NATIONAL COMPREHENSIVE CENTER
FOR TEACHER QUALITY

Linking Assessment and Instruction: Teacher Preparation and Professional Development

DECEMBER 2010

TQ Connection Issue Paper

Linking Assessment and Instruction: Teacher Preparation and Professional Development

A TQ Connection Issue Paper on
Improving Student Outcomes in General and Special Education

December 2010

John L. Hosp, Ph.D.
University of Iowa

The author wishes to acknowledge the expertise and valuable feedback from those who helped to review earlier drafts of this Issue Brief and Innovation Configuration: Kimberly A. Gibbons, Ph.D., St. Croix River Education District, Rush City, Minnesota; Leeann S. Hawken, Ph.D., University of Utah; Bonnie D. Jones, Ed.D., Office of Special Education Programs, U.S. Department of Education; and Paul T. Sindelar, Ph.D., University of Florida.

NATIONAL COMPREHENSIVE CENTER FOR TEACHER QUALITY

CONTENTS

OVERVIEW

Making decisions about instruction is as core a component to teaching as providing the instruction itself. When providing services to students who are at risk for poor educational outcomes or students with disabilities, it is especially salient to ensure that these instructional decisions have the highest likelihood of accuracy as possible and will lead to improving those outcomes. The students with the greatest needs require the most accurate and effective decisions. In addition, recent increases in the need for accountability have put additional pressure on teachers to document their decisions and decision-making processes. Now, more than ever, effective use of assessment data to plan, judge, and modify instruction is a fundamental competency for good teaching.

The purpose of this Issue Paper is to provide a framework and justification for effective ways that teachers can collect and use assessment data to make instructional decisions. This framework is provided as an indication of what effective linking of assessment data to instructional decisions *ought* to look like—rather than a summary or survey of current practices. The framework and respective Innovation Configuration for Linking Assessment and Instruction in Teacher Preparation and Professional Development (provided in the Appendix, pages 31–34) are primarily designed to provide a blueprint for preservice teacher preparation; however, they also may be used as an evaluation rubric or development guide for inservice professional development. Although many schools and districts may not currently have in place the practices discussed in this Issue Paper, these practices are strongly endorsed by the requirements of the 2002 reauthorization of the Elementary and Secondary Education Act (ESEA)—also known as the No Child Left Behind Act—and the competitive grants to states that were made available through the Race to the Top Fund.

This paper begins with a discussion of why assessment and instruction should be linked. It continues with an overview of the innovation configuration, describing essential components in preservice and inservice teacher training to identify the skills and competencies that teachers need to make sound decisions about using assessment information to improve instruction. Next, the major points within the innovation configuration are provided, with a rationale for their importance and elaboration of some of their core characteristics. Last, recommendations are provided regarding how the components of the innovation configuration might be included in teacher preparation and professional development practices.

THE IMPORTANCE OF LINKING ASSESSMENT AND INSTRUCTION

There are different arguments for why assessment and instruction should be closely linked or aligned—some legal, some ethical, and some practical. Each of these reasons is discussed below.

The legal basis for linking assessment and instruction is that federal laws and state regulations have shown an increase in the requirements of collecting assessment data and use of those data for accountability purposes at the state, district, school, teacher, and student levels (Salvia, Ysseldyke, & Bolt, 2010). The 2002 reauthorization of ESEA mandated that assessment is to be used to evaluate schools, districts, and states. Accountability in teacher performance or quality also is being advanced through the Higher Education Opportunity Act of 2008; the influence of this law typically is at the teacher level. At the individual level, the Individuals with Disabilities Education Act (IDEA) of 2004 mandates different types of assessment to document effectiveness for individual students as well as for programs. Although these laws are clearly important influences in the assessment practices of teachers, they are not forces that generally drive day-to-day instructional decisions; nor are many of the assessment methods required by federal or state laws or regulations useful in making decisions about what to teach or how well students are learning the presented material.

The ethical basis for linking assessment and instruction is that most professional organizations include assessment and the use of assessment data to make decisions in their guidelines for ethical and best practices as well as training. As examples, organizations for reading teachers (International Reading Association & National Council of Teachers of English, 2010), mathematics teachers (National Council of Teachers of Mathematics, 1995), and special education teachers (Council for Exceptional Children, 2003) all provide standards for training and practice in the use of assessment. (Although the necessity and role of high-stakes testing is addressed in each of these guidelines, the primary focus is the use of assessment data to make decisions about teaching and learning; this focus is the embodiment of the practical reason for linking assessment and instruction.)

The practical basis for linking assessment and instruction is that teachers need to make screening, progress, diagnostic, and outcome decisions—each of which should link assessment and instruction. In addition, teachers need to make these instructional decisions frequently. Estimates have put the number of instructional decisions that teachers make each day at 1,300 (Jackson, 1968) with about 10 significant, interactive decisions per hour (McKay, 1977), but empirical work also has identified that teachers make 9.6 to 13.9 instructional decisions per lesson (Morine-Dershimer & Vallance, 1975). However, Peterson and Clark (1978) reported that instructional decisions were made only when instruction was not effective; they also indicated that changes were made in only half of the situations in which students were not learning sufficiently. Much of the research on the frequency of teacher decision making was conducted in the 1970s and ’80s (for reviews, see Clark & Peterson, 1986; Shavelson & Stern, 1981). Since that time, the focus of research has changed.

Research on teacher decision making since the early 1980s has often focused on the outcomes of those decisions. The most common outcome is that when teachers use assessment data to make their instructional decisions, student performance increases (Black & Wiliam, 1998;

Fuchs & Fuchs, 1986). The students of teachers who collect systematic progress-monitoring data (and use it to make decisions) score on average a full standard deviation higher than their student peers whose teachers do not collect and use these data (Fuchs, Fuchs, Hamlett, & Allinder, 1991; Stecker & Fuchs, 2000; Wesson, 1991). In addition, teachers using systematic progress-monitoring data make instructional changes more frequently for their students who are experiencing difficulties (Fuchs, Fuchs, Hamlett, & Stecker, 1991). Given the current focus on accountability and outcomes in education, training preservice and inservice teachers to more effectively and efficiently collect and use assessment data to make instructional decisions for their students and classes should be a core component of any type of teacher preparation and professional development.

Description of the Innovation Configuration for Linking Assessment and Instruction in Teacher Preparation and Professional Development

This Issue Paper presents the Innovation Configuration for Linking Assessment and Instruction in Teacher Preparation and Professional Development, which can be used to evaluate general and special education preservice teacher preparation or inservice professional development in terms of content relevant for linking assessment and instruction. This innovation configuration is provided in the Appendix (pages 31–34).

An innovation configuration is a matrix that typically identifies and describes the critical components of a practice that is important to training within a field. The matrix consists of two dimensions: essential components and degree of implementation (Hall & Hord, 1987; Roy & Hord, 2004). The essential components typically are listed as the row headings of the matrix within the leftmost column; additional descriptors or subcomponents also are included for clarification and use with more specific evaluations. The degree of implementation typically is presented as column headings in the topmost row, with multiple levels of implementation specified—ranging from zero (no mention) through progressively higher scores to a maximum that is used to represent exemplary inclusion and implementation of the component. Innovation configurations have been used for more than 30 years as tools to develop, implement, and evaluate education innovations (Hall, Loucks, Rutherford, & Newton, 1975).

The innovation configuration presented in this Issue Paper is designed to provide educators with a tool to evaluate the degree to which their preparation or professional development activities incorporate evidenced-based practices for linking assessment and instruction. It is designed for use with general education teachers, instructional specialists or coaches, special education teachers, paraprofessionals, other specialists or related service providers (e.g., school counselors, school psychologists, speech-language pathologists), or education administrators. Some components of the innovation configuration may be important to elaborate upon and adapt for some specialties, but all components are important considerations for *all* educators.

COMPONENTS OF THE INNOVATION CONFIGURATION

The essential components of the Innovation Configuration for Linking Assessment and Instruction in Teacher Preparation and Professional Development are as follows:

- Fundamentals of assessment
- Standards for comparison of performance
- Considerations for decision making
- Assessment procedures
- Identification of content to teach
- Identification of student response

These six components are based on the research and best practice literature detailing how assessment and instruction can be linked as well as important considerations in assessment and instruction. The following sections briefly describe each component. As stated previously, training for specific roles may warrant additional elaboration of some of the components and some details may vary by the grade level of students with which the educators are being trained to work, but these six components should be addressed in any system of training for educators. Preparation in these components establishes a fundamental competency that is critical for teaching—particularly with at-risk students and students who struggle with academic achievement.

