



Professional Development Guide

Assessment

Authors **Doug Fuchs**
Lynn Fuchs



Columbus, Ohio

Photo Credits: Cover ©SRA File Photos; **1** ©Will & Deni McIntyre/CORBIS; **3** © Robert E. Daemmrich; **4** ©Index Stock/Alamy; **5** ©Jim Cummins/FPG International/Getty Images, Inc.; **6** ©Will Hart/PhotoEdit, Inc.; **10** ©Rob Lewine/CORBIS; ©Jim Whitmer/Stock Boston; **14** ©Robert E. Daemmrich; **16** ©Stewart Cohen/Stone/Getty Images, Inc..

SRAonline.com



Copyright © 2008 by SRA/McGraw-Hill.

All rights reserved. No part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written consent of The McGraw-Hill Companies, Inc., including, but not limited to, network storage or transmission, or broadcast for distance learning. An Open Court Curriculum.

Printed in the United States of America.

Send all inquiries to this address:
SRA/McGraw-Hill
4400 Easton Commons
Columbus, OH 43219

ISBN: 978-0-07-606213-3
MHID: 0-07-606213-9

1 2 3 4 5 6 7 8 9 MAZ 13 12 11 10 09 08 07

Table of Contents

| | |
|------------------------------------------------------------------------------------------------------|-----------|
| Classroom Assessment to Improve Reading Instruction..... | 1 |
| The <i>SRA Imagine It!</i> Conceptual Framework for Assessment..... | 2 |
| General Outcome Measurement | 2 |
| The Curriculum Sampling Approach..... | 2 |
| The Performance Indicator Approach | 3 |
| Three Types of Formal Classroom Assessment..... | 5 |
| Screening | 5 |
| Progress Monitoring..... | 6 |
| Diagnosis..... | 8 |
| The <i>SRA Imagine It!</i> Assessment System | 8 |
| Benchmark Assessments | 8 |
| Lesson Assessments | 9 |
| How <i>SRA Imagine It!</i> Incorporates Both Approaches to Assessment..... | 10 |
| How the <i>SRA Imagine It!</i> Assessment System Addresses the Three Purposes of Assessment | 10 |
| Weekly Progress Monitoring for At-Risk Students..... | 13 |
| A Model for Linking Assessment with Instruction Using a Multitier Prevention System..... | 14 |
| Appendix A | 18 |
| Appendix B | 20 |
| Appendix C | 21 |
| Appendix D | 22 |

Assessment

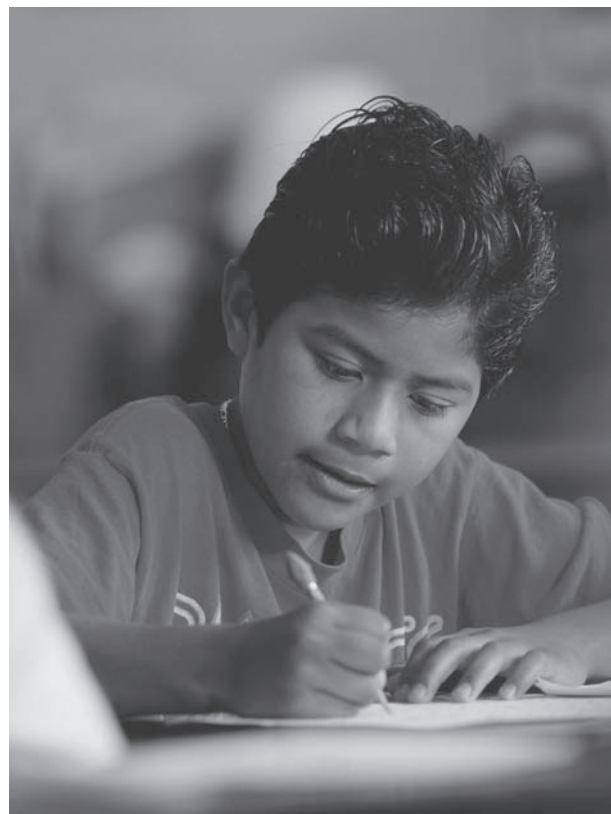
Classroom Assessment to Improve Reading Instruction

Designing an effective reading program for each and every student is one of the most important tasks within education. Reading is essential for learning across all the content areas in school and over the course of a person's entire life. The assessment of students' reading ability is therefore critical. Assessment provides teachers with an accurate understanding of how well a student reads, whether the student is at risk for reading problems in the future, and the degree to which the student has mastered the various component skills that underlie competent reading. This understanding allows teachers to plan reading instruction that matches each student's specific needs and helps the teacher provide the student with appropriate reading material for reading instruction, for recreational reading, and for content-area instruction. The more teachers know about their students' reading abilities, the more likely they will

Assessment provides teachers with an accurate understanding of how well a student reads, whether the student is at risk for reading problems in the future, and the degree to which the student has mastered the various component skills that underlie competent reading.

be able to promote the reading outcomes that make learning in all subject areas successful.

To this end, *SRA Imagine It!* provides a comprehensive, integrated formal system to assess language arts development over the course of the school year. The purpose of this guide is to explain the rationale and structure behind that comprehensive, integrated formal assessment system. With formal classroom assessment, a teacher administers objective tests to measure student abilities, skills, and strategies. Schools use such assessment for four purposes: screening, progress monitoring, diagnosis, and outcome evaluation. The focus of this guide, is restricted to screening, progress monitoring, and diagnosis because these assessment



purposes are most pertinent to the instructional planning decisions teachers make in their classrooms.

This guide begins by explaining the conceptual framework behind the ***SRA Imagine It!*** formal assessment system, including a description of general outcome measurement and an overview of the three major purposes of formal classroom reading assessment. The guide then describes the structure of the ***SRA Imagine It!*** Assessment System and how it fits into that conceptual framework. Finally, it discusses a framework for linking assessment with instruction within a multitier prevention system.

