

Progress Monitoring in an Inclusive Standards-based Assessment and Accountability System

NCEO Synthesis Report 53

Published by the National Center on Educational Outcomes

Prepared by:
Rachel Quenemoen • Martha Thurlow • Ross Moen • Sandra Thompson •
Amanda Blount Morse

February 2004

Any or all portions of this document may be reproduced and distributed without prior permission, provided the source is cited as:

Quenemoen, R., Thurlow, M., Moen, R., Thompson, S. & Morse, A. B. (2003). *Progress monitoring in an inclusive standards-based assessment and accountability system* (Synthesis Report 53). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved [today's date], from the World Wide Web: <http://education.umn.edu/NCEO/OnlinePubs/Synthesis53.html>

Executive Summary

This report describes how progress monitoring—a set of techniques for assessing student performance on a regular and frequent basis—can be an essential and integral part of an inclusive standards-based assessment and accountability system. In order to meet the higher expectations of current standards-based systems, educators need information that can be used to project how students are doing against the grade-level standards throughout the course of the year so they can determine what needs to be done to accelerate student progress toward the proficiency standards. Progress monitoring techniques can provide that information.

While progress monitoring holds much promise for improved outcomes and higher expectations, there are contextual challenges that must be addressed. The challenges that are tied to the progress of students with disabilities that affect the implementation of effective progress monitoring include historical limited access to challenging curriculum, instruction, and assessment; concerns about the target of measurement, that is, whether only basic skills or a full range of rich and challenging content should be measured; and

limited use of data for effective provision of instructional strategies, interventions, and supports.

We discuss the benefits and uses of progress monitoring methods and formative data sources in four general categories: (1) Curriculum-Based Measurement; (2) Classroom assessments (system or teacher-developed); (3) Adaptive assessments; and (4) Large-scale assessments used during the year to monitor growth of individual students and groups of students. We conclude the paper with several recommendations for practice:

- Use multiple measures for progress monitoring.
- Commit necessary resources to build skills and knowledge of all staff on how progress monitoring is used for improvement.
- Find and use available resources.
- Specifically articulate and address as a community the contextual issues of standards-based systems.
- Apply universal design for learning principles to the design of progress monitoring techniques to ensure that individual learner differences are considered from the start.
- Be prepared to have an open discussion of whether the benefit of a comprehensive progress monitoring improvement process is sufficiently large to offset the additional time or cost required for implementation, and enlist the partnership of practitioners who have had success.

Acknowledgements

Our colleagues from three national centers provided comments on earlier drafts of this paper. We thank them for their insights, and encourage readers to make use of their resources. The first is an established center on accessing the general curriculum; the last two are very new progress monitoring centers, funded in 2003.

*National Center on Accessing the General Curriculum (NCAC) CAST, Inc.**

40 Harvard Mills Square, Suite 3

Wakefield MA 01880-3233

phone: 781.245.2212 x233 | fax 781.245.5212

TTY: 781.245.9320

e-mail: ncac@cast.org

<http://www.cast.org/ncac>

National Center on Student Progress Monitoring: Improving Proven Practices in the Elementary Grades

American Institutes for Research
1000 Thomas Jefferson Street, NW
Washington, DC 20007
phone: 202.944.5300 | fax: 202.944.5454
TTY: 877.334.3499
e-mail: studentprogress@air.org
<http://www.studentprogress.org/>

Research Institute on Progress Monitoring

Institute on Community Integration and Department of Educational Psychology
College of Education and Human Development
University of Minnesota
111 A Pattee Hall
150 Pillsbury Dr SE
Minneapolis MN 55455
Phone: 612.626.7220 | fax: 612.625.6619
e-mail: walla001@umn.edu

Overview

This report describes how progress monitoring—a set of techniques for assessing student performance on a regular and frequent basis—can be an essential and integral part of an inclusive standards-based assessment and accountability system. The nation's current emphasis on a standards-based educational system defines success in terms of all children achieving grade-level proficiency. This is a very different context from that in which progress monitoring has been used in the past, where progress was measured against individual starting points, and not to an external criterion. For students with disabilities, the child was expected to make "reasonable" progress defined by the Individualized Educational Program (IEP), which reflected how much progress the school, parent, and student were willing to accept. The emphasis on external criteria of grade-level content and achievement standards in an inclusive standards-based assessment and accountability system has raised the bar of what "reasonable" student progress is. Schools are held accountable for these higher expectations regardless of past practices.

In order to meet the higher expectations of current standards-based systems, educators need information that can be used to project how students are doing against the grade-level standards throughout the course of the year so they can determine what needs to be done to accelerate student progress toward the proficiency standards. Progress monitoring techniques can provide that information. Although the promise of progress monitoring is great, the techniques are not universally used, nor are they universally understood in the context of standards-based reform. If progress monitoring is going to be relevant to the

current context of standards-based assessment and accountability, external criteria must be a part of the progress monitoring system.

We first describe the current context of standards-based assessment and accountability systems in which students with disabilities are to be included. Next we identify and discuss the contextual challenges that must be addressed for the promise of progress monitoring to be reached within an inclusive standards-based assessment and accountability system. We describe a comprehensive progress monitoring improvement process with multiple assessment methods that fit into a standards-based system. Finally, we offer recommendations to consider as progress monitoring approaches are added to inclusive standards-based assessment and accountability systems.

Context of Inclusive Standards-based Assessment and Accountability Systems

With the reauthorization of the Individuals with Disabilities Education Act (IDEA) in 1997, significant new requirements were put into place to ensure that all students had access to and made progress in the general curriculum to the maximum extent possible. One way to support this access to and progress in the curriculum was to require that students with disabilities participate in state and district assessments, with appropriate accommodations if necessary, or in alternate assessments developed for those students unable to participate in general state and district assessments. Assessment of student achievement and the accompanying requirement of state, district, and school accountability for all students' success in the grade-level content is the centerpiece of the No Child Left Behind Act of 2001 (NCLB).