Fundamentals of Assessment

This component consists of fundamental information about assessment and measurement—topics such as reliability and validity, types of scores that might be produced through assessment and their interpretation, legal provisions regarding assessment, issues of cultural and linguistic diversity, statistical bias and fairness, and accommodations and modifications for use with students with disabilities and English learners. This component also consists of information on the types of decisions that teachers and other educators routinely make. These topics are generally covered in any college-level introductory assessment text (e.g., Miller, Linn, & Gronlund, 2008; Popham, 2010; Salvia et al., 2010). As such, these topics will not be detailed here; however, they are important for linking assessment and instruction—particularly when selecting instruments to collect the assessment information on which to base the decisions.

Although many definitions exist, *assessment* is generally considered as the process of collecting information for specific purposes. Within the framework of evaluation or decision making, assessment information can aid in making four types of decisions: screening, progress, diagnostic, or outcome (J. L. Hosp, in press). *Screening* decisions relate to which students are expected to be successful or proficient at the end of the year and which are not. *Progress* decisions relate to whether individuals or groups of students are learning at a sufficient rate to demonstrate proficient end-of-year performance. *Diagnostic* decisions relate to what to teach and how to teach it. *Outcome* decisions relate to which students have or have not met the criterion for proficiency. All four types of decisions should be included in a comprehensive system of linking assessment to instruction and in the preparation of teachers.

Standards for Comparison of Performance

After a student's performance has been measured, a key component to making decisions about his or her performance and planning instruction is the teacher's ability to make comparisons to a standard for performance. Three ways of determining standards are typically used in education: normative, criterion, and ipsative.

Normative standards involve comparing a student's performance on the assessment to that of other students in a comparable peer group. This comparison might be made to other students in the same grade (e.g., 3rd grade), other students taking similar coursework (e.g., high school biology), other students of the same age (e.g., 3-year-olds), or other students with similar demographic characteristics (e.g., students with disabilities).

Criterion standards involve comparing a student's performance to an empirically derived level of proficiency (i.e., a cut score that is used to determine whether or not a student has sufficiently mastered the material). For example, high-stakes accountability tests have cut scores to determine whether or not a student has reached proficiency in a particular area. Typically, these tests are criterion referenced. Another example would be if the core curriculum indicates that students in Grade 1 should be able to compute basic subtraction facts with 90 percent accuracy. This benchmark provides a criterion when giving Grade 1 students a sheet of basic subtraction problems and having them work the problems to determine how many they get correct.

Ipsative standards involve a student's prior performance as the basis for comparison of his or her current performance. Ipsative standards often are used for goal setting and motivation. For example, if a child completed a task such as finishing a sheet of independent-level work (i.e., work the child can perform accurately without support or guidance) in 20 minutes, the teacher could ask the student to complete the task again but try to do it more quickly (i.e., completing it in less than 20 minutes or with a specific goal of 18 minutes). Ipsative standards often are considered when monitoring student progress because the student's current performance can be compared to prior performance (yesterday or last week) as well as future performance (tomorrow or next week).

Considerations for Decision Making

The term *assessment* can have different meanings. It can refer to a specific task or test, the process of assigning numbers to characteristics of people or objects, or the process of making decisions. One way to keep these multiple usages distinct is to use other terms—such as *instrument* to refer to a specific assessment task or test, *measurement* to the process of assigning numbers, and *evaluation* to the process of making decisions. In this framework, the term *assessment* refers to the process of collecting information through measurement (conducted using instruments) for the purpose of evaluation (J. L. Hosp, 2008).

Inside and Outside Decisions

Of course, decision making has many different purposes. A useful framework is to consider these purposes as inside the classroom or outside (J. L. Hosp, in press).

Inside classroom decisions are those that are directly relevant for instructional planning or the day-to-day operations of a classroom. Examples of inside decisions are grouping students for small-group instruction, determining whether or not a student or group of students is making adequate progress, or deciding which method to use to teach a concept or skill.

Outside decisions are those that do not directly impact daily instructional planning. This distinction should not imply that such decisions are not important but only that they do not have a direct or immediate impact on the teaching within a classroom. Such decisions typically are not made by individual teachers but rather are made by groups of which teachers may be members. Examples of outside decisions are student eligibility for specialized programs or services, changes to ensure adequate yearly progress (AYP) of classrooms or schools, or core programs to adopt throughout a school or district.

Summative and Formative Decisions

One of the distinctions occasionally made about types of decisions is the summative/formative dichotomy. These decisions are sometimes considered as summative and formative *assessments* (e.g., Black & Wiliam, 1998; Shepard et al., 2005) and sometimes as summative and formative *evaluation* (Airasian & Madaus, 1972; Fuchs & Fuchs, 1986; Howell, Hosp, & Kurns, 2008).

Summative decisions are made at a single point in time to summarize the learning or performance of a student or group of students. For example, high-stakes tests administered at the end of a school year are for the purpose of summative outcome decisions—determination of whether or not each student met the criterion for mastery of that year's curriculum standards and determination of AYP of the school or district.

Formative decisions are those to help teachers provide the most effective instruction to their students. For example, a curriculum-based measurement of oral reading fluency can be administered once per week to those students experiencing difficulty in order to determine the effectiveness of instruction. When the progress-monitoring data indicate that a student is not learning at a sufficient rate to be proficient by the end of the school year, the educator can alter the instruction to better meet the student's needs.

Decisions for Interim Assessments

Some purposes, however, do not fit the summative-formative dichotomy, requiring the addition of another term—*interim* assessments—to bridge the gap (Perie, Marion, & Gong, 2007). Interim assessments are given less frequently than formative assessments but with more relevance for teaching decisions than summative assessments. As such, they might encompass periodic benchmark or screening assessments. Within this summative/formative framework, *summative* can be conceived of as assessment or decisions *of* learning, whereas *formative* is assessment or decisions *for* learning (Torgesen & Miller, 2009).

Within the framework of evaluation, there is no need to consider "interim" decisions because these assessments would fall under *formative* or *summative*, depending on their frequency and purpose. Within the framework of assessments, however, interim assessments would address a

little bit of both formative and summative characteristics. Interim assessments are administered at periodic intervals to gain snapshots of student performance, but they also can provide some feedback that is useful for instructional planning. For example, benchmark screening measures administered to all students in the fall, winter, and spring can be used for summative decisions about student learning and the effectiveness of instruction; but they also may provide feedback on which students need additional support or which areas of the content need more instruction.

Needs-Based Decision Making

In the context of decision making, teachers have many different needs for making decisions. Some classroom decisions are quick and made immediately (e.g., whether or not to praise a child, which student to call on for response, whether or not to repeat directions). Other decisions require more upfront planning in the collection of data. When a decision has high stakes associated with being wrong (i.e., making an incorrect decision), teachers have an increased need for enough information to make a good decision (J. L. Hosp, 2008). In this case, use of a structured set of procedures for collecting information and making decisions can be useful. Two structured approaches are curriculum-based evaluation (Howell, Hosp, & Kurns, 2008) and the standard treatment protocol approaches of response to intervention (RTI; see Jimerson, Burns, & VanDerHeyden, 2007). These approaches provide explicit guidelines and decision rules for determining what types of information to collect, why it needs to be collected, and how to make decisions—all with explicit links to providing instruction.

Assessment Procedures

Many educators often equate assessment with testing. Yet in meeting the demands of collection and use of information to make decisions about instruction, teachers need to think more broadly about what constitutes assessment. In preservice teacher preparation and inservice teacher professional development, there are many variations in the specific instruments used to collect information. Different procedures are required for measuring reading at the elementary level than mathematics in high school or behavior in early childhood, for example. All methods of assessment can be considered within one of four different categories: review of information, interview, observation, and testing—which fits into the handy rubric, RIOT.

Review of information includes collecting and systematically organizing information that has been collected previously about a student—such as records from his or her cumulative folder, prior test results, and work samples. *Interview* involves talking to others who have knowledge of the student and his or her performance. These people might be other teachers, related service personnel, the student's parents or siblings, and the student himself or herself. Such interviews can be highly structured and even standardized in their administration and scoring, or they can be unstructured or more informal in nature. *Observation* is watching the student perform a task, typically in the learning environment (such as the classroom). Some observations methods are appropriate for classroom teacher to use for collecting observation data on students during instruction; other methods are more appropriate for an external observer to come into the classroom to collect data (Shapiro & Kratochwill, 2000). Similar to interviews, observations can be highly structured or unstructured, depending on the need for information on which to base

decisions. *Testing* is the most common understanding of assessment. It includes methods ranging from informal inventories to individually administered norm-referenced tests.

The RIOT (review of information, interview, observation, and testing) assessment procedures often are discussed in conjunction with different evaluation domains—or areas about which educators need to make decisions. The acronym SCIL refers to the domains of setting, curriculum, instruction, learner (J. L. Hosp, in press); the acronym ICEL is used to refer to the domains of instruction, curriculum, environment, and learning—the same domains (except that *setting* is replaced by *environment*) but differently ordered. *Setting* (or *environment*) refers to where the learning is expected to occur and various characteristics that might be alterable by a teacher in order to facilitate learning. *Curriculum* refers to what is being taught and what the students are expected to learn within the grade or age level. *Instruction* is how the content is being delivered. *Learner* refers to individual student characteristics that might be important to designing instruction.