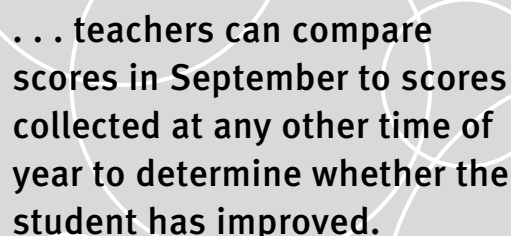
The ***SRA Imagine It!*** Conceptual Framework for Assessment

General Outcome Measurement

General outcome measurement is a form of classroom assessment whereby students' overall competence in the relevant academic domain, such as reading, is measured. General outcome measurement is standardized so that the behaviors to be measured and the procedures for measuring those behaviors are prescribed. In addition, the testing methods and content of general outcome measurement remain constant over the course of an academic year. That is, each test (whether tests are administered once at the beginning of the academic year, periodically over the course of the year, or even on a weekly basis) is of equivalent difficulty and represents competence in the entire school year's curriculum. The primary reason for long-term consistency

is so teachers can compare scores in September to scores collected at any other time of year to determine whether the student has improved.

There are two basic approaches to general outcome measurement.



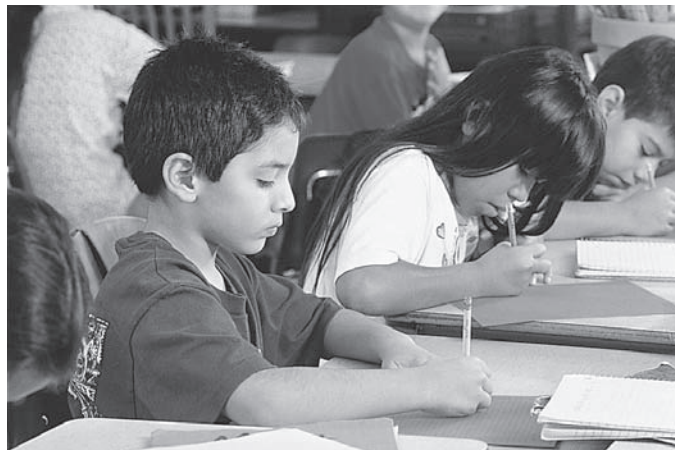
... teachers can compare scores in September to scores collected at any other time of year to determine whether the student has improved.

The Curriculum–Sampling Approach

The first approach to general outcome measurement is called the *curriculum-sampling approach*. With the curriculum-sampling approach, the skills that constitute the annual curriculum are specified, and a set of procedures for measuring each skill in that annual curriculum is designed. Then, the proportion of each item type is determined so that each test samples the various skills to reflect the relative importance of each skill. Each alternate form of the test is created to sample the skills in the same way. That way, the difficulty and content of the test remain constant, and we expect a student's scores to gradually increase over the course of the year as the teacher addresses the skills embedded in the curriculum and as the student gains increasing mastery and competence.

For example, a vocabulary strand might feature questions pertaining to affixes, Greek and Latin roots, synonyms, and

multiple-meaning words. As the student becomes more familiar with these word structure concepts, his or her score will increase while the difficulty of the test remains constant.



The Performance Indicator Approach

The second approach to general outcome measurement is called the *performance indicator approach*. With the performance indicator approach, a single task for measurement is selected because that task correlates highly with the various skills addressed in the annual curriculum. As with the curriculum-sampling approach, the content and difficulty of the task remains constant across the school year although the testing material differs across alternate forms of the test.

An example of the performance indicator approach to classroom assessment is passage reading fluency. With passage reading fluency, a student reads aloud a different passage on each testing occasion. However, each passage represents the difficulty level associated with the end of the school year. Each time, the student reads aloud for one minute; the teacher counts the number of

words read correctly. Although this is a direct measure of oral reading fluency, it also functions well as an indicator of a student's overall reading competence. That is, students who score poorly when asked to quickly read grade-level text aloud are the same students who have poor decoding skill, whose ability to recognize words automatically is problematic, who have limited vocabularies, and who have problems understanding what they read. We expect the student's score on a performance indicator general outcome measure to gradually increase over the course of the year, showing the student's increasing competence in the curriculum. Also, how students perform when asked to read grade-level text aloud in a fixed amount of time has been shown to predict their long-term reading ability as measured on other valued tests.

Measures have been validated to accurately reflect overall reading competence, and as would be expected, those measures differ by grade level. At kindergarten, the major alternatives for the performance indicator approach reading measures are phoneme segmentation fluency, rapid letter naming, and letter-sound fluency. With phoneme segmentation fluency, the examiner says a word; the student says its constituent sounds. The examiner provides as many stimuli within one minute as the rate of the child's response permits. With rapid letter naming, the examiner presents a page of lower- and upper-case letters randomly ordered; the student says as many letter names as he or she can in one minute. With letter-sound fluency, the examiner also presents a page with lower- and upper-case letters randomly ordered; this time, however, the student says sounds

for one minute. Compared to phoneme segmentation fluency, rapid letter naming and letter-sound fluency are easier for practitioners to learn to administer, and their reliability tends to be stronger. On the other hand, letter-sound fluency is a better target for instruction than rapid letter naming because it relates more transparently to what children need to learn to read.

At first grade, two performance indicator reading measures have been studied. One approach involves combining nonsense word fluency and passage reading fluency: Students begin the year on nonsense word fluency and move to the more difficult performance indicator, passage reading fluency, at midyear. With nonsense word fluency, students are presented with a page of consonant-vowel-consonant (with some vowel-consonant) pseudowords and have one minute to decode as many as they can. With passage reading fluency, students are presented with grade-level text (each alternate form is a passage of roughly equivalent difficulty), and students read aloud for one minute. The second approach uses a constant measure across all of first grade: word identification fluency, where students are presented with a page showing one hundred high-frequency words (each alternate form samples words randomly from a pool of grade level words); students read as many words as possible in one minute.

The advantage of nonsense word fluency is that it maps onto beginning decoding instruction, potentially providing teachers with input for instructional planning. The downside of the nonsense word fluency/passage reading fluency combination is that getting a good picture of development

over the course of first grade is problematic because teachers cannot compare scores collected in the first half of the year (with nonsense word fluency) with scores collected after passage reading fluency begins in the second half of the year. By contrast, word identification fluency can be used with strong reliability, validity, and instructional usefulness across the entire first-grade year. This also makes it possible to get a good picture of a student's reading development across the entire time frame.