A decade ago, most states included fewer than 10% of students with disabilities in achievement testing or they simply did not know how many participated (Shriner & Thurlow, 1993). That number often reflected states' expectations about who could achieve and beliefs that achievement requirements could harm some students. At the turn of the century, the average percentage of students in the general assessment was 84% (Thurlow, Wiley, & Bielinski, 2002), and since then, NCLB has required a 95% participation rate in the assessment system. The education system has had to come a long way quickly, based on a belief that the greatest harm to students is caused by what President Bush has called the "soft bigotry of low expectations."

Research suggested that even when students with disabilities participated in assessment systems, the assessment results were not necessarily used in the same way for students with disabilities as the results of other students were used (Krentz, Thurlow, & Callender, 2000). Sometimes they were not reported and sometimes they were not included in accountability systems. Adding to this is considerable evidence that statewide test results for students with disabilities are poor (Thurlow et al., 2002), generally falling significantly below the performance of students without disabilities.

Thus the challenge for states and districts is discovering how to move beyond simply ensuring that students with disabilities participate in assessments and that their results are reported. They need to find ways to improve those results. While many good recommendations can be made that focus on the assessments themselves, such as developing universally designed assessments (Thompson, Johnstone, & Thurlow, 2002) and providing students with appropriate accommodations (Thurlow, Quenemoen, Thompson, & Lehr, 2001), more than this is needed to improve educational results.

Fuchs and Fuchs (1986) are among many (Fuchs & Deno, 1991; Helwig, Heath, & Tindal, 2000; Langenfeld, Thurlow, & Scott, 1997; Lindsey, 1990; Marston, 1989; Paulson, Paulson, & Meyer, 1991; Stecker & Fuchs, 2000; Stiggins, 2001; Wiggins & McTigue, 1998) who have emphasized that it is essential to have assessment that involves the ongoing collection and use of information to evaluate the effectiveness of instruction. The President's Commission on Excellence in Special Education (U.S. Department of Education, 2002) also emphasized the need for implementation of continuous progress monitoring. There seems to be growing recognition among educators, researchers, and policymakers of the need for more widespread use of progress monitoring, that is, a set of techniques for assessing student performance on a regular and frequent basis.

Challenges to Using Progress Monitoring in the Context of Standards-based Assessment and Accountability Systems

While progress monitoring holds much promise for improved outcomes and higher expectations, there are contextual challenges that must be addressed. The challenges that are tied to the progress of students with disabilities and that affect the implementation of effective progress monitoring include:

1. historical limited access to challenging curriculum, instruction, and assessment;
2. concerns about the target of measurement, that is, whether only basic skills or a full range of rich and challenging content should be measured; and
3. limited use of data for effective provision of instructional strategies, interventions, and supports.

Historical Limited Access to Challenging Curriculum, Instruction, and Assessment

A key emphasis of standards-based reform is that all children must have access to varied and challenging curriculum and instruction that is aligned to and focused on the grade-level content and achievement standards. But standards-based systems are relatively recent and students with disabilities historically have not had full access to challenging curriculum, instruction, and assessment. Discussion leading up to the reauthorization of IDEA 1997 suggested that over 20 years of low expectations resulted in dismally low achievement for students with disabilities. The IEP process often contributed to the

lowering of expectations for individual students – essentially defining a separate curriculum for the student - instead of resulting in a plan for the services and supports so that the student could be successful in the general curriculum, the challenging content same age peers were learning. Now as we look toward another reauthorization of the law, we realize that we have not as yet reached the goal of ensuring that all children with disabilities have access to, participate in, and make progress in the general curriculum that is built on grade-level challenging standards. This continuing legacy of low expectations is a foundational contextual challenge within which progress monitoring is implemented.

There is considerable rhetoric in some quarters that asserts all children can learn to grade-level content and achievement standards. There is considerable rhetoric in other quarters that asserts that NOT all children can learn to grade-level content and achievement standards and that it is unfair to schools, educators, and even the children to expect them to achieve to this level. There is general agreement that children in the earliest grades of school should be held to high expectations so that they do not fall behind from the beginning. This philosophy underlies the federal Reading First program.

There is a developing consensus that alternate achievement standards (defined through valid and documented methods to reflect the outcomes of the best possible instruction) are appropriate for a very small percentage of students who have significant cognitive disabilities. These are not lowered expectations; in states that have set these alternate achievement standards with research-based understanding of teaching and learning for students with the most significant cognitive disabilities, these alternate achievement standards set the bar and expectations for student achievement on the challenging content at very high and appropriate levels (Arnold, 2003; Olson, Mead, & Payne, 2002; Wiener, 2002).

There is not consensus on what high expectations mean for students with moderate to mild disabilities, sometimes called gap or gray area students. NCLB requires that states set expectations for the grade-level achievement of all students, and that schools are to be held accountable for the achievement of all students, as measured through the state assessment system. Using emerging data from these standards-based assessment systems in all states, we need to illuminate whether the gap or gray issues for these students are those of curriculum, instruction, assessment, or actual student need (Almond, Quenemoen, Olsen, & Thurlow, 2000; Quenemoen & Almond, 2001). We do not have adequate data to determine this empirically at this time, but as more students with disabilities are provided high quality instruction in the challenging curricula taught to their peers, and as all children are included in standards-based assessment systems, these data will emerge.

Progress monitoring assessment techniques should be part of these assessment systems. Progress monitoring can exist outside of standards-based curriculum, instruction, and assessment, but it will be limited in its effectiveness. As the National Research Council (1999) suggests, alignment between tests and standards is a necessary condition of the theory of action of standards-based reform, and that includes alignment of assessments

used for the purpose of monitoring progress. Any state or district that hopes to use progress monitoring as a tool to ensure the highest possible outcomes for every student needs to ensure that the educational system itself—including curriculum, instruction, formative and summative assessments, professional development and school improvement processes—is aligned, coherent, and focused on ensuring that every child is being taught and is learning the grade-level content.