Educators typically focus much of their decision making on the learner, when it might be more efficient to focus on other factors in addition to the learner. For example, an individual student might be having difficulty learning addition facts and his teacher might devote more time working with him to learn those facts. However, by focusing her assessment on the whole class or grade level, she might determine that a majority of students are having difficulty with addition facts and decide that this situation is due to the new mathematics program not placing enough emphasis on this skill. Therefore, the best solution may reside at the curriculum level rather than with individual learners. Decisions about teaching should incorporate information about the setting, curriculum, and instruction as well as information about the learner. All the RIOT procedures can be useful in considering how to collect the appropriate information to make these decisions.

Identification of Content to Teach

Within the confines of the general classroom and the general curriculum, certain externally predetermined standards indicate what every child is expected to learn within a grade level or at a certain age level. These standards may be the state's core curriculum or standards for grade-level learning. The majority of students will be held to these standards and most likely progress through the expectations at a fairly typical rate. For those students who are not progressing through the curriculum, however, it is important to identify those areas in which they are having difficulty and need extra instruction.

The first step is to compare the student's performance in each broad content area. In the elementary grades, the state or district probably has expectations within areas such as reading/language arts, mathematics, science, and social studies/history, for example. The student's performance should be compared to two different standards—how his or her performance compares to the cutoff for proficiency or mastery (criterion) and to the performance of other students in the classroom (normative). If the student's performance is below the criterion for acceptable performance, he or she needs additional instruction in that area. If the student's performance is similar to the peers' performance (and below the criterion), changes to instruction should involve the entire class and the general, or Tier I, instruction.

Example

As an example, an entire class of Grade 2 students is screened using a curriculum-based measurement (CBM) for mathematics computation. One student of interest has performed below the criterion. The student calculates 10 correct digits in 2 minutes, indicating performance at a "frustrational" level (Burns, VanDerHeyden, & Jiban, 2006). Upon examining the performance of the rest of the class, the teacher notes that 13 of the student's 25 peers also scored in the frustrational range (a total of 14 students in the frustrational range) and the student's score is at the 50th percentile for the class. Rather than developing intervention strategies that focus on that individual student, the teacher examines the broader curriculum (what content to teach) and develops lessons based on providing instruction to larger groups of students or possibly the entire class.

Upon examining student performance, the teacher finds that many students are still having difficulty with addition facts, which makes it likely that they will have trouble with more complex addition problems. Using addition fact-specific CBMs, she finds that eight of the students know the addition facts accurately but cannot compute them fluently and the other six do not yet know their addition facts. She decides to break the class into smaller groups for some of their mathematics time; she will work on accuracy of addition facts with one group and fluency of addition facts with the other.

If the performance of the student of interest is below that of his or her peers (as well as below the criterion), the instruction should be supplementary. Conducting additional assessments is necessary to determine more specifically where the breakdown in learning is occurring.

Skills to Be Examined

When a student's performance is significantly below the criterion for acceptable performance as well as his or her peers' performance, it is necessary to identify more specifically what difficulty the student is experiencing. This area of decision making can encompass three types of skills to examine: prerequisites, related skills, and subskills.

Prerequisites are abilities that the student must have in order to perform the task at hand, but they are not necessarily skills that would be taught previously. This term includes visual acuity (i.e., being able to read the materials), language proficiency, and other personological characteristics that may impact the student's ability to access the learning materials. Such characteristics are important to the learning action and might need to be accommodated in order to allow the student access. For example, a student with poor vision might need to wear corrective lenses, sit closer to the board, or have larger print materials. These interventions would accommodate the prerequisite of being able to see the materials.

Related skills are skills that the student must be able to perform or areas of knowledge that the student must have mastered, which are related to the content area of interest but are not included within it. Such skills often should have been taught or learned previously but in

a different content area. For example, many mathematics instructional materials require reading skills. The student must read the problems in order to derive the information for computation or application. Reading is not a component of mathematics, per se, but is important when students need to solve story problems, geometry theorems, or other mathematical applications within sciences such as biology or physics. As such, being able to decode the text in order to comprehend the information contained therein and associate it with one's vocabulary and prior knowledge covers a series of related skills and subskills.

Subskills are skills that are actually components of the content area of focus that must be learned before being able to master that content. They are sometimes derived through a task analysis of a skill (i.e., explicit identification of the subskills necessary to complete it) or through an explicit scope and sequence of a curriculum. For example, the student experiencing difficulty in mathematics may actually be having a specific difficulty with computation—particularly with double-digit addition with regrouping. This subskill is a relatively specific skill within the curriculum; however, there are other subskills that are critical to being able to add two double-digit numbers with regrouping. The student must understand the concepts of regrouping, conservation of quantity, and place value. The student must know procedures for regrouping and column addition. The student must have number sense and know basic addition facts as well as understand the concepts behind and procedures for adding two numbers.

Forms of Knowledge

When considering which procedures to use to collect information about content areas, prerequisites, related skills, or subskills, teachers must ensure that the assessment procedures used are aligned with the form of knowledge that is expected: fact, concept, or strategy (Howell & Nolet, 2000).

Facts (also called *rote* or *declarative knowledge*; see Marzano et al., 1988) are types of information that are discrete and stand alone. For example, knowing that the capital of the United States is Washington, D.C., does not give any information about the capital cities of states within the United States, capital cities of other countries, or details about Washington, D.C., such as where it is, how to get there, or how many residents it has.

Concepts are groups of objects, events, or actions that share a set of distinguishing characteristics. These characteristics are generally defined through rules for differentiating examples and nonexamples of the concept. For example, the concept of "squares" would be defined by the following rules: two-dimensional figure, four sides of equal length, and four right angles where the sides meet. Nonexamples would include *near distracters* (i.e., those that are similar in that they share one or two rule-traits but not all—such as a rectangle) and *far distracters* (i.e., those that share few or no rule-traits—such as a sphere).

Strategies often are defined as processes of work rather than products (Marzano et al., 1988). As such, they can be considered knowledge of how to do something or procedures for its demonstration. Strategies involve applying or generating other forms of knowledge (i.e., facts and concepts). In mathematics, for example, there are strategies for conducting numeric operations; in reading, there are strategies for decoding a word that the reader does not recognize. Such strategies are procedures for conducting an action or solving a "problem" of sorts.

To put all these ideas together, consider the case of trying to determine the area of a circle. Concepts involved include knowledge of what a circle is and that mathematical equations can be used to represent physical attributes. Facts involved would be the equation for determining the area of circle, multiplication facts, and the value of π. Strategies involved would be to find the radius of the circle and substitute that for *r* in the equation as well as the process of solving the equation (which involves application of facts such as when to multiply π by r^2 and to square *r*—i.e., multiply it times itself). So the smooth performance of this seemingly simple activity requires the learner to combine different forms of knowledge in rule-governed ways but also to know when and how to apply them.

Structured Systems of Evaluation

There are a few approaches for putting together these types of information and decisions into a structured system of evaluation. *Instructional assessment* (Gravois & Gickling, 2008), which is sometimes referred to as *curriculum-based assessment for instructional design* (Burns & Mosack, 2005), is an approach that relies heavily on subskill mastery measurement to align a student's prior knowledge to the instructional tasks and level of difficulty. *Curriculum-based evaluation* (Howell, et al., 2008; Howell & Nolet, 2000) is an approach that emphasizes the nature of thinking and decision making in a structured fashion. Some approaches to RTI also fall into the category of structured systems of evaluation through the use of standard protocols, particularly when a student is having difficulty and has not responded sufficiently to previous instruction and intervention (Jimerson, Burns, & VanDerHeyden, 2007). All these approaches share some common features of problem solving and data-based decision making, yet they each manifest in different ways—sometimes to achieve different ends.

Skill Deficits and Performance Deficits

When a student does not perform a task or subskill to proficiency (i.e., above the criterion for acceptable performance), it is important to determine whether the student *cannot* perform the task or *will not* perform the task—because remediation of each situation requires different instructional methods (Noell et al., 1998). Determining if the student's difficulty is the result of a skill deficit or a performance deficit (Gresham, 1981; J. L. Hosp & Ardoin, 2008) is important.

A *skill deficit* occurs when the student is not able to perform the task at the level of proficiency required for successful performance. A *performance deficit* occurs when the student does not have sufficient motivation to perform the task at a proficient level or to sustain performance enough to complete a task. When exhibiting a performance deficit, the student is capable of performing the task when there is sufficient motivation but the difficulty lies within generating the motivation. Note that although it is possible that some students actively decide to *not* perform a task, more often there are other reasons that negatively impact the student's motivation. Identification of a performance deficit should not be used to automatically indicate that a student is willfully not performing.

