficiency measures AIMSweb Early Numeracy Quantity Discrimination and Missing Number, and STAR Math.
- The tool with the **greatest time requirements for administration and scoring** was Group Math Assessment and Diagnostic Evaluation (GMADE).
- The **most expensive** tool was Acuity Mathematics.

Figure 1.3: Comparisons of Tool Quality

Tool	Unit Title	Evidence of Classification Accuracy	Generalizability	Evidence of Reliability	Evidence of Validity	Disaggregated Data	Grades Served
A+ LearningLink: Progress in Math	Math	Partially convincing	Moderate Low	Partially convincing	Partially convincing	No Data Available	Grades 1-8
Acuity	Mathematics	Partially convincing	Moderate High	Partially convincing	**Convincing**	No Data Available	Grades 3-8
AIMSweb	Math-CBM	Partially convincing	Moderate high	Partially convincing	Unconvincing	No Data Available	Grades 1-8
	Early Numeracy-Missing Number	Unconvincing	**Broad**	**Convincing**	**Convincing**	No Data Available	Grades K-1
	Early Numeracy-Number Identification	Unconvincing	**Broad**	**Convincing**	Unconvincing	No Data Available	Grades K-1
	Early Numeracy- Oral Counting	Unconvincing	Moderate Low	Unconvincing	Unconvincing	No Data Available	Grades K-1
	Early Numeracy-Quantity Discrimination	Unconvincing	**Broad**	**Convincing**	**Convincing**	No Data Available	Grades K-1
	Mathematics Concepts and Applications	Partially convincing	Moderate Low	**Convincing**	Unconvincing	Unconvincing	Grades 2-8
Classworks Universal Screen	Math	Partially convincing	Moderate High	**Convincing**	Unconvincing	No Data Available	Grades K-10
Discovery Education Predictive Assessment	Math	**Convincing**	Moderate High	**Convincing**	Partially convincing	Partially convincing	Grades K-12
easyCBM	Mathematics	**Convincing**	Moderate High	Partially convincing	**Convincing**	**Convincing**	Grades K-8
Group Assessment and Diagnostic Evaluation	Group Math Assessment and Diagnostic Evaluation	Partially convincing	Moderate Low	**Convincing**	Partially convincing	No Data Available	Grades K-12
Iowa Tests of Basic Skills	Math	**Convincing**	Moderate High	Partially convincing	**Convincing**	No Data Available	Grades K-8

mCLASS	Math	Unconvincing	Moderate high	**Convincing**	Unconvincing	Unconvincing	Grades K-3
Measures of Academic Progress (MAP)	Mathematics	**Convincing**	Moderate high	Partially convincing	Partially convincing	Partially convincing	Grades 2-10
MAP for Primary Grades	Mathematics	**Convincing**	Moderate High	Partially convincing	Partially convincing	Partially convincing	Grades K-2
STAR	Math	**Convincing**	**Broad**	**Convincing**	**Convincing**	**Convincing**	Grades 1-12

Source: National Center on Response to Intervention[68]

[68] Adapted From "Screening Tools Chart." Center on Response to Intervention. http://www.rti4success.org/resources/tools-charts/screening-tools-chart

Figure 1.4: Comparison of Implementation Parameters for RTI Math Screening Tools

Tool	Unit Title	Administration Format	Technology Requirements for Testing	Administration Scoring & Time	Cost	Training Requirements
A+ LearningLink: Progress in Math	Math	Group	Internet access	35-40 Minutes	$5 per student per year	1-4 hours
Acuity	Mathematics	Group	Internet access	50 Minutes	$5,200 for setup + $9.50 per student per year	Less than 1 hour
AIMSweb	Math-CBM	Group	Internet access	2 Minutes	$3-$5.00 per student per year	4-8 hours
	Early Numeracy- Missing Number	Individual	Internet access	2 Minutes	$3-$5.00 per student per year	4-8 hours
	Early Numeracy- Number Identification	Individual	Internet access	2 Minutes	$3-$5.00 per student per year	4-8 hours
	Early Numeracy- Oral Counting	Individual	Internet access	2 Minutes	$3-$5.00 per student per year	4-8 hours
	Early Numeracy- Quantity Discrimination	Individual	Internet access	2 Minutes	$3-$5.00 per student per year	4-8 hours
	Mathematics Concepts and Applications	Group Individual	Internet access	11-13 Minutes	$4.00 per student per year	1-4 hours
Classworks Universal Screen	Math	Group	None listed	30 Minutes	$4.00 per student per year	Less than 1 hour
Discovery Education Predictive Assessment	Math	Group	Internet access	40 Minutes	$8-$8.50 per student per year. Optional $2,500 for on-site professional development, $250 for webinar.	Less than 1 hour
easyCBM	Mathematics	Group Individual	Internet access	30 Minutes	Teacher version: Individuals can sign up for free. District version: $1.00 per student per year. Options for each version differ slightly.	1-4 hours