Concerns About the Target of Measurement

The target of measurement is another issue that generates controversy when measuring progress over time. Despite strong correlations between performance on basic skills and more complex skills, many educators express concerns about the need to balance the focus on developing basic skills in reading and mathematics with ensuring progress and achievement in the rich and challenging content beyond basic skills. Researchers (e.g., Lane & Silver, 1995; Oakes, 1990) have questioned whether the drill and practice of basic factual knowledge and computational skills in mathematics comes at the expense of work on complex tasks requiring mathematical reasoning and problem-solving, particularly for some groups of students. Similarly, Vinovskis (1996) argues that while providing books with limited vocabulary may help decoding, it stunts vocabulary development and prevents access to other challenging academic standards. Walmsley and Allington (1995) found that students with print disabilities spent a great deal of their language arts instruction in remedial programs designed to circumvent their decoding problems. Remediation programs are not, per se, a harmful activity for students with print disabilities; rather, the harm comes to students from opportunity lost for other important content.

As Walmsley and Allington (1995) described, students are commonly taken out of content area courses or even grade-level reading courses to create time for direct instruction of basic reading skills. One argument for this practice is that until the students can read independently, they cannot benefit from the existing curriculum in the content areas or from grade-level reading instruction. CAST, an educational research and development group, and the National Center on Accessing the Curriculum (NCAC) at CAST have worked to develop the concept of Universal Design for Learning (UDL) as an alternative to this common practice. Instead, individual learner differences are considered from the start in the instructional process, and instruction, materials, and tools are made accessible for all students. Hitchcock, Meyer, Rose, and Jackson (2002), in describing UDL approaches, provide an example that illustrates how important content can be accessed by any student. In addition, they stress that what tools are used to provide access depends on what is being taught, that is, on the goal of instruction and measurement in a standards-based system. Their example is:

Suppose a student is assigned an Aesop's fable to read. The purpose of this assignment determines the appropriate steps for making it accessible. Is the goal to learn to decode text, to learn comprehension strategies for extended passages, to build vocabulary, to learn the moral or point of the fable, to learn the common elements of any fable, to learn

how to compare and contrast fables with news reports, to articulate the relationship between the fable and the overall culture? The scaffolds and supports that might be appropriate depend entirely on the purpose of the assignment.

If for example the purpose of the fable assignment were to become familiar with the elements commonly found in fables, then supporting word decoding, vocabulary, and comprehension of the story itself would not interfere with the learning challenge. Supports such as text-to-speech, linked vocabulary, or animations illustrating interactions between characters would support different students but still leave the appropriate kind of challenge for all learners. But if the goal were to provide practice in decoding and reading fluency, providing those same supports could undermine the learning challenge and actually impede access to learning. The reading support would eliminate the students' opportunity to practice and work towards reading independence. (Hitchcock et al., 2002, p.15)

If some students are prevented from accessing the challenging content beyond decoding and fluency, their standards-based progress will suffer even as their basic skills develop. This does not discount the need for, or the value of teaching basic skills. Students need access to ALL essential skills and knowledge, and practices in teaching and in assessment will have to reflect that requirement. The CAST/NCAC example demonstrates the complex issues surrounding the target of measurement debate and how these issues relate to the need for effective instructional strategies, interventions, and support.

Limited Use of Data for Effective Provision of Instructional Strategies, Interventions, and Supports

Gersten, Keating, and Irvin (1995) argue that assessment is only valid if it results in improved learning; in many systems, assessment data from district, school, classroom, and individual levels are not used effectively to improve outcomes. Systems level school improvement processes are essential to ensuring that all children have access to rich and challenging content on a daily basis. School improvement processes must look at classroom and schoolwide formative and summative data to determine research-based practices, including UDL, that can improve outcomes for the entire student population. Progress monitoring methods can give us data to determine whether some groups of children or individual children need additional interventions and supports. Challenges remain in using the data at all levels. School improvement teams, intervention specialists, those involved in pre-referral processes, and IEP teams must have access to and an understanding of research-based practices to effectively use all available data to make empirically based decisions about instructional programming.

Many states and districts are working to ensure that school improvement processes required by state and federal law consider the achievement of all students, and resulting school improvement plans provide universal access to challenging and effective curriculum and instruction. States and districts are also working to improve educators' skills in designing interventions and managing pre-referral processes. Ultimately, if

students are identified as requiring special education services and supports, then training and support on decision-making processes and best practices need to be in place for IEP team members. Learning how to balance competing learning needs while maintaining a longitudinal view of student learning is essential (McLaughlin, 2000). Thus, partnerships with experts and organizational resources on pre-referral, intervention, and IEP team processes are necessary. These need to be designed in a way that constantly emphasizes increased data, thought, and attention to longitudinal implications of instructional decisions. This will define the necessary changes in instruction, policy, and accountability for teachers and schools.

Challenge Implications: High Expectations

These three interrelated contextual challenges lead to an essential question that needs to be articulated and answered. That is, how do we build research-based understanding and stakeholder consensus on high student expectations that are essential to the success of all children? The expectation in *No Child Left Behind* is that all children will achieve at state defined proficiency levels. Yet, "by the end of kindergarten, the distribution of mathematical development in typical public school classrooms spans more than four grade levels, with many children, especially those from high-poverty backgrounds, demonstrating large deficits in their mathematical thinking and skills" (Fuchs, Fuchs, Yazdian, & Powell, 2002, p. 569). This is consistent with literature on early reading, which suggests that children from poor or minority backgrounds have a higher risk of poorer reading outcomes than white, middle class children (National Research Council, 1998). Simply put, children come to school as kindergarteners with extremely varied preparation for learning, regardless of any risk for disability determination.