It also is possible that a student exhibits a *combined skill and performance deficit,* wherein the student cannot quite perform the task to proficiency but also has difficulty sustaining motivation to perform the task. The type of performance deficit can be distinguished through the use of a "can't do/won't do" assessment (VanDerHeyden & Witt, 2008). This approach uses repetition of the task (using parallel materials) combined with implementation of reward conditions in order to determine whether or not the student cannot or will not perform the task to proficiency.

Stages of Learning

In addition to determining whether a student cannot or will not perform the task to proficiency, the teacher or educator should consider the stage of learning at which the student can perform the task (Idol, 1989). The stages of learning are sometimes referred to as the *instructional hierarchy* (Haring & Eaton, 1978) and are related to the work of Benjamin Bloom (1971). Students go through five stages or levels of learning before mastering a task or skill.

As a student begins to learn a task, he or she is in the *acquisition stage*. This stage is marked by the student becoming increasingly accurate at performing the task. After achieving accuracy of 90 percent to 100 percent, the student moves into the *proficiency* or *fluency* stage, which is marked by high accuracy as well as an increasing rate of performing the task (i.e., being able to perform the task more quickly while maintaining high accuracy). Next, the student enters the stage of *maintenance*, which is marked by retention of high rate and accuracy. Then the student moves to the next stage, *generalization*. This stage is marked by the student beginning to transfer performance of the task to new settings or applications. Last, the student enters the stage of *adaptation*, wherein he or she is able to capitalize on the knowledge and use that knowledge to solve problems in various settings—particularly using new or novel applications of the task.

One reason the stages-of-learning approach is important to consider in assessment is that if the student is in the accuracy stage of learning and can perform the task with 70 percent accuracy, yet the instrument being used to measure the student's performance requires performance at rate (i.e., at the proficiency or fluency stage), the assessment results might suggest that the student cannot perform the task, when in reality the student can perform the task but at a different level of learning. It is especially important to consider when the assessment requires a late stage of demonstration (i.e., generalization or adaptation) and the student is in the early stages of learning (i.e., acquisition or proficiency).

Individualized Education Programs

If the student has an individualized education program (IEP), a Section 504 plan, or any other document that explicitly determines education goals and objectives, the methods of assessment must align with the student's goals and objectives. Preservice teachers should learn what types of plans or documents might exist for their future students and know where to find them. They also should know which specialists in their school would be primarily responsible for these plans or documents (if they are not the ones responsible). The state laws and rules guiding development of these documents vary from state to state, so situating the preservice training (or inservice professional development, particularly for new teachers coming from out of state or district) in the laws and regulations specific to that state and district is important.

Judgments of Student Work

During the course of a typical school day, students generate a lot of work—some of it transitory (e.g., oral responses to questions that are not recorded or written down) and some of it permanent (e.g., written or audio- or video-recorded work). Good teachers are always looking at (or listening to) student work with an evaluative focus to judge the sufficiency of the student's performance. Much of the time, this evaluation is informal—including subjective judgments of quality, inferences about

the difficulty of the task for the student, and determinations of whether or not the work was completed within the allotted time. Although all of these on-the-spot evaluative judgments may be incorporated into the teacher's overall impression of the student's performance, sometimes it is important to use more standard judgments of student work in order to include the permanent products into the student's cumulative folder or to share it with others who are involved in decision making about the student (e.g., parents, related service personnel, administrators).

Example

As an example of linking assessment to identification of content to teach, consider the case of a 10th-grade student (Hubert) having difficulty in an American History class. At the beginning of the year, the teacher (Ms. Washington) gives a test of content from the year's curriculum to all students to determine their prior knowledge of the material. The test is a screening decision using evaluation characteristics that are both summative (determining prior knowledge) and formative (determining a baseline for all students and identifying gaps in knowledge). This test gives Ms. Washington an idea of what the students in her class already know, but it also serves as a guide for what she will need to teach during the year. During the first month of school, Ms. Washington gives weekly quizzes to all students to monitor their progress and make formative decisions about the effectiveness of her instruction. She notices Hubert does not participate in class and has failed every weekly quiz. Ms. Washington decides to review Hubert's records to evaluate his prerequisite skills, which she determines are important. She finds that his vision and hearing are both excellent and that his attendance is good. She also evaluates related skills that might impact his performance. She notes that no previous teacher has documented a difficulty with attention or focus and that his reading skills (particularly comprehension) are good.

At this point, Ms. Washington decides to examine the specific subskills she has been focusing on in the American history class. There have been two main foci: facts (such as names, dates, and locations of colonial America) and concepts (such as colonialism). On measures of American history facts, Hubert scores above 90 percent, can recall the facts at rate, and is doing so for the facts from the prior units. This result suggests to Ms. Washington that Hubert's learning of these facts is at a maintenance stage of learning and is where she expects it to be (i.e., it is similar to that of other students in the class). On measures of the conceptual information, however, Hubert has difficulty identifying the core characteristics of the concepts as well as providing nonexamples. This result suggests to Ms. Washington that Hubert is having difficulty acquiring the conceptual knowledge that she is teaching. Next, she wants to determine whether this difficulty represents a skill deficit or a performance deficit, so she uses "can't do/won't do" procedures with Hubert. Ms. Washington determines that his difficulties arise from a skill deficit—he is having trouble grasping the concepts involved. Now she understands that she needs to provide Hubert with additional instruction in acquiring the concept of colonialism.

Identification of Student Response

Assessment for identification of the content to teach is primarily about determining a student's *level of performance* in different areas or with different skills. When comparing this performance to various standards, the teacher typically collects the assessment at a single point in time to describe the student's performance. The teacher also needs to consider how the student's performance changes over time. This information is collected through assessment of student *progress*. Progress decisions (or progress assessment) are one of the types of decisions (J. L. Hosp, in press) included in comprehensive frameworks for assessment and decision making in education that are typically included with the fundamentals of assessment (see Salvia et al. 2010). Monitoring student progress and making progress decisions are core features of RTI (Reschly & Wood-Garnett, 2009). This type of formative evaluation is really the driving force for linking assessment and instruction because it represents decision making for learning—that is, decisions used to plan instruction (Torgesen & Miller, 2009). (See "Attributes of Progress-Monitoring Instruments Used to Identify Student Response" below.)

Attributes of Progress-Monitoring Instruments Used to Identify Student Response

The specific choices of instruments used to collect information on student progress will differ by content area (e.g., reading, mathematics) and by the grade or age level of the student (prekindergarten, elementary, secondary). When selecting instruments, preservice or inservice teachers should be aware of many common attributes, including core characteristics (such as reliability and validity), efficiency, consistency (J. L. Hosp, in press).

Core Characteristics

The first consideration is about the *core characteristics* of the instrument—its reliability, validity, and nondiscrimination against subgroups of students (i.e., general fairness and no statistical bias; see National Center on Response to Intervention, 2010, for a review of progress measures). Depending on the level of aggregation being examined (i.e., individuals or groups such as classrooms or grade levels), there are different standards for reliability: .60 or better for group decisions and .80 for individual decisions (Salvia et al., 2010). There also are considerations for different types of reliability. For progress assessment, interrater reliability and alternate form reliability are crucial whereas internal consistency often is not as important. The reliability and validity of both the level and slope scores also should be considered (National Center on Response to Intervention, 2010). An instrument for progress monitoring also should be nondiscriminatory such that it is generally fair in its content and the reliability and validity of the measure are not different for various subgroups of students. Because instruments for progress monitoring typically are developed to be closely aligned with the content that the student is expected to learn (generally measuring the same skills and response type expected in the curriculum), such instruments fare well when examined for nondiscrimination (see National Center on Response to Intervention, 2010, for examples).

Efficiency

The second consideration is *efficiency*. Instruments for progress monitoring should be quick and easy to administer and score (Deno, 2003). In general, if a progress measure requires more than 3–5 minutes per student to administer and score, it will take too much instructional time to be

useful for progress decisions. Progress measures can be used as dynamic indicators of growth over time (Shinn, 2008)—similar to how at every visit to the doctor's office, the patient's temperature, weight, height, and blood pressure are measured; these measures are quick, efficient indicators over time of overall health rather than an in-depth assessment of specific issues. Part of the efficiency of progress measures is consideration for interpretation and communication of performance. The results of many progress measures can be illustrated through the use of graphs. In particular, line graphs are useful for showing change over time. (For an example of a line graph, see Figure 2 on page 19.) With inclusion of a standard for comparison (e.g., the rate of growth that is expected to meet a later goal), interpretation of how the student's progress compares is simple.