Tool	Unit Title	Administration Format	Technology Requirements for Testing	Administration Scoring & Time	Cost	Training Requirements
Group Assessment and Diagnostic Evaluation	Group Math Assessment and Diagnostic Evaluation (GMADE)	Group Individual	Internet access	46-95 Minutes	Test forms: $116.50-$309.95 per 30 students. Additional packages of 10 for $35.50. Software: $411.95-$2,695.95	Less than 1 hour-4 hours
Iowa Tests of Basic Skills	Math	Group	Computer	60 Minutes	$20.26 per student in year 1, $4.67 per student per year thereafter	Less than 1 hour
mCLASS	Math	Group Individual	Internet access and a hand-held computer	1-12 Minutes	$13.90 per student per year, $75 per kit, $400 installation fee, $1,400 on-site installation	4-8 hours
NWEA Measures of Academic Progress (MAP)	Mathematics	Group Individual	Computer: Can deliver with or without internet access.	40 Minutes	$13.50 per student per year + $3,700 optional for on-site training	1-4 hours
NWEA MAP for Primary Grades	Mathematics	Group Individual	Computer: Can deliver with or without internet access.	40 Minutes	$13.50 per student per year + $3,700 optional for on-site training	1-4 hours
STAR	Math	Group	Computer	10 Minutes	Year 1: Assuming minimum of 203 students: $1,799. Total and $.99 per student every year thereafter	Less than 1 hour

Source: National Center on Response to Intervention[69]

[69] Information on individual tools linked to via "Screening Tools Chart," Op. Cit.

SECTION II: SCHOOL PROFILES

In this section of the report, Hanover examines the RTI frameworks of five school districts with characteristics similar to Northwest Independent School District. These districts are identified by the National Center for Education Statistics as being part of Northwest ISD's 100-member peer group.[70] This means they are among the 100 schools found to be most similar to Northwest ISD across various characteristics, including district type, locale, total enrollment, student/teacher ratio, and the percentage of children in poverty.[71] Each profile provides an overview of the rationale and framework used for RTI by each district, including particular emphasis on universal screening components and relevant progress monitoring components in mathematics.

ROCKWALL INDEPENDENT SCHOOL DISTRICT, TX

Rockwall ISD is characterized as a rural fringe district, with a student population of 14,330 and a student/teacher ratio of 16.38.[72] The three-tier RTI program is operated out of the Curriculum department. Tier I consists of Core Instruction, which relies on RISD curriculum and evidence-based instruction. Students are screened in the fall, winter, and spring to determine which students need additional assistance; and students who meet the criteria are referred to a Student Support Team (SST). The SST develops a Team Learning Plan, which is considered the first level of RTI intervention. Students who continue to fail to meet screening requirements, or who are unsuccessful in Tier I coursework, are referred to Tier II, in which students receive supplemental instruction through flexible grouping. Progress monitoring is also implemented at this stage, and data collected through that process is considered in subsequent SST meetings. Referral to Tier III is considered a de facto referral to special education, and as such, requires the SST to intensively review instruction and intervention history as well as a consult with parents.[73]

Figure 2.1 illustrates the math screening that is conducted at each assessment point by grade. Rockwall uses a combination of First Steps Diagnostics Tasks, curriculum based assessments (essentially CBMs), performance on state standardized testing, and district-developed screening tools to determine the appropriate tier for each student.