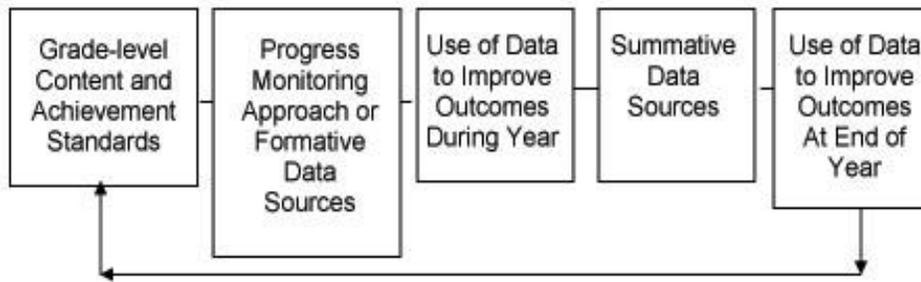
The challenge of implementing a comprehensive progress monitoring improvement process to ensure that all children learn to high levels regardless of how prepared they are when they come to school is considerable but must be met. Addressing this basic question of expectations is central to the challenge, and is the foundation on which all efforts must rest. In the next section, we describe what such a process would look like, and identify multiple progress monitoring techniques that can be used.

A Comprehensive Progress Monitoring Improvement Process

Figure 1 shows a comprehensive progress monitoring improvement process to ensure that each child achieves to state defined proficiency levels. The essence of progress monitoring is that assessment data inform educators when students are not progressing as they should so that the educators can take action to improve student progress, whether they are performing above or below the grade-level targets. These actions might include changing instructional practices, adjusting curricula, and adding necessary services, supports, and interventions that will ensure that each child makes the progress expected for all children. This requires multiple methods that carefully measure progress on the

development of basic skills in reading, math, and writing, as well as the challenging content in rich and varied content areas. These progress measures can provide data for whole school and district improvement, classroom improvement, and improvements for subgroups of children including those with disabilities, English language learners, and individual children.

Figure 1. Comprehensive Progress Monitoring Improvement Processes



In the past, some approaches to progress monitoring looked to expectations that were set individually through the IEP process, or to local or national normative data to gauge what the expectations for students should be, and thus what benchmarks were set throughout the year (Carnine & Granzin, 2001; Deno, Fuchs, Marston, & Shin, 2001). Placing these approaches within the context of standards set for the grade level, as required in current standards-based systems, is essential. Deno et al. (2001) addressed the implications of a normative approach to setting expectations, which assumes that the current typically observed growth rates for students with disabilities are reasonable and what should be expected. Following this line of reasoning, students with learning disabilities will learn at a slower rate than other students. The researchers speculate that this kind of reasoning reflects the "well-accepted fact that special education, as typically practiced in this country, fails to regularly incorporate demonstrably effective methods" (p. 515).

Under current federal requirements, for each classroom and for each child other than those included in the small exception for alternate achievement standards, the current achievement levels must be compared with yearend standards-based grade-level expectations. The learning slope represents the targeted rate of progress, and the student's actual rate of progress should be monitored daily, weekly, monthly, or quarterly to determine whether yearend mastery is on target. If not, the teacher intervenes on instruction in hopes of accelerating learning. Then progress monitoring continues, and the efficacy of the intervention is determined, adjusted, and so on. Aggressive interventions can be implemented and monitored, with special attention paid to students who continue to be resistant to intervention treatment (Vaughn, Linan-Thompson, & Hickman, 2003).

We discuss the benefits and uses of progress monitoring methods and formative data sources in four general categories: (1) Curriculum-Based Measurement; (2) Classroom assessments (system or teacher-developed); (3) Adaptive assessments; and (4) Large-

scale assessments used during the year to monitor growth of individual students and groups of students.

Benefits and Uses of Multiple Progress Monitoring Methods

A comprehensive progress monitoring improvement process includes multiple assessment methods that have different strengths and meet different needs. As these methods are implemented, we find that they tend to be blended in practice. For example, curriculum-based measures are used to target focused skill development which in turn helps inform the design of teacher-developed classroom assessments across content areas; classroom assessment strategies are informed by large-scale assessment data both at grade level and at instructional level (if it is different); reports of student progress to parents tend to reflect data from all measures in use. This is the ideal use of multiple measures for ensuring that all students succeed.

Curriculum-Based Measurement (CBM)

Experimental research on the use of progress monitoring to enhance student performance has focused primarily on one form of progress monitoring: Curriculum-Based Measurement (CBM). Academic areas frequently assessed using CBM include reading, mathematics, written expression, and spelling. Criteria that describe the measures used as part of CBM include valid and reliable indicators of generalized performance, short duration to facilitate frequent administration, a focus on direct and repeated measures of student performance, multiple forms that are inexpensive to create and produce, and sensitivity to changes in student achievement over time (Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993). Especially important for ongoing classroom use is the demonstrated technical adequacy of the approach for accurate, easy, and repeatable measures (Crawford, Stieber, & Tindal, 2000; Crawford, Tindal, & Stieber, 2001; Helwig, Anderson, & Tindal, 2002; Helwig, et al., 2000; Ketterlin-Geller, McCoy, Twyman, & Tindal, 2003). Shinn and Bamonto (1998) estimated that over 150 articles have been published since 1988 on Curriculum-Based Measurement. CBMs have been used primarily in basic skill areas and in elementary school grades, but development work is taking place to expand their use to other content and to other age and grade levels (Espin, Busch, Shin, & Kruschwitz, 2001; Espin & Foegen, 1996; Foegen, 2000; Foegen & Deno, 2001; Gansle, Noell, & VanDerHeyden, 2002).

There are a number of strengths of CBM that other approaches view as challenges. For example, CBM has the advantage of simple hand scoring. Other classroom assessments often involve more time consuming, sometimes subjective scoring procedures. CBM has viable ways to conduct error analyses to identify specific targets for intervention efforts. Other approaches have not developed similar focused intervention tools. Graphing of student progress (baseline to target and trend line) has been used effectively in CBM (Calhoun & Fuchs, 2003; Deno, 1992; Fuchs, Fuchs, Hamlett, Thompson, Roberts,

Kubek, & Stecker, 1994; Pemberton, 2003). Graphing provides a pictorial representation of the student's progress so that teachers, parents, and students themselves can see exactly what student progress looks like. This can eliminate the false perceptions of where the child is academically and could potentially motivate students to continue working to reach the target. Although many forms of classroom assessment provide visual representations of student work compared to exemplars, these other approaches do not lend themselves to graphing. Other approaches to progress monitoring need to meet the challenge of providing similar high quality data tools and strategies.