At Grades 2–4, the passage reading fluency measure provides the strongest source of information on reading development as a form of performance indicator general outcome measurement. Each week, one test is administered, with the student reading aloud from a different but equivalent passage for one minute; the examiner counts the number of words read correctly within the one-minute time frame. The reliability, validity, and instructional utility of this simple measure have been demonstrated repeatedly. Some teachers can think of students whose comprehension is weak despite good fluency. It is important to note that these students are so memorable because they are so few and far between. In the vast majority of cases, students who have

strong fluency are the same ones who demonstrate good comprehension.

Some research indicates that the validity of the passage reading fluency performance indicator begins to decrease somewhere around Grade 5. So, beginning in the fifth grade (or in Grade 4 for high-performing schools), teachers should consider using a different measure that more directly taps comprehension.

In the vast majority of cases, students who have strong fluency are the same students who demonstrate good comprehension.

One alternative for the higher grades is maze fluency. With maze fluency, students are presented with a passage from which approximately every seventh word has been deleted and replaced with three possible replacements, only one of which is semantically tenable. The student has three minutes to read and replace blanks, and the score is the number of correct replacements. Some research indicates that maze fluency demonstrates strong reliability and validity and models reading development beginning at Grade 4 and continuing through the eighth grade.

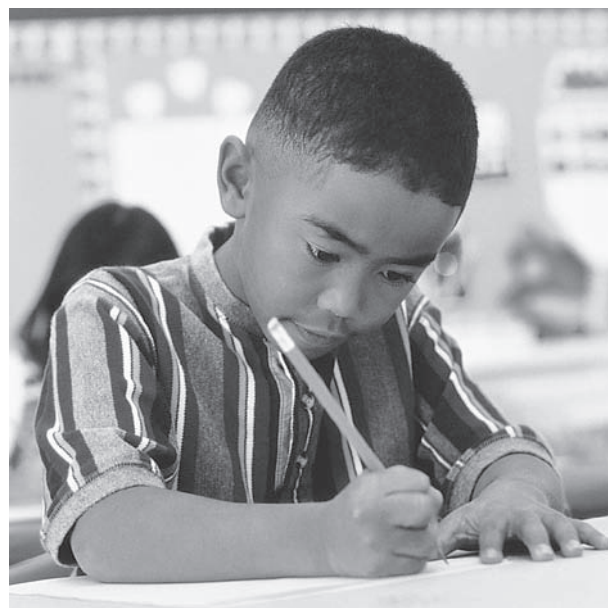
Three Types of Formal Classroom Assessment

SRA Imagine It! relies on general outcome measurement as the assessment framework, and it combines the curriculum-

sampling approach with the performance indicator approach. *SRA Imagine It!*, uses general outcome measurement for three purposes: screening, progress monitoring, and diagnosis.

Screening

Screening is the process of measuring all students in a class to identify the subset of students who, without special attention, are in danger of scoring poorly on the end-of-year high-stakes tests and long-term reading failure. With screening, students are identified as having potential problems early, well before the end of the year; therefore, teachers can allocate special attention to these students as quickly as possible. The hope is that the teacher's early attention will improve the long-term outcomes of these students, avoiding the chronic and severe difficulties that students with reading disabilities face throughout school and in their adult lives. Once a screening assessment has been selected, a cut-point is determined. When a student's score falls below the cut-point, the student



is suspected to be at risk of poor long-term reading outcomes. Performing below the cut-point signals the teacher to pay extra attention to this student, with more frequent progress monitoring and with small-group instruction to remediate the student's difficulties.

Some schools or programs rely on screening at a single point in time, usually at the beginning of a school year, so that students suspected to be at risk are identified early and can be monitored and receive special attention throughout the school year. Other schools or programs rely on periodic screening, which occurs at regular intervals across the school year. When periodic screening is used, an alternate form of the screening test is administered at each screening. However, the difficulty and nature of the screening instrument remain the same across the year. So the screening test administered in September is an alternate form representing the same difficulty and content as the screening test administered in November or January or March or May. When periodic screening is used, the cut-point on each subsequent screening gradually increases. So the cut-point for what indicates risk is lower in September than in November, and so on.

Performing below the cut-point signals the teacher to pay extra attention . . . with more frequent progress monitoring and with small-group instruction to remediate the student's difficulties.

Progress Monitoring

In common usage, the term *progress monitoring* may refer to any type of ongoing classroom assessment, informal or formal. ***SRA Imagine It!*** however, uses the term *progress monitoring* in a stricter sense to refer specifically to systematic formal assessments in which students are assessed on a regular basis. Teachers can use results from progress monitoring assessments (a) to formulate decisions about how to make classroom-level instruction more responsive to individual student needs, (b) to determine whether a student is responding adequately to the instructional program, and (c) for students who are unresponsive to a validated or research-based instructional program, to inductively design individualized instructional plans.

In most reading programs, progress monitoring is accomplished via mastery measurement. ***SRA Imagine It!*** does *not* use mastery measurement, but it takes a few minutes to explain mastery measurement to distinguish it from ***SRA Imagine It!***'s general outcome measurement approach. With mastery measurement, teachers assess mastery of a sequence of skills. Designing a mastery measurement progress-monitoring system requires two major tasks. The first requires determining the hierarchy of skills for instruction. For example, with first-grade reading, one might specify the following sequence of skills: letter-sound correspondence, decoding phonemically regular consonant-vowel-consonant words, decoding phonemically regular consonant-vowel-consonant-final *e* words, automatically recognizing the one hundred most frequent Dolch words, decoding

phonemically regular *r*-controlled words, and so on. The second major task in designing a mastery measurement progress-monitoring system is to design a criterion-referenced test for each skill in the instructional hierarchy. So, for example, for letter-sound correspondence, the criterion-referenced test might involve presenting students with a shuffled deck of letters, one letter at a time for one minute; the student responds by saying the sounds associated with the letters; the score is the number correct. The criterion for mastery might be twenty-six letter sounds in one minute on two consecutive tests. The test might be administered once weekly, and when the student achieves the mastery criterion, instruction and testing simultaneously shift to the next skill in the hierarchy, decoding phonemically regular consonant-vowel-consonant words, for which a criterion-referenced test is also designed. Most classroom-based reading assessments fall into the category of mastery measurement.