Consistency

A third consideration for selection of progress-monitoring instruments is *consistency* of administration, scoring, and materials. This attribute also is referred to as *standardization*. The use of standardized directions and scoring rules enable most instruments to demonstrate good reliability and validity. Such consistency can be compared to weights and measures having standard definitions. (For example, if the length of a foot were allowed to vary among rulers or tape measures, it would be nearly impossible to build things or communicate dimensions of objects.) Use of consistent materials ensures that when the teacher measures growth in student performance, that growth is due to learning and not to changes in the materials. This outcome is especially important with progress assessment because the instruments must be able to be administered frequently to the same student.

Achieving consistency is possible by using the exact same materials—but only if the student is not expected to learn from or remember the specific materials; otherwise, his or her growth could represent a "practice effect" (O'Connor, White, & Swanson, 2007). In most academic areas, consistency of progress materials is achieved through the use of alternate, parallel forms—versions of the same task that include the same form of task at an equivalent difficulty but with different specific items included. In mathematics operations, this parallelism would include the same types of problems (e.g., multiplication facts) but with different numbers. In reading and content areas, it would include the same difficulty of the content but a different focus (e.g., one story on the life of sea turtles, another on whether or not bears hibernate).

One benefit of this consistency is that progress-monitoring instruments must be *sensitive to growth* (i.e., they need to be able to accurately measure changes in performance). When the materials are sufficiently consistent, the teacher can be reasonably certain that the changes are not the result of using different materials or different levels of difficulty of the material but rather from real differences in student performance of the task.

Another benefit of this consistency is that if all students in a class or grade level are doing the same task, under the same conditions, with the same scoring, those data can be used to make *multiple decisions*. The data can be used to make decisions about that individual student, but they also can be aggregated to make decisions about the progress of small groups of students (e.g., different reading groups, English learners), the classroom as a whole (e.g., to determine if the instruction is effective at increasing everyone's performance), or an entire grade level across the school or district (e.g., to judge the adequacy of the curriculum). Although decisions at larger levels of aggregation might be beyond the control of preservice or inservice teachers, such decisions are important considerations and ones to which the teacher can then contribute.

Using Progress Data to Examine the Effectiveness of Curricula and Instructional Practices

Progress data are useful for examining the effectiveness of curricula and instructional practices. Figures 1A and 1B present curriculum-based reading data from two Grade 1 classrooms using the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) nonsense word fluency measure (Good & Kaminski, 2011).

Figure 1A. Grade 1 Nonsense Word Fluency Progress Monitoring in Classroom A

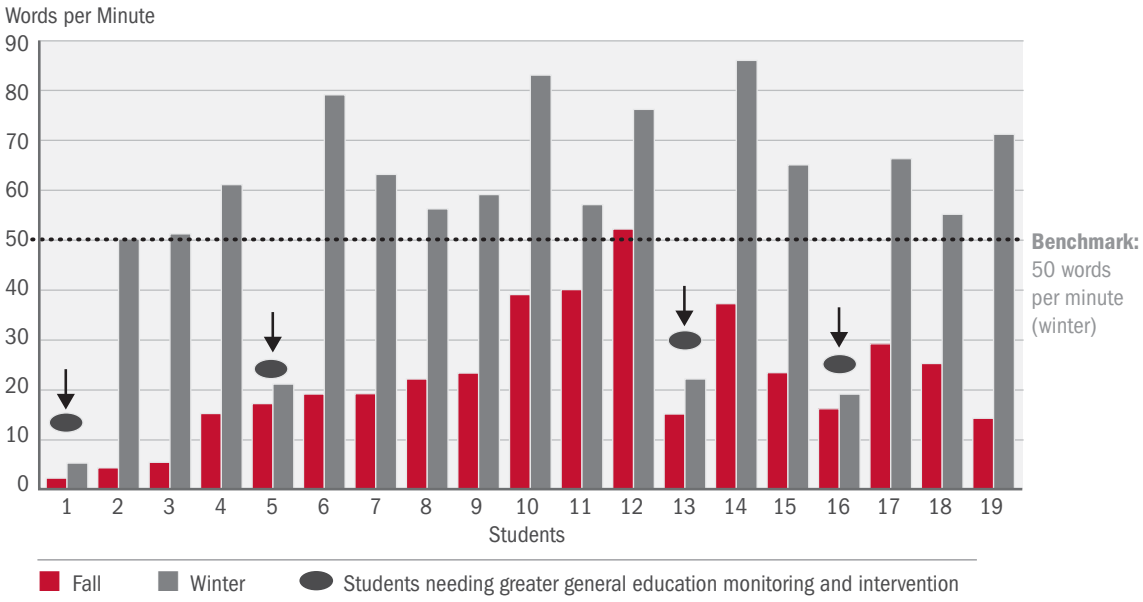
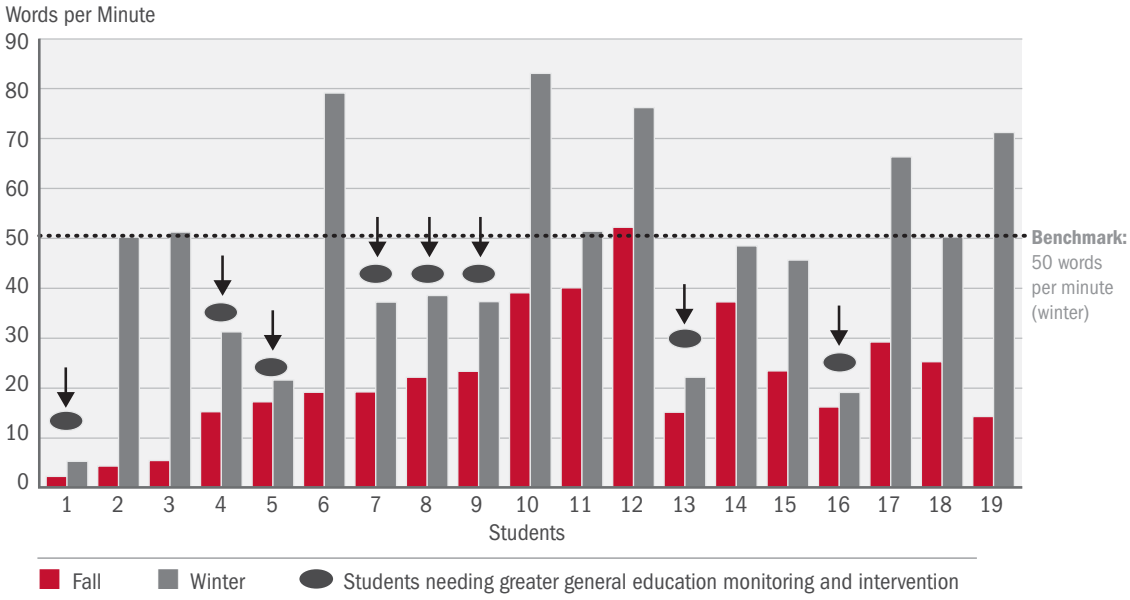


Figure 1B. Grade 1 Nonsense Word Fluency Progress Monitoring in Classroom B



The classrooms are adjacent to each other and draw from the same population of students. In Classroom A (see Figure 1A), 80 percent of the students meet the benchmark of 50 correct nonsense words per minute. In Classroom B (see Figure 1B), only 40 percent meet this benchmark.

Moreover, the rate of student growth in the two classrooms differs significantly. Based on research relating curriculum-based measurement results to performance on high-stakes Grade 3 reading tests (Good, Simmons & Kame'enui, 2001), most of the students in Classroom B are at risk for failure while most of the students in Classroom A are likely to pass these high-stakes tests.

The results in Figures 1A and 1B are not unusual. Instructional effects on the acquisition of reading skills vary this dramatically in typical classrooms across the nation. The results are highly valuable for several important decisions. First, the findings are useful to monitor the course of reading development and provide the basis for interventions early in the student's career, when such interventions are likely to be more effective. Second, the results for Classroom B suggest that the reading curriculum needs to be assessed to determine if the right content is being taught (National Reading Panel, 2000). Third, the instructional practices in Classroom B should be carefully evaluated to determine if the most effective approaches are being utilized (Snow, Burns, & Griffin, 1998). Fourth, classwide interventions are needed in Classroom B to assist students in meeting reading benchmarks and achieve a trajectory toward success in reading by the end of Grade 3. Fifth, the lowest performing students in each classroom should be identified for additional instructional opportunities through grouping within the classroom; additional instructional time on reading; or, for those farthest behind, pull-out programs such as Tier II in a RTI system.