Those who have been conducting the strong program of research on curriculum-based measures (e.g., see Deno et al., 2001) recognize that there are many issues yet to be addressed. These issues may include the best method for setting long-range goals, the frequency with which assessment should occur, the need to revise measures according to student performance level, and ways to help teachers use the data for instructional decision making. The clear commitment by researchers to continued rigorous debate, recognized by the editors of the 2002 special issue *School Psychology Review: Special Topic Development in Academic Assessment and Intervention*, is an outstanding resource to efforts to ensure that all children succeed (Daly & McCurdy, 2002).

Classroom Assessments (System or Teacher Developed)

Stiggins (2001) noted that the current "dismal state of classroom assessment" has "kept classroom assessment from even approximating its immense potential as a school improvement tool" (p. 5). Yet, progress monitoring in the context of federal legislation may be the catalyst needed to bring about a transformation in classroom assessment practices. Stiggins reinforces this, concluding that "we have an excellent foundation from which to develop strong classroom assessments" (p. 12). The same message is proclaimed by many other leaders in the field: that classroom assessment is important, that it typically is not done well now, but that teachers can be taught to do it well. Marzano (2003) bases his work on 35 years of research that shows progress monitoring (labeled goals and feedback) as one of five school-level factors and one of nine teacher-level factors that can improve student learning. Black (2003) and Shepard (2000) elaborate on effective feedback from classroom assessments. Wiggins and McTigue (1998) show teachers how to use standards-based classroom assessment as the starting point and then how to design backward from the skills and knowledge to be assessed to plan for instruction.

The importance of classroom assessment at this time is underscored by the recent formation of the National Research Council's Committee on Assessment in Support of Instruction and Learning (see *Bridging the Gap Between Large-Scale and Classroom Assessment*, Petit, 2003). Projecting success on large-scale assessments by monitoring progress on classroom assessments will lead educators to improve the quality of classroom assessments and the alignment between classroom assessments and large-scale tests. Classroom assessment can take many forms. That is a feature that fits particularly well with an emphasis of the National Research Council's Committee on the Foundations

of Assessment (Pellegrino, Chudowsky, & Glaser, 2001, p. 293) on the "need for comprehensive systems of assessment consisting of multiple measures, including those that rely on the professional judgments of teachers." Several states continue to work toward a combination of summative state large-scale assessments and formative local classroom assessments as the best method of accountability. The kinds of observations to be made and the process for interpreting these observations vary according to the nature of what is to be measured.

A less teacher-dependent form of classroom assessment relies on assessments provided by publishers for use with books and other instructional materials. Publisher-supplied tests are valuable for progress monitoring only to the extent that the published materials and accompanying assessments are aligned with the target standards. Yet the potential is great. Pellegrino et al. (2001) cite developments in statistical modeling methods (e.g., Bayesian inference networks, or Bayes nets) that allow for more complex reasoning about complex student competencies – far beyond basic skills, and tailored to specific content domains. They speculate that by building these complex assessment approaches into intelligent tutoring systems or other instructional materials, teachers will have increased ability to understand student progress at multiple levels.

Adaptive Assessments

Adaptive assessments are tests that determine the items to which a student responds based on the student's performance levels. Although this approach can be implemented without computers, it works best with computers. On computers it is called Computerized Adaptive Testing (CAT) (Thompson, Thurlow, Quenemoen, & Lehr, 2002). CAT improves testing efficiency and precision because instant item scoring lets the computer exclude tasks that are too easy or too hard for a student and focuses only on reasonably challenging tasks. Depending on their design and implementation, CATs can be administered several times each year. CAT systems typically have an Item Response Theory (IRT) foundation that provides a built-in scale suitable for reporting progress. One of the challenges for CAT is that unless the range of the assessment has been constrained, results may say very little about a student's progress on grade-level standards (Olson, 2002). This is a serious constraint in the context of NCLB accountability requirements, and also could result in inadvertent lowering of expectations. If a core assessment is given that is controlled to remain at grade level only, and then additional items added to provide more precise diagnostic information, this serious limitation can be addressed.

Large-scale Assessments

The final category of progress monitoring approaches is different from the others described here. It involves the use of large-scale tests that are usually administered one to three times per year to show growth over time, and thus to monitor student progress. This is one goal of many states' current efforts to vertically scale their annual tests and to

maintain longitudinal student data. This effort involves implementing vertical scaling, which involves creating a single scale on which the test for each grade can be placed. There is considerable debate on whether the goal of vertical scaling is achievable or desirable. For example, there are fundamental questions about whether grade-level content lends itself to vertical explication, or if there are content "chunks" that are not coherently linked from grade to grade (e.g., panel presentations by Orr, Chin-Chance, Rabinowitz, & Vukminovic, 2003). Beyond this debate among measurement and curriculum experts, the extent to which this approach would meet the desired characteristics of frequent checks on student performance is questionable given the infrequent and summative nature of most of these assessments, and the time required for administration. Yet many states and districts continue to explore how a large-scale test can be used for multiple purposes, including tracking progress and growth of individual students as well as for diagnostic purposes. Whether that is a reasonable expectation is debatable, but there may be a place for use of large-scale assessment data in progress monitoring as one of multiple measures to be considered.

Broader Progress Monitoring Issues

Successful implementation of progress monitoring is not just a matter of picking an approach or a combination of approaches. Regardless of methods used, progress monitoring approaches in a standards-based assessment and accountability system must include defined strategies for scoring, analyzing, reporting, and tracking data, and defined strategies for creating meaning from the data gathered across all sources to develop effective improvement plans.