Most basal reading assessments provide unit tests to assess mastery of skills addressed in each unit. Nevertheless, research indicates that few classroom teachers adhere to mastery rules, based

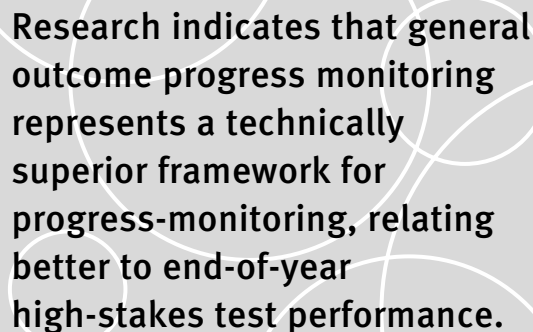


on those tests, for advancing students to new instructional content. In addition, some technical problems plague mastery measurement systems.

For example, it is possible that achievement as demonstrated on the mastery measurement system (that is, mastery of many skills across a school year) fails to relate well to performance at the end of the year on high-stakes tests. Poor correspondence, with some children who have mastered many skills performing surprisingly poorly on the high-stakes end-of-year testing, can occur because of poor retention of previously mastered skills. For example, a student might master consonant-vowel-consonant words, but when instruction and testing simultaneously shift to the next skill in the hierarchy, the student's accuracy in decoding consonant-vowel-consonant words deteriorates. Another source of poor correspondence between the number of skills mastered and final performance on the high-stakes test is mastery measurement's reliance on a *single-skill* testing framework. Some students can perform a skill competently only because they know that every item on the test represents an example of the target skill. For example, some students decode consonant-vowel-consonant words accurately when all the words on the test represent that pattern; they know to incorporate the short vowel for every word on the test. Some subset of these students cannot, however, decode consonant-vowel-consonant words when they are presented in mixed-skill fashion along with words of a variety of phonetic categories. In contrast to mastery measurement, high-stakes tests do not rely on single-skill measurement.

Hence the potential exists for a lack of correspondence between performance on mastery measurement's single-skill testing framework and high-stakes tests' sampling across many skills.

For these and other reasons, *SRA Imagine It!* relies on general outcome measurement for progress monitoring. As already explained, general outcome measurement simultaneously assesses performance across the many skills represented in the annual curriculum, either via systematic curriculum sampling or by relying on a performance indicator of overall competence. Research indicates that general outcome progress monitoring represents a technically superior framework for progress-monitoring, relating better to end-of-year high-stakes test performance.



Research indicates that general outcome progress monitoring represents a technically superior framework for progress-monitoring, relating better to end-of-year high-stakes test performance.

Diagnosis

This guide, defines diagnosis to mean assessment that describes a student's strengths and weaknesses with respect to skills or strategies. The goal is to identify productive targets for instruction. Hence, if a diagnostic assessment instrument indicates that the student is strong in decoding but lacks fluency, the teacher

would direct the instruction toward fluency. If a diagnostic assessment instrument indicates that the student decodes *r*-controlled words accurately but cannot decode two-syllable words with an open first syllable, the teacher would direct instruction toward this type of two-syllable word. A curriculum-sampling approach to general outcome measurement provides teachers with a strong basis for describing students' strengths and weaknesses in the curriculum because all the skills embedded in the annual curriculum are assessed on each testing occasion.

The *SRA Imagine It!* Assessment System

The *SRA Imagine It!* formal assessment system includes two major components: the Benchmark Assessments and the Lesson Assessments.

Benchmark Assessments

Description

At each grade level, there are periodic Benchmark Assessments. At kindergarten and first grade, there are six Benchmark Assessments, one to be administered at the beginning of the school year and one to be administered after every other unit for the remainder of the school year. At the other grades, there are seven Benchmark Assessments, one to be administered at the beginning of the year and one to be administered at the end of each unit.

Within a grade level, each Benchmark Assessment over the course of the year is of equivalent difficulty, and each one

samples the entire year's curriculum. The specific test items on the Benchmarks vary from assessment to assessment, but each Benchmark is equivalent to the others within the grade level in terms of the content covered, the question formats, and difficulty. This equivalency means that rising scores reflect increased student competence in the curriculum. These assessments differ from traditional assessments that (a) cover only the material in the most recently taught unit and (b) consequently become increasingly difficult over the course of a school year.

Benchmark Assessments include the following components.

... each Benchmark is equivalent to the others within the grade level in terms of the content covered, the question formats, and difficulty. This equivalency means that rising scores reflect increased student competence in the curriculum.

100-Point Skills Battery

Students answer questions relating to comprehension, vocabulary, grammar, usage, and mechanics, spelling, phonics, and phonemic/phonological awareness. Each grade level has a 100-Point Skills Battery that is weighted and administered according to that grade level's focus. For example, Grades 4–6 are group-administered and feature 20 percent of the total score in the Grammar, Usage, and Mechanics section whereas Grade 1 has

both individual and group-administered assessments and apportion only 5 percent of the total score to Grammar, Usage, and Mechanics but assign Phonics and Phonemic/Phonological Awareness 30 percent of the total score.

Fluency Assessment

The second component is the Fluency Assessment, which is a global indicator of a student's reading ability. At Grades 5 and 6, teachers have the option of using passage reading or a group-administered pencil-and-paper maze task to measure fluency. (See Appendix A for examples of both measures.) For Grades 2–4, students' oral fluency is measured by passage reading. At first grade, word identification fluency is used, and letter-sound fluency is used at kindergarten.

Writing Assessment

In Grades 2–6, a writing assessment is included with the first, fourth, and seventh Benchmark Assessments. Each writing assessment includes an expository prompt and a checklist for students. (See Appendix B for an example of a prompt.) These prompts are similar to the type of prompt found in high-stakes tests.

Lesson Assessments

Description

As students complete each lesson, they will be assessed on their understanding of the instructional content and the literature in each lesson. The results of the assessments will then be used to inform subsequent instruction. How students score on the assessments offers a picture

of current student achievement status while also guiding you toward appropriate instructional decisions.

The Lesson Assessments are in a structure similar to that of the Benchmark Assessments in terms of assessment sections that equal 100 points, a fluency assessment, and a writing assessment.