Standards for Comparison of Performance

As previously discussed, standards for comparison are an important consideration when selecting progress measures. Usage differs for benchmarks (which are criterion referenced) and norms, based on the purpose of comparison. Many progress measures use benchmarks that have been empirically derived in order to reliably predict proficient performance on a meaningful or important outcome measure such as the state's high-stakes accountability measure. If the progress measure is being used to ensure that each student's growth keeps him or her on track for proficient performance at the end of the year, benchmarks would be a good standard to use. Norms would be useful when attempting to compare a student's performance to his or her peers. If the progress measure were being used to determine when a student receiving special education services can be reasonably reintegrated into the general classroom, or when he or she should be exited from special education services, the use of norms allows a comparison of that student's performance to that of his or her peers. This comparison is an important consideration for changing the level or intensity of service for a child (Powell-Smith & Ball, 2008). In standards for comparison of progress, ipsative standards can be used to compare the student's current progress to prior progress; however, this comparison is appropriate only if the student's prior progress was sufficient or if the comparison is to determine how much change (in rate of progress) has occurred as a result of an instructional change (M. K. Hosp, Hosp, & Howell, 2007).

Selection of Instruments for Progress Monitoring

A note of caution about the selection of instruments for progress monitoring is needed. In order to provide sufficient, technically adequate information on which to base progress decisions, an instrument must be quick to administer and score (3–5 minutes), reliable and valid for the purpose of determining rate of improvement over time, and able to be administered quite frequently (at least weekly or even more frequently). Many instruments used for this purpose

are not actually sufficient for the task. For example, informal reading inventories typically are not standardized and sometimes take longer than 5 minutes to administer and score. Instruments developed to be more diagnostic—such as the Developmental Reading Assessment—cannot be administered frequently enough and do not provide sufficiently valid information about rate of improvement over time. Other procedures that are instructional in nature, such as guided reading, are sometimes used inappropriately to monitor progress. These instruments have other purposes, but they are neither designed nor validated as instruments to collect information that can be used to make reliable progress decisions.

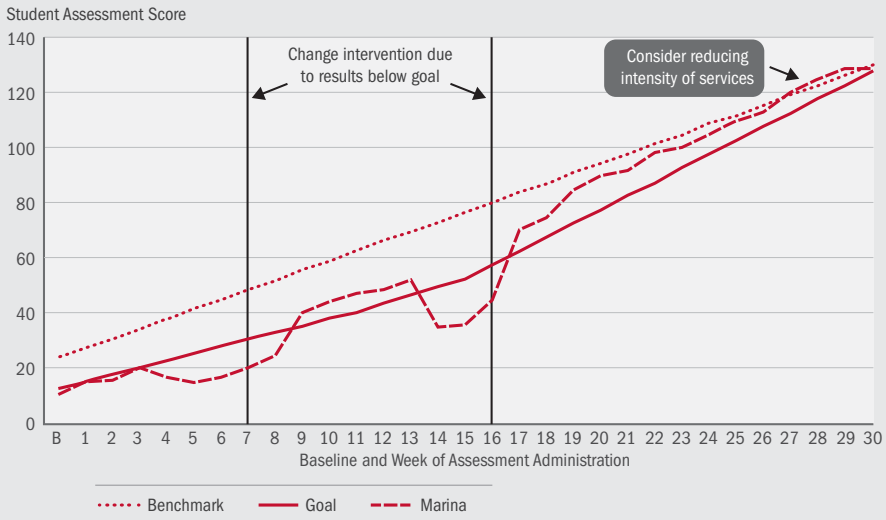
Preservice and inservice teachers need to receive the proper training to ensure selection of suitable measures for the purposes for which they are intended and are valid and to avoid the mistakes of improper usage. Selecting an inappropriate instrument to collect information to make a decision generally results in the old computer adage "garbage in, garbage out." The National Center on Response to Intervention (2010) provides a list of appropriate progress-monitoring instruments along with commentary on their strengths and weaknesses.

Example

As an example of how identification of student response aids in linking assessment and instruction, consider the case of a Grade 3 student (Marina) having difficulty with reading. Marina's teacher (Mr. Jones) uses a standardized measure of reading for both screening and progress monitoring. He has chosen published materials created in the vein of curriculum-based measurement (Deno, 1985, 2003) because of their good reliability, validity, and ability to predict mastery on the end-of-year state-mandated test. For reading, he is using a measure of oral reading fluency because it is efficient (taking only 1 minute per student per week) and consistent (in that the publisher has 30 alternate forms available so that he can use a different one each week). Through the other data he has collected, Mr. Jones knows that Marina is having great difficulty with reading, as indicated by her low scores compared to developmental benchmarks. He has identified the specific areas in which Marina needs help and has planned the instruction to provide her with the skills she is missing. Mr. Jones is measuring Marina's response to the instruction that he is providing.

Mr. Jones first identifies Marina's current level of performance and marks it on a graph. This level is indicated by the first point at the left on the goal line in Figure 2 (page 19). He also identifies the end-of-year goal for Marina and marks it at the right on the graph. He then draws a line to connect these points because that line shows the average weekly rate of progress that Marina needs to demonstrate in order to meet the end-of-year goal. Mr. Jones begins implementing the additional instruction that he is providing to Marina. Once per week, he has Marina read aloud from one of the passages and counts the number of words she reads correctly in that minute. As the weeks go on, he can see how she is responding to his instruction. After six weeks, Mr. Jones sees that Marina's reading is not progressing at the rate she needs to be successful by the end of the year. He draws an intervention line to indicate that he made an instructional change.

Figure 2. Time Series Graph of Reading Progress Monitoring With Instructional Change Decision Rules (see M. K. Hosp, Hosp, & Howell, 2007)



The assessment data do not tell Mr. Jones what to change or how to change it. Instead, he needs to use his professional judgment, expertise, and other sources of information to make that decision. Once he does, he implements that instruction and continues to monitor Marina's progress to ensure that she is on track to meet her goal.

An important point to make about Mr. Jones and Marina is the value of continuous monitoring. If Mr. Jones had not been monitoring Marina's progress weekly, at the end of the year (given her rate of progress and poor response to his instruction) he would have found that Marina was even farther behind than she was at the beginning of the year. At that point, it would have been too late for him to do anything about it. The situation would have been frustrating for Mr. Jones and demoralizing for Marina and her parents. Fortunately, the progress monitoring was successful in helping Marina reach her reading goal.

Teachers who use measures that meet the standards of reliability, validity, efficiency, and consistency have been shown to make more frequent instructional decisions (Fuchs & Fuchs, 1986) and effect greater student learning (Black & Wiliam, 1998) than teachers who do not use such data to make decisions. However, it is not just the act of collecting information that effects greater student learning. Teachers need to actively use the information to critically evaluate their instruction in order to determine how it could be changed to better meet the student's needs (Fuchs, Fuchs, Hamlett, & Stecker, 1991).

RECOMMENDATIONS

This section provides three recommendations for how to integrate the components of the Innovation Configuration for Linking Assessment and Instruction into a program of study for preservice teachers or professional development for inservice teachers.

Recommendation 1: Structure the Courses Appropriately

A series of preservice training courses or inservice activities can be structured in many ways to cover the range of topics linking assessment and instruction. The innovation configuration in the Appendix of this report is useful to identify redundancies and gaps in each of the following course structures: sequential method, infused method, and hybrid method. With explicit use of cognitive maps or scope and sequences of the interrelated nature of topic (Darling-Hammond et al., 2005) any of these methods could meet the "connected and coherent" criteria of effective teacher preparation programs (Zeichner & Gore, 1990). No one of these methods has been shown to be better than the others, so it is up to the program organizers to determine which fits best into other requirements as well as the needs of the program and students—keeping a consistent focus on the core conceptual ideas and practical skills required (Wideen, Mayer-Smith, & Moon, 1998).

Sequential Method

The sequential method of training or development for linking assessment and instruction involves separate courses or activities for different areas. For example, a preservice program of study might involve an introductory assessment course (to cover the fundamentals), a separate course on decision making (or an advanced assessment course to cover application and implementation), and then coursework that focuses on content-area instructional methods. An elementary education program may have separate methods courses for specific subjects such as reading/language arts, mathematics, science, and social studies. Secondary education programs generally will be more content specific (e.g., science, mathematics) unless the degree is for a more general focus such as special education.

One benefit of the sequential method is that the coursework can clearly build on prior courses; this sequencing provides the repeated practice that effects deeper learning and development of expertise (Gick & Holyoak, 1983). A potential disadvantage is when integration of the courses becomes more difficult due to fragmented structure or when a consistent faculty message (Gore & Zeichner, 1991) or explicit application within the content methods courses is lacking (Ericsson, Krampe, & Tesch-Romer, 1993).

Infused Method

The infused method of training or development for linking assessment and instruction involves infusing that information into the content methods courses (rather than having a stand-alone course for assessment or decision making). One benefit of this method is that the examples used and practice activities can be specifically aligned with that content area, and practice can be used to reinforce the concepts both of assessment and of instruction in order to align them.