Strategies for Scoring, Analyzing, and Tracking Data

Solid strategies for local implementation of scoring and analysis of procedures depend on the nature of the assessment (e.g., hand scoring, computerized scoring). Identifying procedures for quick turn-around of scores is desired to provide feedback for instruction. It is important for districts and schools to keep longitudinal data for each student. At the classroom level, this could be relatively simple using paper-and-pencil or gradebook-like computer programs. But to allow the data to be rolled up or down to and from the school, district, or even state level, and to allow for year-to-year progress monitoring, more sophisticated computer programs and a system for uniquely identifying each student is needed. Using a relational database that would link progress in one area to progress in other areas and to other student or school information would allow educational administrators and researchers to try to understand differential student progress within the bounds of student privacy limitations. Web sites such as the American Association of School Administrators (<http://www.aasa.org/>) list a range of such technology resources.

Deriving Meaning From the Data to Develop Effective Improvement Plans

The primary purpose of data, as shown in Figure 1, is to identify classrooms, subgroups of students within classrooms, and individual students who are not showing expected progress in the content areas. The benefit of progress monitoring relies on effective use of this information by teachers to identify areas of needed support, more instruction, or changes in instruction. It also supports effective communication with students and parents to help particular students or the entire classroom make greater progress. Teachers need help learning how to do this. Leadership, information and staff development are needed for currently employed teachers; changes in teacher education are needed to help teacher candidates (Stiggins, 2001).

The difficulty of understanding the meaning of data from multiple measures of progress monitoring for individual students and for groups of students and then choosing the best research-based interventions for a particular situation cannot be overemphasized. Numerous educators, researchers, and policymakers are grappling with this reality. Principals, teachers, parents, and students need the data on which instructional planning is based to be meaningful, understandable, and useful in making decisions on instructional interventions, services, and supports. Hand charting sets of data for entire classrooms has also been found to be effective for helping teachers and school level administrators get a feel for changes that may be needed at that level (Sargent, 2001). More complex graphics and statistical reports are needed to help administrators and researchers understand group progress from the classroom level on up to the state level (Pellegrino et al., 2001). In the long term, preservice training will play a role to ensure that teachers and school psychologists are equipped with the necessary knowledge and skills to use these tools and strategies when they start their careers.

District level administrators also need help in using data, as suggested by the quote from a superintendent in the forward to *Using Data to Improve Schools: What's Working*: "We spend a lot of time on testing but not much time on what to do with the results" (AASA, 2002, p. iii). Administrators who use progress information effectively examine progress data to determine whether some teachers or schools are making more progress than others in particular curricular areas or with particular groups of students. Special attention is given to any areas where students are not on course to show proficiency on state standards. These data are combined with other information to, for example, make data-guided decisions about changes in organizational structure and staffing, allocating resources for additional help, and staff development. Staff and schools that seem to be particularly strong in some areas or with some students can be used as resources to help others in those areas. Similarly, resources might be committed to revising curricula or in other ways strengthening areas that seem to be a weakness for the entire system.

Recommendations

There are several recommendations to consider as you implement progress monitoring in an inclusive standards-based assessment and accountability system.

1. Use multiple measures for progress monitoring. Each of the approaches discussed has strengths and drawbacks and none is a magic solution to the challenges facing schools. Commitment to just one measure can derail a well intentioned effort and may cause harm. Ensure the consistent use of monitoring devices over time.
2. Commit necessary resources to build skills and knowledge of all staff on how progress monitoring is used for improvement including time for training, practice, coaching, and collaborative problem-solving. Include parents and students in training as appropriate to ensure a full partnership with all the key stakeholders. Building a community of practice around a comprehensive progress monitoring improvement process—within the school and across the community—will ensure that a commitment builds for effective use of the process.
3. Find and use available resources. Technical assistance and research centers can provide resource maps and literature reviews for use by states and districts. Existing state and district initiatives in the content areas can also be essential resources to schools. Examples of these include Reading First programs and state and local chapters of professional organizations such as the Association for Supervision and Curriculum Development, the National Staff Development Council, National Council of Teachers of Mathematics, International Reading Association, and the Eisenhower Regional Consortia for improving K-12 math and science education. (See the acknowledgement page for contact information for two national centers on progress monitoring, and one on accessing the general curriculum.)
4. Specifically articulate and address as a community the contextual issues of standards-based systems including: the historical limited access to challenging curriculum, instruction, and assessment; concerns about the target of measurement; and the need for effective provision of instructional strategies, interventions, and supports. Focus on research based curricular and instructional practices that are most effective in helping most students reach proficiency on grade-level standards. In addition to being used for identifying generally effective curricular and instructional strategies, assessment data also needs to be used to help identify students and staff who need something other than these generally effective strategies (e.g., more time, more help, a different approach). In identifying the particular needs of some students and staff, avoid meeting one academic need in a way that prevents a child from making progress on other academic needs (e.g., decoding vs. access to other content). Once an intervention strategy is determined, continue progress monitoring throughout the school year, to ensure interventions are working. Finally, when identifying students who need something besides the typical strategies, systematically address and circumvent educators' tendency to translate an individualized need into a lowered expectation for performance.

5. Apply universal design for learning principles to the design of progress monitoring techniques to ensure that individual learner differences are considered from the start. As fully accessible curriculum materials become widely used in schools, and technology systems are developed that collect, present, and offer authentic and timely student data, the reality of on-going progress monitoring can be realized.
6. Given the resources required to implement these recommendations, there will be questions about whether the benefit of a comprehensive progress monitoring improvement process is sufficiently large to offset the additional time or cost required for implementation. Be prepared to have this discussion openly and enlist the partnership of practitioners who have had success.