[70] "Public School District Finance Peer Search: Northwest ISD." National Center for Education Statistics. http://nces.ed.gov/edfin/search/peergroupdata.asp?dataid=1&mt=0&subdataid=1&bleaid=4833180&jobid={2DD5A68B-253B-4807-B880-D0408F4E7BE7}

[71] "How to Use the Peer Search Tool." National Center for Education Statistics. http://nces.ed.gov/edfin/search/info.asp?t=faq

[72] "Search for Public School Districts: Rockwall ISD." National Center for Education Statistics. http://nces.ed.gov/ccd/districtsearch/district_detail.asp?Search=1&details=1&InstName=rockwall&DistrictType=1&DistrictType=2&DistrictType=3&DistrictType=4&DistrictType=5&DistrictType=6&DistrictType=7&NumOfStudentsRange=more&NumOfSchoolsRange=more&ID2=4837650

[73] "Process of the Three-Tier RtI Model." Rockwall ISD. pp.1-2. http://cdn.rockwallisd.com/intranet/files/2012/01/Process_of_Three_Tier_RTI.pdf

Figure 2.1: Math Screening Tools

	Beginning of Year	Middle of Year	End of Year
Kindergarten	First Steps Diagnostic Tasks	First Steps Diagnostic Tasks	First Steps Diagnostic Tasks
Grade 1	First Steps Diagnostic Tasks	First Steps Diagnostic Tasks	First Steps Diagnostic Tasks
Grade 2	First Steps Diagnostic Tasks	First Steps Diagnostic Tasks Curriculum Based Assessment	First Steps Diagnostic Tasks Curriculum Based Assessment
Grade 3	First Steps Diagnostic Tasks District Developed Test	First Steps Diagnostic Tasks Curriculum Based Assessment	Did NOT Meet Expectations on Grade 3 STAAR Math
Grade 4	Did NOT Meet Expectations on Grade 3 TAKS Math	Curriculum Based Assessment	Did NOT Meet Expectations on Grade 4 STAAR Math
Grade 5	Did NOT Meet Expectations on Grade 4 TAKS Math **OR** Appear on 110% report in AWARE	Curriculum Based Assessment	Did NOT Meet Expectations on Grade 5 STAAR Math

Source: Rockwall Independent School District[74]

Huntington County Community School Corporation, IN

Huntington County Community School Corporation (HCCSC) is characterized as a rural fringe district, with a student population of 5,883 and a student/teacher ratio of 15.69.[75] RTI programs are coordinated out of the office of Special Programs, which also serves English Language Learners and homeless children.[76] HCCSC's RTI guidebook specifies that one of the goals of the program is to integrate and collaborate across a variety of educational initiatives "to ensure students are well prepared for their career and life."[77] HCCSC is explicitly opposed to the practice of tracking or grouping students on the basis of ability. The only exception is when groups are short term, and instruction is tied to a specific skill deficit.[78]

[74] "RISD Tier II Universal Screening: Math 2011-2012." Rockwall ISD. http://cdn.rockwallisd.com/intranet/files/2012/01/RtI-Math_Universal_Screening.pdf
[75] "Search for Public School Districts: Huntington Co Com Sch Corp." National Center for Education Statistics. http://nces.ed.gov/ccd/districtsearch/district_detail.asp?ID2=1804710&details=1
[76] "Special Programs." Huntington County Community School Corporation. http://specialprograms.hccsc.k12.in.us/modules/groups/integrated_home.phtml?&gid=1575917&sessionid=32026c9b8c6ad61f295c8c3798ffb6ce&t=#
[77] "Response to Intervention (RTI) Guidebook Summary." Huntington County Community School Corporation, 2009-2010. p.3. http://specialprograms.hccsc.k12.in.us/modules/locker/files/get_group_file.phtml?gid=1575917&fid=8019861&sessionid=32026c9b8c6ad61f295c8c3798ffb6ce
[78] Ibid., p. 5.

The HCCSC model consists of three tiers, each of which focuses on both academic and behavioral components. The academic component within Tier 1 provides all students with evidence-based instruction and screenings at least three times per year, including benchmark assessments.[79] HCCSC has identified two purposes for universal screening, the first of which is to assess core curriculum and instruction: if 80 percent of students are not successful on the universal screening, the problem is assumed to be with the curriculum or with the instruction. The second purpose is to identify students who need further interventions at Tier 2, which is implemented as soon as students show signs of falling behind peers in universal screens. At Tier 2, the appropriate intervention is selected from the HCCSC Standard Treatment Protocol menu based on screening data. Progress monitoring of student achievement begins at this tier. Students who show no improvement in Tier 2 are moved to Tier 3, where they are referred to an RTI Team. The RTI Team reviews all pertinent screening and progress monitoring data, and selects a Tier 3 intervention from the HCCSC Standard Treatment Protocol menu. The RTI Team becomes responsible for progress monitoring at this tier.[80]