This method also aligns well with Bruner's (1977) notion of a spiral curriculum that returns to emphasize basic ideas repeatedly and in different contexts to promote a deeper understanding of the material. A disadvantage of this approach, however, is that the core assessment information (i.e., the fundamentals) often must be repeated across the methods courses or included with a single course from which it then diverts valuable instructional time (such that that area does not get equal coverage as the others that do not include assessment fundamentals).

Hybrid Method

The hybrid method involves aspects of both the sequential and infused methods. In this method, there is a stand-alone assessment course to cover the fundamentals of assessment. This course often is used as a prerequisite for the instructional methods courses. Afterward, preservice teachers take the instructional methods courses in which the decision making and application instruction of assessment and its linking with instruction is infused. Ideally, it provides a spiral curriculum (Bruner, 1977) with repeated opportunities for practice (Gick & Holyoak, 1983), provided there is consistent structure (Zeichner & Gore, 1990) and a consistent message (Wideen et al., 1998).

Within any of these methods of course sequencing, it is imperative that teacher preparation programs incorporate the use of the practicum (supervised, practical application in the classroom) across the courses so that the preservice or inservice teachers have ample opportunity to practice the skills and apply the knowledge that they are developing in their coursework; this approach enables them to implement new practices more effectively in the classroom (Lieberman & Wood, 2003).

Recommendation 2: Use a Variety of Practice Activities

Practice is an important part of any effective training or professional development (Darling-Hammond et al., 2005). When providing preservice teacher training or inservice teacher professional development on linking assessment and instruction, it is important to include a variety of practice activities that are appropriate for the skills being covered (Ball & Cohen, 1999). In addition to other evaluative activities, practice activities can be used to determine if the preservice or inservice teachers have learned the factual information about the fundamentals of assessment. The training or professional development should be structured so that opportunities for practice and learning are ongoing (rather than the traditional one-time training), cover topics and skills in a cyclical manner (coming back to provide additional opportunities for practice and a chance to incorporate new topics with previous ones), and have ample support and mentoring so that the preservice or inservice teachers can get immediate corrective feedback (Hammerness et al., 2005). This approach will help ensure that preservice and inservice teachers have practice in applying the skills and knowledge in the same ways that they will be required to perform such activities in their classrooms. Practice should cover at least four areas: selecting the instruments, administering the instruments, scoring the instruments, and reporting and interpreting the results to parents or other professionals.

Selecting the Instruments

Practice in selecting the instruments is often aided by providing a checklist for the preservice or inservice teachers to use in order to ensure that they are considering the most relevant characteristics that they need to make accurate decisions (see Recommendation 3 on page 25). The instrument selection process also can include activities such as locating and researching different instruments that are available and accessible and that provide information aligned with instructional decisions they need to make. Offering potential scenarios to preservice and inservice teachers or allowing them to use actual scenarios that arise in their classroom, practicum, or student-teaching site provides opportunities for practice that will be relevant to the decisions they need to make. Sharing among groups or individuals also allows them to build a sort of toolbox, expanding on each others' work.

Ideally, practice selecting instruments would be heavily scaffolded with explicit transfer, starting with some case studies or scenarios in which the instructor is demonstrating and heavily guiding the application of standards or a checklist. Repeated practice could move to small group and individual practice in applying the standards to cases or scenarios and application within a practicum setting where the preservice or inservice teacher has the opportunity to discuss the process with other educators. These educators could be cooperating or mentor teachers, grade-level team members, or problem-solving team members; they should have the expertise necessary to provide expert input and guidance.

Administering the Instruments

Practice in administering the instruments may best be achieved through different levels. First, it is important for preservice and inservice teachers to administer an instrument to others in the training and to receive feedback from both the instructor and the other preservice or inservice teachers. This experience allows them not only to get the perspective of others (the instructor and their peers) but also to have a chance to watch others administer the instrument and compare their performance to the standardization rules. Use of checklists for fidelity of implementation is an easy, structured way to make sure that everyone is looking for the same characteristics while still allowing space for personal observations, such as quality of implementation and aspects that are performed particularly well. Such checklists often are available with published instruments and can be created for instruments lacking them.

Next, the preservice or inservice teachers can practice administering the instrument to a student for whom the data are not needed. (*Note:* Administering the instrument to a student who recently has taken the measure or will take it in the near future should be avoided because this administration may affect his or her results.) When practicing with a student, preservice or inservice teachers should not share the results with the student or his or her teachers or parents because these results are for training purposes only. If the preservice or inservice teacher has a teaching certification or is being specifically observed and checked by a certified teacher, some programs and districts will allow the use of those student's results. When in doubt, it is preferable to err on the side of caution and differentiate between administration for practice and administration for actual data collection and decision making. Over time, supervision and scaffolding of administration and scoring support can be gradually released

(Lampert, 2001). An additional note is that although these examples include administration of an instrument to an individual student, the same process holds true for reviews, observations, and interviews and is equally relevant for practicing such administration to groups as well as individuals.

Scoring the Instruments

For practice in scoring the instruments, it often is useful to start with simulated (or sample) results that the preservice or inservice teachers do not have to collect themselves. This approach allows the instructor to calculate reliability among the preservice or inservice teachers (which can be a useful exercise in demonstrating the importance of standardization as well as the concept of error in measurement). After the preservice or inservice teachers have administered the instruments to each other or to students (for practice), these results can be scored. Having these teachers exchange the raw results to rescore each others' work also can be used to check reliability and consistency of use with standardized scoring rubrics. As with the other areas, it is important to conduct this application in practicum or mentored settings where the preservice or inservice teacher can receive some coaching and guidance before having to work independently.

Reporting and Interpreting the Results

Practice reporting and interpreting the results is the last step, but it certainly is important as a component of practice. In their classrooms, preservice and inservice teachers will be required to share assessment results with parents and other educators. Practicing aspects of presenting the results will facilitate this process. Such aspects include describing the assessment tasks, explaining how the results are reported, explaining the standards for comparison, using graphs and charts as much as possible, and explicitly detailing how the results allow each teacher to make instructional decisions about individual students. Preservice and inservice teachers also need practice asking for feedback and interpretations of the results from other educators. These practice activities should begin with presentations to each other. Ideally, these activities should include presenting to individuals without the same training or experience (e.g., parents who are not educators), but concerns about confidentiality must be navigated. *Note:* Practice-activity results should not be presented to parents because these results are for practice purposes only and not for actual decision making about a student.

Example

As an example of putting these steps into practice, an introductory assessment course within a hybrid course structure might serve as the foundations course in which preservice teachers learn about the fundamentals of assessment as well as practice selecting, administering, and scoring different instruments. The course could be linked with a 3-hour practicum to provide access to practice opportunities. After the fundamentals of reliability, validity, types of scores, and decisions have been covered, the instructor can have the preservice teachers gather in small groups of 3–5 and critique an assessment instrument. A follow-up

activity is for each preservice teacher to do a critique individually. All the critiques are compiled and distributed to all the preservice teachers for future reference. In addition, each preservice teacher is assigned to interview his or her supervising or mentor teacher about the criteria used by that teacher to select instruments when working with students.

Preservice teachers also go through a scaffolded process of administration. First, they observe the instructor administering an instrument, observe their mentor teacher administering the same instrument, and then practice administering it to each other. After they have administered it to a peer three times, they select a student at their practicum site with whom to practice administering the instrument. This selected student must not have had a history of difficulty in school and must not have been administered the instrument within the past 12 months.

While practicing the administration, preservice teachers also have been working with simulated protocols to practice scoring. The instructor begins with the whole class scoring together and discussing how some of the decision rules are applied in a standardized fashion. Next, preservice teachers pair up and score another simulated protocol before scoring individually. After each of these activities, the instructor calculates each preservice teacher's reliability in scoring and notes areas in which mistakes are made consistently. In addition, the preservice teacher shadows the mentor teacher when the mentor teacher is scoring a protocol in order to be able to ask questions about decision rules and the link to instructional planning. After the preservice teacher has completed each of the practice administrations, these protocols also can be scored. An important point to note, however, is that the results should not be shared with anyone other than the preservice teacher's instructor and mentor teacher because these results are for practice for the preservice teacher (rather than for making decisions about the student's performance).

The last component, reporting and interpreting results, is practiced in this same scaffolded way: First, the instructor and mentor teacher demonstrate; next, preservice teachers practice with each other; and, finally, preservice teachers practice reporting to others (possibly parents or other teachers but not those of the students with whom they worked). These activities will then form the basis for the assessment and instructional planning activities in their other coursework—ensuring that they have the opportunity to practice selecting, administering, and scoring instruments in order to interpret the results and link their instructional development to them. These activities should be supervised by the course instructor as well as mentor teacher so that the preservice teacher can have sufficient chances to get feedback as well as observe how someone else might interpret the results.

Recommendation 3: Develop and Use Checklists for Selecting, Administering, and Scoring an Instrument

Structured decision-making guidelines can ease professionals through complex processes. Similarly, structured checklists to complete when selecting, administering, and scoring an instrument can be a useful tool for the preservice or inservice teacher who is not yet fully proficient at these activities.