The perspective that can guide this discussion is that progress monitoring encourages continued persistence and high expectations for all students. It moves educators and society as a whole away from accepting the notion that there are always going to be students who will be underachieving (i.e., at the low end of the spectrum), away from a belief that is "just the way it is." Instead, continual progress monitoring that informs instruction and intervention focuses attention on the essential effort to help each student reach grade-level expectations. It encourages continual tweaking of instruction to find what works for a student or group of students. It also encourages a sense of effectiveness in teachers—they are not powerless to help their students and they cannot simply accept the belief that they can do nothing. That sends a powerful message to the students: "We are not going to give up on you (as an individual or as a group of students). We are going to continue to monitor your progress because we want to see you reach the goal, and we are going to do what is necessary (make the necessary changes and add the necessary supports) to facilitate your progress because we believe in you."

This perspective—and this message—can profoundly change the dynamics of teaching and learning in a district and school. Progress monitoring in a standards-based system can be the key to unlock powerful skills and knowledge for teachers and for students and can result in success for the school, district, and state in inclusive standards-based assessment and accountability systems.

References

American Association of School Administrators. (2002). *Using data to improve schools: What's working*. Retrieved from

<http://www.aasa.org/cas/UsingDataToImproveSchools.pdf>

Almond, P., Quenemoen, R., Olsen, K., & Thurlow, M. (2000). *Gray areas of assessment systems* (Synthesis Report 32). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Available at

<http://www.cehd.umn.edu/NCEO/OnlinePubs/archive/Synthesis/Synthesis32.html>

Arnold, N. (2003). *Washington alternate assessment system technical report on standard setting for the 2002 portfolio* (Synthesis Report 52). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Available at <http://education.umn.edu/nceo/OnlinePubs/Synthesis52.html>

Black, P. (2003, April). *The nature and value of formative assessment for learning*. Paper presented with the King's College London Assessment for Learning Group at AERA, Chicago.

Calhoon, M.B., & Fuchs, L.S. (2003). The effects of peer-assisted learning strategies and curriculum-based measurement on the mathematics performance of secondary students with disabilities. *Remedial and Special Education, 24*(4), 235-245.

Carnine, D., & Granzin, A. (2001) Setting learning expectations for students with disabilities. *School Psychology Review, 30*(4), 466-472.

Crawford, L., Stieber, S., & Tindal, G. (2000). *Using timed oral readings to predict students' performance on statewide achievement tests*. Eugene: RCTP.

Crawford, L., Tindal, G., & Stieber, S. (2001). Using oral reading rate to predict student performance on statewide achievement tests. *Educational Assessment, 7*(4), 303-323.

Daly, E., & McCurdy, M. (2002). Getting it right so they can get it right: An overview of the special series. *School Psychology Review, 31*(4), 453-458.

Deno, S. L. (1992). The nature and development of curriculum-based measurement. *Preventing School Failure, 36*(2), 5-10.

Deno, S., Fuchs, L., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with disabilities. *School Psychology Review, 30*(4), 466-472.

Espin, C.A., Busch, T., Shin, J., & Kruschwitz, R. (2001). Curriculum-based measures in the content areas: Validity of vocabulary-matching measures as indicators of performance in social studies. *Learning Disabilities Research and Practice, 16*(3), 142-151.

Espin, C.A., & Foegen, A. (1996). Validity of three general outcome measures for predicting secondary students' performance on content-area tasks. *Exceptional Children, 62*, 497-514.

Foegen, A. (2000). Technical adequacy of general outcome measures for middle school mathematics. *Diagnostique, 25*, 175-203.

Foegen, A., & Deno, S. L. (2001). Identifying growth indicators for low-achieving students in middle school mathematics. *The Journal of Special Education, 35*, 4-16.

- Fuchs, L., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, 53(3), 199-208.
- Fuchs, L.S., Fuchs, D., Hamlett, C.L., Thompson, A., Roberts, P.H., Kubek, P., & Stecker, P.S. (1994). Technical features of a mathematics concepts and applications curriculum-based measurement system. *Diagnostique*, 19(4), 23-49.
- Fuchs, L., Fuchs, D., Hamlett, C.L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review*, 22, 27-48.
- Fuchs, L., Fuchs, D., Yazdian, L., & Powell, S. (2002). Enhancing first-grade children's mathematical development with peer-assisted learning strategies. *School Psychology Review*, 31(4), 569-583.
- Fuchs, L.S., & Deno, S.L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children*, 57, 488-501.
- Gansle, K.A., Noell, G.H., & VanDerHeyden, A.M. (2002). Moving beyond total words written: The reliability, criterion validity, and time cost of alternate measures for curriculum-based measurement in writing. *School Psychology Review*, 31(4), 477-497.
- Gersten, R., Keating T.J., & Irvin, L.K. (1995). The burden of proof: Validity as improvement of instructional practice. *Exceptional Children*, 61, 510-519.
- Helwig, B., Anderson, L., & Tindal, G. (2002). Using a concept-grounded, curriculum-based measure in mathematics to predict statewide test scores for middle school students with learning disabilities. *The Journal of Special Education*, 36(2), 102-112.
- Helwig, R., Heath, B., & Tindal, G. (2000). *Predicting middle school mathematics achievement using practical and efficient measurement instruments*. Eugene: RCTP
- Hitchcock, C., Meyer A., Rose, D., & Jackson, R. (2002). Providing new access to the general curriculum: Universal design for learning. *Teaching Exceptional Children*, 35(2), 8-17.
- Ketterlin-Geller, L.R., McCoy, J.D., Twyman, T., & Tindal, G. (2003). How do critical thinking measures fit within standards-based reform? *Assessment for Effective Intervention*, 28 (3&4), 37-48.
- Krentz, J., Thurlow, M., & Callender, S. (2000). *Accountability systems and counting students with disabilities* (Technical Report 29). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Available at <http://www.cehd.umn.edu/NCEO/OnlinePubs/archive/Technical/Technical29.htm>