How students score on the assessments offers a picture of current student achievement status while also guiding you toward appropriate instructional decisions.

How *SRA Imagine It!* Incorporates Both Approaches to Assessment

Both the curriculum-sampling and the performance indicator approaches to assessment (described above) have important strengths, and *SRA Imagine It!* capitalizes on those strengths by incorporating both assessment approaches. The Benchmark Assessments incorporate the curriculum-sampling approach with the 100-Point Skills Battery. The Benchmark and Lesson Assessments also incorporate the performance indicator approach with their Fluency Assessments.

For the Benchmark 100-Point Skills Battery, critical skills that constitute the program have been identified; the relative importance of each skill has been determined within each area of the language arts curriculum. For each skill,

assessment items were designed. Then the items were organized into six or seven equivalent tests, so that the number of items for each skill on each of the six or seven equivalent tests reflects the relative importance of that skill in the annual curriculum. The 100-Point Skills Battery provides an overall indicator of a student's performance with the total score on the 100-Point Skills Battery. However, because the 100-Point Skills Battery includes a separate section on each major strand within the curriculum, it also gives teachers information about students' performance in specific areas of the curriculum. This type of detailed information is not available from Fluency Assessments alone. Another important advantage is that the 100-Point Skills Battery mirrors the format of end-of-year tests. This can help students learn strategies for performing well on those high-stakes tests even as the curriculum-sampling screener helps the teacher design instruction to emphasize the various skills that constitute her reading program and appear on the high-stakes tests.



For the Benchmark Fluency Assessment, a single performance indicator assessment has been selected at each grade level, based on the strongest research available, and six or seven equivalent Fluency Assessments (depending on grade level) have been created, one for each alternate form of the Benchmark Assessments. The specific type of fluency measure used varies across the grades.

At kindergarten, letter-sound fluency is used because it is easy for teachers to learn to administer, it is reliable, and also it can provide guidance for the teacher's instructional behavior. At first grade, *SRA Imagine It!* uses word-identification fluency because it has strong reliability, validity, and instructional usefulness across the entire first-grade year. In Grades 2–4, *SRA Imagine It!* incorporates passage-reading fluency as its ongoing progress-monitoring measure because of the body of research that supports its reliability, validity, and instructional utility. *SRA Imagine It!* provides maze fluency as its ongoing progress-monitoring measure at Grades 5–6 because research indicates that at the upper grades it may be a better indicator of overall reading ability than oral reading fluency. Because the maze test is a pencil-and-paper measure, it can also be administered in a group setting, which is easier for schools in which students rotate between classrooms for different subjects.

The Lesson Assessments feature performance indicator passage reading fluency assessments in the second half of Grade 1 and throughout Grades 2–6. These assessments allow the teacher to monitor student fluency more frequently than then the Benchmark Fluency

Assessments—eighteen opportunities in Grade 1 and thirty opportunities in Grades 2–6.



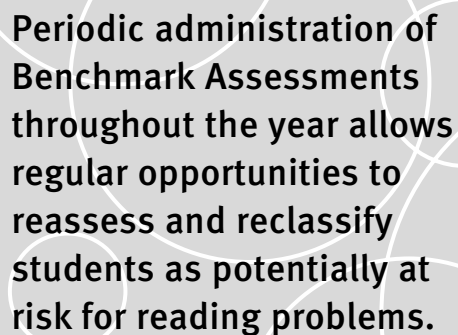
How the *SRA Imagine It!* Assessment System Addresses the Three Purposes of Assessment

Screening

With respect to screening, the six or seven (depending on grade level) alternate forms of the Benchmark Assessments allow them to be used as a periodic screening tool. For each Benchmark Assessment at each grade level cutoff scores for the 100-Point Skills Battery and the Fluency Assessment are provided. (See Appendix C for cutoff scores.) Any student who falls below the cutoff on either of these measures can be classified as potentially at risk for reading problems. This early screening allows teachers to identify children with potential problems very early in the school year, permitting them to focus special attention on those students from the beginning.

Periodic administration of Benchmark Assessments throughout the year allows

regular opportunities to reassess and reclassify students as potentially at risk for reading problems. A student who scores above the cut-point at the beginning of the year but then scores below the cut-point later will be brought to the teacher's attention. On the other hand, a student who begins the school year with performance below the cut-point would exit the risk category at that time, freeing up the teacher's attention for other students who require it. This permits teachers to allocate special attention in a manner that reflects each individual student's response to the instructional program. In this way, the Benchmark Assessments are essentially a periodic screening tool to be used throughout the school year.



Periodic administration of Benchmark Assessments throughout the year allows regular opportunities to reassess and reclassify students as potentially at risk for reading problems.

Progress Monitoring

At each grade, the Benchmark Assessments include six or seven alternate forms to be administered at the beginning of the school year and at regular intervals throughout the remainder of the school year.

Each Benchmark samples the entire year-long curriculum in the different components of the language arts program, including reading

comprehension, vocabulary, spelling, word structure, and at the lower grades in phonics and phonemic awareness. Because the tests are equivalent in difficulty and content, a student's scores across the year can be compared. Improving scores indicate increasing mastery of the curriculum. (See Appendix D for an example of a Benchmark Tracking Chart.) The overall score indicates increasing overall mastery, but the teacher can also use scores within specific curriculum areas to look for improvement in those specific areas as well as in overall progress. When the Benchmark Assessments indicate that a student is potentially at risk for reading difficulties, the teacher will monitor the progress of the student more closely.

One way to do this is to measure student achievement in the Lesson Assessments. These assessments reflect the instructional content taught during a particular week. For example, a Grade 4 student whose comprehension scores on a Benchmark test are low will encounter five comprehension assessments in the Lesson Assessments before another Benchmark test is administered. The information from the Lesson Assessments can be a useful tool in reinforcing an opinion of a student's at-risk status or it can provide data of growth and greater comfort within instructional strands. Students' performances in the Lesson Assessments also offer data of student achievement independent of the Benchmark Assessments. Students who perform below the recommended levels in the 100-point scale, the Oral Fluency Assessments and the writing assessments should be monitored and presented with opportunities for additional support.