HCCSC uses multiple types of academic screens at each grade level. The district does not explicitly distinguish between tools used solely for math or for a combination of math and literacy, so the full list of tools used is included in Figure 2.2. In addition, HCCSC uses the Pearson Inform Data Warehouse and Academic Intervention Plan to "house, manage, and manipulate all of the universal screening and progress monitoring data, as well as additional RTI documentation (i.e.: interventions, anecdotal notes, target groups, goals, etc.)."[81]

Figure 2.2: HCCSC Screening Tools

GRADES K-2	
▪ HCCSC Kindergarten Screening Tool ▪ Marie Clay Letter ID (K) ▪ DIBELS (K-1) ▪ Fountas & Pinnell Reading Benchmark	▪ ELA Standard 7 Rubric ▪ Quarterly Writing Prompts ▪ NWEA
GRADES 3-5	
▪ Fountas & Pinnell Reading Benchmark ▪ ELA Standard 7 Rubric	▪ Quarterly Writing Prompts ▪ NWEA

Source: HCCSC[82]

AUGUSTA COUNTY PUBLIC SCHOOLS, VA

Augusta County Public Schools is characterized as a rural fringe district with a student population of 10,743 and a student/teacher ratio of 13.[83] Within the district, RTI is referred to interchangeably as Response to Intervention or Response to Instruction. RTI programs

[79] Ibid., pp. 4, 6, 12
[80] Ibid., pp. 7-9.
[81] Ibid., pp. 31-32.
[82] Ibid., p. 32
[83] "Search for Public School Districts: Augusta Co Pblc Schs." National Center for Education Statistics. http://nces.ed.gov/ccd/districtsearch/district_detail.asp?ID2=5100300&details=1

are operated out of the Instructional Department, through which teachers, administrators, and other staff collaborate to develop school curricula and provide materials and instructional resources.[84]

The Augusta County Schools RTI framework consists of three levels. Level I, or Core Instruction, covers standard class curricula; poor performance on a Tier I screen triggers the move into Level II, or Strategic Instruction. At this level, students receive additional small group instruction targeted at specific areas of weakness. Finally at Level III, or Intensive Instruction, students receive small group instruction and individualized interventions. Progress monitoring using CBM measures is conducted beginning at Levels II and extends into Level III.[85]

Level I screenings are conducted three times a year. Kindergarten and grader 1 students are tested for numeracy proficiency in the areas of rote counting, missing number identification, number recognition, and quantity discrimination. Students in grades 1 through 5 are tested on basic computation skills using ProEd Monitoring Basic Skills Progress (MBSP) grade level assessments. Starting in grade 2, students are also assessed against locally generated math benchmarks.[86]

CAMPBELL COUNTY PUBLIC SCHOOLS, KY

Campbell County Public Schools is characterized as a rural fringe district with 4,987 students and a student/teacher ratio of 16.31.[87] The motto of Campbell County Schools is "Whatever it takes," and that philosophy guides RTI implementation in the district. Campbell County strives to ensure that all students are learning at a high level, and focuses RTI efforts on behavior, academics, and physical development areas such as hearing and vision. The district emphasizes that RTI is to be used for intervention and prevention, as opposed to disability identification and verification of special education eligibility. RTI planning and strategy teams operate at the building, grade, and student levels.[88] Core RTI principles include:

- We can effectively teach all children
- Intervene early
- Use a multi-tier model of service delivery

[84] "Instruction." Augusta County Public Schools. http://www.augusta.k12.va.us/Page/153
[85] Ibid., p. 2
[86] [1] "Mathematics Universal Screenings- Local Norms." Augusta County Public Schools, 2010. http://www.augusta.k12.va.us/cms/lib01/VA01000173/Centricity/Domain/22/UniversalScreeningNorms-Math-2010.pdf
[2] "Assessment Calendar for 2010-11." Augusta County Public Schools, 2010-2011. http://www.augusta.k12.va.us/cms/lib01/VA01000173/Centricity/Domain/22/Assessment_Record_2010-2011-2.xlsx
[87] "Search for Public School Districts: Campbell County." National Center for Education Statistics. http://nces.ed.gov/ccd/districtsearch/district_detail.asp?ID2=2100900&details=1
[88] "Response to Intervention Handbook 2013-2014." Campbell County Public Schools. pp. 2, 5-6. Available for download from http://www.campbellcountyschools.org/content_page.aspx?cid=338