- Lane, S., & Silver, E. (1995). Equity and validity considerations in the design and implementation of a mathematics performance assessment: The experience of the QUASAR project. In M. Nettles, & A.L. Nettles (Eds.), *Equity and excellence in educational testing and assessment* (pp. 185-219). Boston: Kluwer.
- Langenfeld, K., Thurlow, M., & Scott, D. (1997). *High stakes testing for students: Unanswered questions and implications for students with disabilities* (Synthesis Report 26). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Available at <http://www.cehd.umn.edu/nceo/OnlinePubs/archive/Synthesis/Synthesis26.htm>
- Lindsey, O.R. (1990). Precision teaching: By teachers for children. *Teaching Exceptional Children*, 22(3), 10-15.
- Marston, D. (1989). Curriculum-based measurement: What is it and why do we do it? In M.R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18-78). New York: Guilford.
- Marzano, R. (2003). *What works in schools: Translating research into action*. Virginia: ASCD.
- McLaughlin, M. (2000). Access to the General Education Curriculum: Paperwork and Procedure or Reinventing Special Education. Retrieved from online *CASE Journal of Special Education*.
- National Research Council. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- National Research Council. (1999). *Testing, teaching, and learning: A guide for states and school districts*. (Committee on Title I Testing and Assessment, R. F. Elmore & R. Rothman, eds). Board on Testing and Assessment, Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.
- Oakes, J. (1990). *Multiplying inequalities: The effects of race, social class, and tracking on opportunities to learn mathematics and science*. Santa Monica, CA: The RAND Corporation.
- Olson, B., Mead, R., & Payne, D. (2002). *A report of a standard setting method for alternate assessments for students with significant disabilities* (Synthesis Report 47). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Available at <http://www.cehd.umn.edu/nceo/OnlinePubs/Synthesis47.html>
- Olson, L., (2002). Ed. dept. hints Idaho's novel testing plan unacceptable, *Education Week*, February 6, 2002.

Orr, C., Chin-Chance, S., Rabinowitz, S., & Vukminovic, Z. (2003, June). *Integrating the elements of Annual Yearly Progress: Vertical scaling and standard setting*. Presentation at the Council of Chief State School Officers Large Scale Assessment Conference, San Antonio, TX.

Paulson, F.L., Paulson, P.R., & Meyer, C.A. (1991). What makes a portfolio a portfolio? *Educational Leadership*, 48, 60-63.

Pellegrino, J.W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.

Pemberton, J.B. (2003). Communicating academic progress as an integral part of assessment. *Teaching Exceptional Children*, 35(4), 16-20.

Petit, M. (2003, January). *Bridging the gap between large-scale and classroom assessment*. Presentation at the National Research Council, Washington, DC.

Quenemoen, R.F., & Almond, P. (2001). Gray areas of assessment systems. *Perspectives* (Journal of the International Dyslexia Association), 27(4), 15-19.

Sargent, J. (2001). *Data Retreat Facilitator's Guide*, Naperville, IL: NCREL.

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.

Shinn, M.R., & Bamonto, S. (1998). Advanced applications of curriculum-based measurement: "Big ideas" and avoiding confusion. In M.R. Shinn. (Ed.), *Advanced applications of Curriculum-Based Measurement*. New York, NY: The Guilford Press.

Shriner, J.G., & Thurlow, M.L. (1993). *State special education outcomes 1992: A report on state activities in the assessment of educational outcomes for students with disabilities*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Stecker, P.M., & Fuchs, L.S. (2000). Effecting superior achievement using curriculum-based measurement: the importance of individual progress monitoring. *Learning Disabilities: Research & Practice*, 15(3), 128-34.

Stiggins, R.J. (2001). The unfulfilled promise of classroom assessment. *Educational Measurement: Issues and Practice*, 20(3), 5-15.

Thompson, S.J., Johnstone, C.J., & Thurlow, M.L. (2002). *Universal design applied to large scale assessments* (Synthesis Report 44). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Available at <http://www.cehd.umn.edu/nceo/OnlinePubs/Synthesis44.html>

Thompson, S.J., Thurlow, M.L., Quenemoen, R.F., & Lehr, C.A. (2002). *Access to computer-based testing for students with disabilities* (Synthesis Report 45). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Available at <http://www.cehd.umn.edu/nceo/OnlinePubs/Synthesis45.html>

Thurlow, M., Quenemoen, R., Thompson, S., & Lehr, C. (2001). *Principles and characteristics of inclusive assessment and accountability systems* (Synthesis Report 40). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Available at <http://education.umn.edu/nceo/OnlinePubs/Synthesis40.html>

Thurlow, M.L., Wiley, H.I., & Bielinski, J. (2002). *Biennial performance reports: 2000-2001 state assessment data*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Available at <http://education.umn.edu/nceo/OnlinePubs/BPRsummary.12.29.02.pdf>

U.S. Department of Education (2002). *A new era: Revitalizing special education for children and their families*. Washington, DC: Author

Vaughn, S., Linan-Thompson, S., & Hickman, P. (2003). Response to instruction as a means of identifying students with reading/learning disabilities. *Exceptional Children*, 69 (4), 391-409.

Vinovskis, M.A. (1996). An analysis of the concept and uses of systemic educational reform. *American Educational Research Journal*, 33, 53-85.

Walmsley, S.A., & Allington, R.L. (1995). Redefining and reforming instructional support programs for at-risk students. In R.L. Allington & S.A. Walmsley (Eds.), *No quick fix: Rethinking literacy programs in America's elementary schools* (pp. 19-44). Newark, DE / New York: International Reading Association and Teachers College Press.

Wiener, D. (2002). *Massachusetts: One state's approach to setting performance levels on the alternate assessment* (Synthesis Report 48). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Available at <http://education.umn.edu/nceo/OnlinePubs/Synthesis48.html>

Wiggins, G., & McTigue, J. (1998). *Understanding by design*. Alexandria, Virginia: Association for Supervision and Curriculum Development.

This document is provided for the user's convenience. Inclusion does not constitute an endorsement by the U.S. Department of Education of any views, products or services offered or expressed.

*Please note: at time of publication all links were current. To maintain the fidelity of the original resource, we have kept the links to show original sources