... information from the Lesson Assessments can be a useful tool in reinforcing an opinion of a student's at-risk status or it can provide data of growth and greater comfort within instructional strands.

Diagnosis

Lesson Assessments and Benchmark Assessments also incorporate diagnostic assessment. Besides providing an overall indicator of a student's performance in the language arts curriculum, each 100-Point Skills Battery in a Benchmark Assessment provides the teacher with a breakdown of each student's performance on each major component of the curriculum. Teachers can use this skill profile to determine which aspects of the reading program are important to target for the class and which aspects of the reading program require attention for individual or small groups of students.

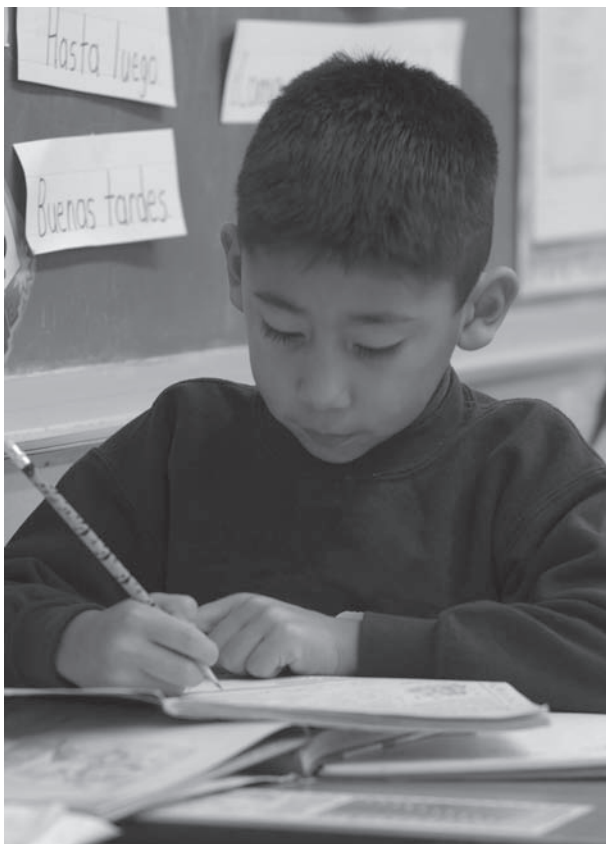
Students' achievement in the components of the Lesson Assessments helps determine which students need additional skill instruction. That additional instruction may be **Reteach** support for students who would benefit from further practice with a skill or concept. It could involve another intervention program for students who have difficulty understanding what is required of them. Additionally, the Lesson Assessments can help identify students who continually score above the recommended performance levels and might benefit from **Challenge Activities**

exercises to push their abilities even further.

Weekly Progress Monitoring for At-Risk Students

Strategies using fluency-based performance indicators for characterizing reading growth have been shown to be more sensitive for detecting student improvement than other classroom reading tests. Also, these assessments can be easily administered. For these reasons, **SRA Imagine It!** provides enough opportunities for fluency assessments so that an Oral Fluency Assessment can be used to conduct weekly progress monitoring. This is strongly recommended for the subset of students suspected to be at risk. Weekly progress monitoring for these students can help teachers gain insight about the effectiveness of their attempts at remediation for these at-risk students. Although not all students in the class receive progress monitoring, the periodic screening on the Benchmark Assessments helps teachers pick up students who are falling into the risk category. When that happens, those students then begin to receive weekly progress monitoring. It is important to note that after a student is first designated as at possible risk through periodic screening, then student should continue to receive weekly progress monitoring for the remainder of the year (even if the student scores above the risk cut-point on a subsequent screening).

With weekly progress monitoring, a student's scores are graphed against time, and a line of best fit is drawn through the graphed points. For this line of best fit, a slope (i.e., the weekly rate of increase) is calculated to quantify the rate of learning.



Teachers can use the progress-monitoring slope, along with the student's current score, for three purposes: (a) to formulate decisions about how to make classroom-level instruction more responsive to individual student needs, (b) to determine whether a student is responding adequately to the instructional program, and (c) for students who are unresponsive to a validated or research-based instructional program, to inductively design individualized instructional plans.

Weekly progress monitoring is important for at-risk students. It has been shown to reveal student improvement irrespective of the teacher's approach to reading instruction. Also, weekly progress monitoring using performance indicators is independent of the instructional methods used. Therefore, teachers can incorporate a wide range of instructional

methods, including, for example, decoding instruction, repeated readings, vocabulary instruction, story grammar exercises, and semantic-mapping activities. This permits teachers to experiment with instructional adaptations or major program revisions for their at-risk students in an attempt to effect progress for students for whom the standard reading program proves ineffective. Research shows that when teachers use weekly progress monitoring to help them tailor their instruction to match students' needs, it results in a better instructional program and has a positive effect on student achievement.

Weekly progress monitoring has been shown to reveal student improvement irrespective of the teacher's approach to reading instruction.

When special attention by the classroom teacher or small-group tutoring fails to enhance reading performance, schools need to consider the possibility that the student requires an individually tailored program to meet his or her special needs.

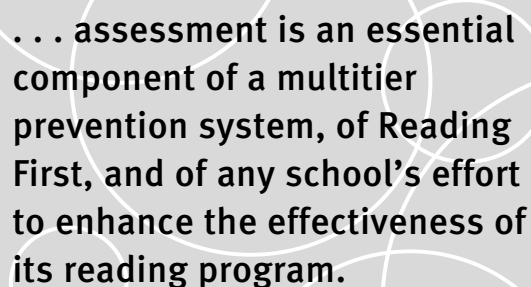
A Model for Linking Assessment with Instruction Using a Multitier Prevention System

Reading First has been contextualized as a multitier prevention system that typically incorporates three tiers, with each tier providing increasingly more intensive instruction. General education is the first tier. The second tier is typically

accomplished with a standard, validated, small-group tutoring protocol. When this tutoring proves unsuccessful for a given student, then a third tier of intervention is implemented. This third tier is typically individualized to meet the student's special learning needs.