- Use a problem-solving methodology
- Use research-based, scientifically validated interventions/instruction
- Monitor student progress to inform instruction
- Use data to make decisions
- Use assessments for three different purposes: (1) screening; (2) diagnostics; and (3) progress monitoring[89]

Campbell's RTI model is structured using the traditional three-level framework described in Section I. Level 1 consists of core instruction, and universal screening is conducted three times a year. Students begin to receive regular progress monitoring at Level 2, which provides core and supplemental instruction to small groups of no more than six to eight students. Finally, Level 3 provides core, supplemental and/or intensive instruction to smaller groups of up to three students. Students who continue to struggle at Level 3 are referred for special education assessment.[90]

Universal RTI screening is conducted in the fall, winter and spring using Measures of Academic Progress (MAP) to benchmark student achievement. All regular education teachers, special education teachers, intervention teachers, and para-educators have been trained to administer and interpret assessments.[91] Subsequent progress monitoring uses CBM measures. In kindergarten and first grade, math monitoring focuses on oral counting, number identification, quantity discrimination, and missing number identification, and is conducted using Intervention Central NumberFly Customizable probes. AIMSweb probes are used to monitor skill acquisition in math computation as well as math concepts and applications in all grades beyond kindergarten.[92]

HAYS CONSOLIDATED INDEPENDENT SCHOOL DISTRICT, TX

Hays Consolidated Independent School District is characterized as a rural fringe district with 15,932 students and a student/teacher ratio of 15.84.[93] RTI services in the district are designed to address both academic and behavioral issues.[94] Intervention programs are managed by the office of Intervention Services, whose mission is to assist all students with special needs.[95]

Hays CISD applies a three-tier tier framework to RTI services. In Tier 1, core classroom instruction is aligned with Texas Essential Knowledge and Skills (TEKS). The district

[89] Taken verbatim from "RTI-Response to Intervention." Campbell County Public Schools. http://www.campbellcountyschools.org/content_page.aspx?cid=338
[90] "Response to Intervention Handbook 2013-2014," Op. Cit., pp. 5-6.
[91] "RTI-Response to Intervention," Op. Cit.
[92] "Response to Intervention Handbook 2013-2014," Op. Cit., p. 20.
[93] "Search for Public School Districts: Hays CISD." National Center for Education Statistics. http://nces.ed.gov/ccd/districtsearch/district_detail.asp?ID2=4800010&details=1
[94] "Response to Intervention." Hays Consolidated ISD. http://www.hayscisd.net/district.cfm?subpage=9838
[95] "Intervention Services." Hays Consolidated ISD. http://www.hayscisd.net/district.cfm?subpage=9828

emphasizes that it uses data-based progress monitoring to determine who requires additional instruction while noting that approximately 80 percent of students are successful in this tier.[96] AIMSWeb screenings are given three times during the year,[97] and students who perform "slightly below average" are given classroom interventions with biweekly progress monitoring. Students who do not respond to these class-based interventions are moved up to Tier 2 to receive diagnostic testing, more personalized intervention strategies, and continued progress monitoring. Finally, those who do not respond to Tiers 1 or 2 are provided with "specific, custom-designed individual or small group instruction" designed to assist students with "identified difficulties in academics or behavior."[98]

[96] "Response to Intervention," Op. Cit.
[97] "Student Assessment." Hays Consolidated ISD. http://www.hayscisd.net/district.cfm?subpage=320
[98] "Response to Intervention," Op. Cit.

Project Evaluation Form

Hanover Research is committed to providing a work product that meets or exceeds partner expectations. In keeping with that goal, we would like to hear your opinions regarding our reports. Feedback is critically important and serves as the strongest mechanism by which we tailor our research to your organization. When you have had a chance to evaluate this report, please take a moment to fill out the following questionnaire.

http://www.hanoverresearch.com/evaluation/index.php

Caveat