Webinar Handout

Types of Reading Assessment Data Used in Schools

	Data Source (including other common names & examples)	Description	Purpose	Designed to Answer Questions	Guidance for Selection & Use	Important Considerations
FORMATIVE	**UNIVERSAL SCREENING** Also referred to as benchmark assessments, interim assessments, dyslexia screener, etc. Examples (not endorsements) include NWEA MAP, Acadience Reading, iReady, STAR, etc.	• Typically administered three times per year (beginning, middle, end) to all students. • Captures broad academic performance and progress across students. • Brief, standardized, and efficient. • Criterion-referenced or norm-referenced.	• Evaluate the effectiveness of the school's core curriculum, environment, and instruction. • Identify students not making expected progress who may need additional intervention.	• Which students are meeting, at risk of not meeting, or not meeting grade-level reading benchmarks? • Is the core (Tier 1) reading instruction effective for the majority of students? • Who may require additional Tier 2 or Tier 3 intervention? • How do screening outcomes compare to prior years or other schools in the district?	• Because screening results can influence access to support, high-quality screeners demonstrate strong reliability, validity, and classification accuracy. • **Tools Chart Overview** (NCII) • **Core Considerations for Selecting a Literacy Screener** (NCIL) • **Screening for Early Literacy Milestones and Reading Disabilities The Why, When, Whom, How, and Where** (Petscher & Gaab)	• Not intended to be diagnostic tool. • Not intended to determine special education eligibility. • Often provide large grain data (i.e., what component of reading is weak, such as word recognition and/or language comprehension), but don't answer why a student is struggling or which specific skills require attention.
	CURRICULUM-EMBEDDED ASSESSMENTS Also referred to as unit tests, exit tickets, assignments, or independent work, etc. Examples (not endorsements) include weekly spelling checks over taught phoneme-grapheme correspondences, multiple-choice comprehension questions over a grade-level text.	• Cover course-specific content directly tied to curriculum. • Administered after exposure to targeted material. • These are brief, teacher-administered, and scored immediately, providing actionable classroom data.	• Evaluate students' understanding of explicitly taught content. • Inform grading and classroom instruction. • Guide reteaching and inform short-term instructional decisions.	• To what extent did students master the specific concept, skill, or text that was taught in this lesson/unit? • Does student performance suggest that the instructional sequence or materials should be adjusted (i.e., reordering lessons to better build prerequisite knowledge)? • For students who easily meet expectations, what opportunities exist to extend or enrich instruction to maintain engagement and growth?	• Evidence supports the use of these to enhance teaching and learning. Meta-analyses demonstrate that these practices yield moderate to large positive effects on student achievement when teachers use results to adapt instruction, and when that alignment with instructional content is critical to the reliability and validity.	• Not designed to measure long-term growth or broader patterns of achievement. • Results may not generalize beyond the instructional context or allow for valid comparisons across classrooms/schools. • May lack the technical rigor of standardized measures, making them less suitable for program evaluation or high-stakes decision-making.

FORMATIVE **PROGRESS MONITORING** Also referred to as curriculum-based measurement (CBM), general outcome measurement, mastery measurement, progress probes, etc. Examples (not endorsements) include oral reading fluency, nonsense word fluency, and maze probes using CBM systems such as DIBELS 8th, Acadience, etc.	• Short, frequent probes of targeted academic skills. • Administered as part of instruction or intervention. • These measures are sensitive to short-term growth and instructional responsiveness, they are not designed to provide a comprehensive picture of overall achievement or to assess mastery of specific content standards.	• Assess students' performance relative to long-term learning goals. • Evaluate responsiveness to instruction and need for instructional adjustment.	• Is the current intervention or instructional approach producing meaningful growth in student performance over time? • Is the student's rate of improvement sufficient to meet grade-level benchmarks or individualized goals within the expected timeframe? • Do results suggest a need to intensify, modify, or maintain the current intervention? • Are patterns across students or groups indicating that the intervention program or instructional materials should be revised/replaced?	• **US Department of Education** • **Reading Progress Monitoring and Instructional Decision Making** (NCII) • **Page 2: Progress Monitoring**	• Extensive empirical evidence supports the reliability, validity, and instructional utility of progress monitoring measures, particularly CBM. • Teachers who use CBM data to guide instruction see greater student achievement gains than those who do not. • The strength of progress monitoring lies in tracking change over time and informing ongoing instructional adjustments - not in replacing more comprehensive assessments.
DIAGNOSTIC ASSESSMENTS Also referred to as diagnostic testing, error analysis, phonics inventory, etc. Examples (not endorsements) include CORE phonics survey, Acadience CFOL, Quick Spelling Screener, etc.	• Administered when progress monitoring data show limited growth or when a student's learning profile is unclear. • Provide detailed information about specific skill strengths and weaknesses. • Emphasize depth rather than breadth; they are time-intensive.	• Determine underlying reasons for academic difficulties. • Inform individualized intervention planning. • May include word recognition and language comprehension skills; as well as related behavioral, motivational, or strategic factors that influence reading.	• Which specific component reading skills or underlying processes are contributing to the student's reading difficulties? • What hypothesis can be generated about a student's learning needs? • How should instruction be tailored or intensified to meet those needs?	• **Diagnostic and Mastery Assessment in Reading** (NCII) • **Example Diagnostic Tools** • **Next STEPS in Literacy Instruction: Connecting Assessments to Effective Interventions**	• The usefulness of diagnostic assessment depends on skilled interpretation and clear instruciotnal hypothesis; when used in isolation, they risk overassessment or misinterpretation. • Not intended to be collected for all learners. Used as needed rather than on a fixed schedule; often following screening or progress monitoring when a student fails to make expected gains. • Not appropriate for frequent use or evaluating program effectiveness.

SUMMATIVE	**SUMMATIVE STATE ASSESSMENTS** Also referred to as outcome measures, high-stakes tests, accountability assessments, etc. These test names vary by state.	• Annual, comprehensive assessment of curriculum across an entire year. • Capture achievement levels of student groups in broad content areas.	• Determine if students meet grade-level benchmarks. • Evaluate effectiveness of schoolwide curriculum and instruction. • Inform accountability and equity monitoring and district or state level.	• How are students performing relative to state standards? • Where do achievement gaps exist across schools or student subgroups? • Where should we target resources to address persistent weaknesses?	• Well-designed state assessments yield reliable and valid measures of achievement at the system level. • **A Policymaker's Guide to State Summative Assessment Systems** • **Why State Reading Tests Are Poor Benchmarks of Student Success**	• Provide consistent metrics for evaluating performance across years and population. • Should not be used for diagnostic or placement decisions at the individual student level.

Content adapted from "**Measuring What Matters: Understanding What Reading Assessments Really Tell Us**". (Toste, et al., 2026).

Discussion Questions:

- Which examples from the article reflect what happens in our school or district?
- Where do we see assessment data being used outside their intended purpose?
- Where do we have gaps or overlaps in our assessments?
- What questions about students' literacy and language skills remain unanswered?
- Are there any tests we should stop using because they aren't validated for the intended purpose?