Why is a multitier prevention system important? It is important because no instructional method, even those validated using stringent experimental studies, works for all students. For this reason, as schools implement validated interventions, even those that have been scientifically validated, the effects of those interventions on children's academic performance must be assessed. This is why assessment is an essential component of a multitier prevention system, of Reading First, and of any school's effort to enhance the effectiveness of its reading program. With systematic assessment that is integrated in a meaningful way with the life of a reading instruction program, children who do not respond adequately to the classroom program can be identified promptly for a second tier of more intensive intervention. For students who fail to respond to a second, more intensive level of programming, a third tier of instruction, with greater individualization, is implemented, and response continues to be assessed. This iterative process, with which interventions of increasing intensity and individualization are conducted and their effects are continuously assessed, constitutes a multitier prevention system. Within a multitier prevention system, therefore, assessment plays three important roles: (a) identifying students who should be targeted for attention, (b) among those targeted for

attention, quantifying responsiveness to intervention, and (c) for the most unresponsive subset of children, tailoring individualized instructional programs.



... assessment is an essential component of a multitier prevention system, of Reading First, and of any school's effort to enhance the effectiveness of its reading program.

Within this scheme, the first assessment function is identifying a subset of the school population that is suspected to be at risk for poor reading outcomes. These students become the focus of the multitier prevention system. As described earlier, this function is screening, where an initial assessment is administered to all students in a school or within targeted grade levels (usually the primary grades) in a school.

A screening cut-score is applied, and it specifies what screening score is associated with inadequate performance on a valued outcome measure, such as a high-stakes test, at a later time. All students scoring below this cut-score are designated as at risk for poor outcome. However, a one-time assessment for the purpose of screening carries a significant danger of identifying students for tutoring when, in fact, those students would go on to develop strong academic skills without tutoring. It can also miss students who in fact require that tutoring.

Because a one-time screening assessment, especially at kindergarten and first grade, typically makes too many errors of these

types, we suggest that the initial screening assessment constitute only the first step in the process of designating risk status. That is, we recommend that students who are first *suspected* to be at risk based on the first Benchmark Assessment be followed with six to eight weeks of progress monitoring while Tier 1 general education is implemented. The purpose of this short-term progress monitoring is to gauge response to general education and thereby confirm that the suspected risk, based on Benchmark screening, probably constitutes actual risk. Such short-term progress monitoring has been shown to increase the precision of designating who requires a Tier 2 intervention and therefore avoids wasting costly Tier 2 services on students whose academic skills would develop nicely without that special intervention. Of course, best practice means using progress-monitoring and diagnostic assessment even within Tier 1 toward the goal of differentiating instruction.

Within a multitier prevention system, a second purpose for progress monitoring occurs at Tier 2, where tutoring is based on standard treatment protocols. With a standard treatment protocol, a validated or research-based approach to tutoring is implemented in small groups. The assumption is that a vast majority of students should respond well to the validated or research-based small-group tutoring.

If a child responds poorly to tutoring that benefits most students, then the responsiveness-to-instruction assessment process eliminates instructional quality as a viable explanation for poor academic growth and instead provides evidence of a disability. The purpose of weekly progress monitoring at Tier 2 is to quantify student

progress over the course of tutoring—that is, to determine which students do and which students do not respond to validated small-group tutoring. The students who fare well (i.e., whose weekly progress-monitoring slopes of improvement indicate adequate progress) are returned to Tier 1. There, weekly progress monitoring is continued to assess whether the Tier 2 tutoring “inoculated” the student against further learning difficulties or whether the student instead requires another round of Tier 2 tutoring.

... short-term progress monitoring has been shown to increase the precision of designating who requires a Tier 2 intervention and therefore avoids wasting costly Tier 2 services on students whose academic skills would develop nicely without that special intervention.

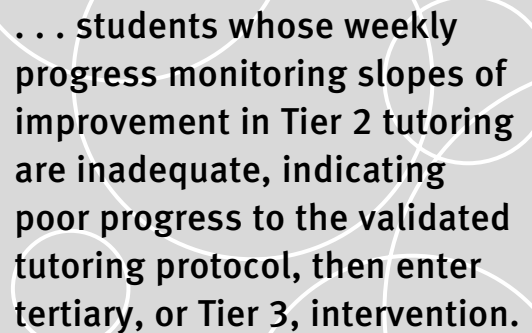
By contrast, students whose weekly progress monitoring slopes of improvement in Tier 2 tutoring are inadequate, indicating poor progress to



the validated tutoring protocol, then enter tertiary, or Tier 3, intervention. This is typically conducted with special education resources and personnel.

At the tertiary or Tier 3 level, instruction differs from Tier 2 because it is (a) more intensive (usually involving longer sessions and conducted in smaller groups, if not one to one) and (b) individualized (rather than relying on a validated treatment protocol). At Tier 3, progress monitoring and diagnostic assessment are therefore essential for two purposes: (a) to inductively formulate the individualized instructional program that is optimal for an individual student (i.e., with progress monitoring and diagnostic assessment) and (b) to determine (this time with progress-monitoring assessment, not diagnostic assessment) when the student's response to Tier 3 instruction

is adequate to warrant a return to Tier 1 primary prevention (general education, with or without accommodations or modifications) or to Tier 2 small-group tutoring (using a standard treatment protocol). With a return to Tier 1 or Tier 2, progress monitoring continues so that secondary or tertiary intervention can be reinitiated as needed.



. . . students whose weekly progress monitoring slopes of improvement in Tier 2 tutoring are inadequate, indicating poor progress to the validated tutoring protocol, then enter tertiary, or Tier 3, intervention.

Appendix A

Passage Reading Fluency

The world's oceans are filled with shipwrecks. Most of them lay hidden and forgotten. When one is found, it is a big deal. Historians and explorers can study these shipwrecks to find out about past cultures.

1–9
10–22
23–31
32–36

One such shipwreck was found in 2003 in the Black Sea. The ship sank more than two thousand years ago. This was the time when the ancient Greeks lived. Explorers hoped to find out more about the Greek civilization from the objects they found in the wreck.

37–47
48–58
59–68
69–78
79–83

One thing they found in the shipwreck was a pile of huge clay jars. The Greeks used these jars to pack and transport all kinds of things. The historians wondered what kind of things had been carried in these jars.

84–95
96–106
107–116
117–123

The answer to that question became the real “treasure” of this shipwreck. When the jars were first studied, fish bones were found inside. Fish bones are usually not very interesting. But this information was very helpful to the explorers and historians. It helped them to understand the types of foods ancient Greeks ate. The fish were probably going to feed hungry Greek soldiers.

124–132
133–142
143–152
153–161
162–170
171–180
181–186

The oceans are still filled with shipwrecks waiting to be found. Some of them hold amazing treasure. Others will give information about the lives of people who lived long ago.

187–196
197–206
207–216

READING RATE AND ACCURACY

Total Words Read: _____
Number of Errors: _____
Number of Correct Words _____
Read Per Minute (WPM): _____
Accuracy Rate: _____
(Number of Correct Words Read per
Minute ÷ Total Words Read)

READING FLUENCY

| | Low | Average | High |
|------------------|-----------------------|-----------------------|-----------------------|
| Decoding Ability | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Pace | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Syntax | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Self-correction | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Intonation | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Maze Fluency

The world's oceans are filled with shipwrecks. Most of them lay hidden and [**yesterday / permission / forgotten**]. When one is found, it is a [**zip / ace / big**] deal. Historians and explorers can study [**wrong / these / blank**] shipwrecks to find out about past [**whether / cultures / accident**].

One such shipwreck was found in 2003 [**in / of / am**] the Black Sea. The ship sank [**more / can't / dust**] than two thousand years ago. This [**ash / sea / was**] the time when the ancient Greeks [**house / lived / blame**]. Explorers hoped to find out more [**trust / about / quote**] the Greek civilization from the objects [**part / sure / they**] found in the wreck.

One thing [**they / mind / sews**] found in the shipwreck was a [**sing / such / pile**] of huge clay jars. The Greeks [**bowl / used / pass**] these jars to pack and transport [**cot / all / gel**] kinds of things. The historians wondered [**what / band / case**] kind of things had been carried [**ox / in / of**] these jars.

The answer to that [**question / convince / material**] became the real "treasure" of this [**fortunate / conclude / shipwreck**]. When the jars were first studied, [**break / fish / quote**] bones were found inside. Fish bones are [**problem / teacher / usually**] not very interesting. But this information [**era / was / fun**] very helpful to the explorers and [**discrimination / historians / considerations**]. It helped them to understand the [**types / upper / shout**] of foods ancient Greeks ate. The [**sale / thin / fish**] were probably going to feed hungry Greek [**soldiers / backfire / valuable**].

The oceans are still filled with [**tournament / reputation / shipwrecks**] waiting to be found. Some of [**sill / talk / them**] hold amazing treasure. Others will give [**information / surrender / responsible**] about the lives of people who [**boost / lived / storm**] long ago.

Appendix B

Writing Prompts

Expository Writing Prompt

Directions for Writing

Think about some outdoor places you have visited. They can be parks, a farm, the beach, or any other outdoor place that comes to mind. Write about the outdoor place you like best. Tell where the place is, what it looks like, and what you did there. Explain why you like this place so much.

Checklist

You will earn the best score if you

- think about your favorite outdoor place.
- think about your audience as you plan your writing.
- write so that your ideas will help the reader understand why you like this place.
- have an opening paragraph that gets the attention of readers.
- write paragraphs that have a topic sentence and focus on related ideas.
- use transition words to go from one idea to another.
- avoid words and phrases that are overused.
- delete ideas that are not important.
- write more sentences and longer sentences when you revise.
- read your writing after you finish and check for mistakes.

Appendix C

Benchmark Assessment Cutoff Numbers

Fluency

Letter-Sound Fluency (Grade K)

| | Test 1 | Test 2 | Test 3 | Test 4 | Test 5 | Test 6 |
|---------|--------|--------|--------|--------|--------|--------|
| Grade K | 0 | 6 | 12 | 18 | 24 | 30 |

Word Identification Fluency (Grade 1)

| | Test 1 | Test 2 | Test 3 | Test 4 | Test 5 | Test 6 |
|---------|--------|--------|--------|--------|--------|--------|
| Grade 1 | 10 | 16 | 22 | 28 | 34 | 40 |

Passage Reading Fluency (Grades 2–6)

| | Test 1 | Test 2 | Test 3 | Test 4 | Test 5 | Test 6 | Test 7 |
|---------|--------|--------|--------|--------|--------|--------|--------|
| Grade 2 | 25 | 40 | 55 | 70 | 85 | 100 | 117 |
| Grade 3 | 44 | 60 | 75 | 91 | 106 | 122 | 137 |
| Grade 4 | 68 | 82 | 96 | 110 | 124 | 138 | 152 |
| Grade 5 | 85 | 98 | 112 | 126 | 140 | 154 | 168 |
| Grade 6 | 98 | 111 | 124 | 137 | 150 | 163 | 177 |

Maze Fluency (Grades 5 and 6)

| | Test 1 | Test 2 | Test 3 | Test 4 | Test 5 | Test 6 | Test 7 |
|---------|--------|--------|--------|--------|--------|--------|--------|
| Grade 5 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| Grade 6 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |

100-Point Skills Battery

| | Test 1 | Test 2 | Test 3 | Test 4 | Test 5 | Test 6 | Test 7 |
|---------|--------|--------|--------|--------|--------|--------|--------|
| Grade K | 7 | 23 | 38 | 54 | 70 | 85 | N/A |
| Grade 1 | 10 | 25 | 40 | 55 | 67 | 85 | N/A |
| Grade 2 | 20 | 30 | 42 | 54 | 66 | 78 | 90 |
| Grade 3 | 20 | 30 | 42 | 54 | 66 | 78 | 90 |
| Grade 4 | 20 | 30 | 42 | 54 | 66 | 78 | 90 |
| Grade 5 | 20 | 30 | 42 | 54 | 66 | 78 | 90 |
| Grade 6 | 20 | 30 | 42 | 54 | 66 | 78 | 90 |

Writing (Grades 2–6)

| | Test 1 | Test 2 | Test 3 | Test 4 | Test 5 | Test 6 | Test 7 |
|----|--------|--------|--------|--------|--------|--------|--------|
| 11 | N/A | N/A | 13 | N/A | N/A | 15 | |

Appendix D

Benchmark Tracking Chart