Preliminary Questions to Consider

Before attempting to select an instrument, it also is important for the preservice or inservice teacher to ask questions such as the following:

- Why am I administering this instrument?
- Is there a more efficient way to get this information?
- Will this instrument lead to better instruction and outcomes for this student?

If the purpose for administering an instrument cannot be explicitly and emphatically stated before selecting it, other questions will need to be answered—rather than whether or not it is reliable.

Checklist for Selecting an Instrument

A checklist for scoring an instrument should cover the following general topics and principles outlined in this document:

- Fundamentals of assessment (e.g., Is the measure sufficiently reliable, valid for this purpose, and appropriate for this population?)
- Standards for comparison (e.g., Which type of standards are appropriate, and where can I find them?)
- Considerations for decision making (e.g., For what purpose do I need this instrument? Does this fit into my decision-making framework?)
- Assessment procedures (e.g., Are there other ways I could collect this information?)
- Identification of the content (e.g., Do the measurement tasks align with those expected to be taught? Does the instrument measure skill deficits or performance deficits?)
- Identification of student progress (e.g., Will this instrument provide a level of performance only, or can it also be used to index growth over time?)

Checklist for Administering or Scoring an Instrument

Providing preservice or inservice teachers with a checklist for administering or scoring an instrument is useful. Sometimes these checklists are similar to the implementation checklists for specific measures (see Good & Kaminski, 2002), and sometimes other resources are specific to an instrument (see M. K. Hosp et al. 2007). When there are not specific resources, other general checklists for setting up and preparing to administer an instrument with a student are available (see M. K. Hosp & Hosp, 2000).

CONCLUSION

Assessment and instruction are two key components of effective teaching and, therefore, are necessary components of preservice teacher training and inservice teacher professional development. These components should be intricately linked. Although there is great variation in the details of how information is collected, what it is used for, and the effect it has, research has consistently shown that teachers who base their instructional decisions on assessment data effect greater student learning (Black & Wiliam, 1998; Fuchs & Fuchs, 1986).

Not all components of this Issue Paper or the Innovation Configuration on Linking Assessment and Instruction will be equally important for all training activities, but they are important concepts and skills for all teachers and educators to have. As the field of education moves increasingly to evidence-based practice, the role of teachers as data-based decision makers also will increase. Through a detailed understanding and applied use of linking assessment and instruction, teachers will be well situated for this role.

REFERENCES

Airasian, P. W., & Madaus, G. F. (1972). Functional types of student evaluation. *Measurement and Evaluation in Guidance, 4*, 221–233.

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 3–32). San Francisco: Jossey-Bass.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Educational Assessment: Principles, Policy, and Practice, 5*, 7–74.

Bloom, B. S. (1971). Learning for mastery. In B. S. Bloom, J. T. Hastings, & G. F. Madaus (Eds.), *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.

Bruner, J. S. (1977). *The process of education: A landmark in educational theory* (Rev. ed.). Cambridge, MA: Harvard University Press.

Burns, M. K., & Mosack, J. L. (2005). Criterion-related validity of measuring sight-word acquisition with curriculum-based assessment. *Journal of Psychoeducational Assessment, 23*, 216–224.

Burns, M. K., VanDerHeyden, A. M., & Jiban, C. (2006). Assessing the instructional level for mathematics: A comparison of methods. *School Psychology Review, 35*, 401–418.

Clark, C. M., & Peterson, P. L. (1986). Teachers' thought processes. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 255–296). New York: Macmillan.

Council for Exceptional Children. (2003). *What every special educator must know: Ethics, standards, and guidelines for special education* (5th ed.). Arlington, VA: Author.

Darling-Hammond, L., Hammerness, K., Grossman, P., Rust, F., & Shulman, L. (2005). The design of teacher education programs. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 390–441). San Francisco: Jossey-Bass.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*, 219–232.

Deno, S. L. (2003). Curriculum-based measures: Development and perspectives. *Assessment for effective intervention, 28*, 3–12.

Ericsson, K. A., Krampe, R., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363–406.

Fuchs, L., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children, 53*, 199–208.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Allinder, R. M. (1991). Effects of expert system advice within curriculum-based measurement on teacher planning and student achievement in spelling. *School Psychology Review, 20*, 49–66.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Stecker, P. M. (1991). Effects of curriculum-based measurement and consultation on teacher planning and student achievement in mathematics operations. *American Educational Research Journal, 28*, 617–641.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 12*, 306–355.

Good, R. H., & Kaminski, R. (Eds.). (2002). *Dynamic indicators of basic early literacy skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.

Good, R. H., III, & Kaminski, R. A. (2011). *DIBELS Next assessment manual*. Eugene, OR: Dynamic Measurement Group.

Good, R., H., III, Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5,* 257–288

Gore, J. M., & Zeichner, K. M. (1991). Action research and reflective teaching in preservice teacher education: A case study from the United States. T*eaching and Teacher Education, 7*, 119–136.

Gravois, T. A., & Gickling, E. E. (2008). Best practices in instructional assessment. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology, V* (pp. 503–518). Bethesda, MD: National Association of School Psychologists.

Gresham, F. M. (1981). Assessment of children's social skills. *Journal of School Psychology, 19*, 120–133.

Hall, G. E., & Hord, S. M. (1987). *Changes in schools: Facilitating the process*. New York: State University of New York Press.

Hall, G. E., Loucks, S. F., Rutherford, W. L., & Newton, B. W. (1975). Levels of use of the innovation: A framework for analyzing innovation adoption. *Journal of Teacher Education, 26*, 52–56.

Hammerness, K., Darling-Hammond, L., Bransford, J., Berliner, D., Cochran-Smith, M., McDonald, M., & Zeichner, K. (2005). How teachers learn and develop. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 390–441). San Francisco: Jossey-Bass.

Haring, N. G., & Eaton, M. D. (1978). Systematic procedures: An instructional hierarchy. In N. G. Haring, T. C. Lovitt, M. D. Eaton, & C. L. Hansen (Eds.), *The fourth R: Research in the classroom*. Columbus, OH: Merrill.

Higher Education Opportunity Act of 2008, Pub. L. No. 110-315, § 1001, 122 Stat. 3083 (2008). Retrieved October 7, 2010, from http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_public_laws&docid=f:publ315.110.pdf

Hosp, J. L. (2008). Best practices in aligning academic assessment with instruction. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology, V* (pp. 363–376). Bethesda, MD: National Association of School Psychologists.

Hosp, J. L. (in press). Using assessment data to make decisions about teaching and learning. In K. Harris, T. Urdan, & S. Graham (Eds.). *The APA handbook of educational psychology*. Washington, DC: American Psychological Association.

Hosp, J. L., & Ardoin, S. P. (2008). Assessment for instructional planning. A*ssessment for Effective Intervention, 33*, 69–77.

Hosp, M. K, & Hosp, J. L. (2000). Hospitable hints. In K. Howell, S. Fox, S. Zucker, & M. Morehead, *Resources for implementing curriculum-based evaluation* (p. 12). Belmont, CA: Wadsworth.

Hosp, M. K., Hosp, J. L., & Howell, K. W. (2007). *The ABCs of CBM: A practical guide to curriculum-based measurement*. New York: Guilford Press.

Howell, K. W., Hosp, J. L., & Kurns, S. (2008). Best practices in curriculum-based evaluation. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology, V* (pp. 349–362). Bethesda, MD: National Association of School Psychologists.

Howell, K. W., & Nolet, V. (2000). *Curriculum-based evaluation: Teaching and decision making.* Belmont, CA: Wadsworth.

Idol, L. (1989). The resource/consulting teacher: An integrated model of service delivery. *Remedial and Special Education, 10*, 38–48.

Individuals with Disabilities Education Improvement Act of 2004, Pub. L. No. 108-446, § 1400, 118 Stat. 2647 (2004). Retrieved October 7, 2010, from http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=108_cong_public_laws&docid=f:publ446.108.pdf

International Reading Association & National Council of Teachers of English. (2010). *Standards for the assessment of reading and writing* (Rev. ed.). Newark, DE: International Reading Association.

Jackson, P. (1968). *Life in classrooms*. New York: Holt, Rinehart, & Winston.

Jimerson, S. R., Burns, M. K., & VanDerHeyden, A. M. (Eds.). (2007) *Handbook of response to intervention: The science and practice of assessment and intervention*. New York: Springer.

Lampert, M. (2001). *Teaching problems and the problems of teaching*. New Haven, CT: Yale University Press.

Lieberman, A., & Wood, D. (2003). *Inside the national writing project: Connecting network learning and classroom teaching*. New York: Teachers College Press.

Marzano, R. J., Brandt, R. S., Hughes, C. S., Jones, B. F., Presseisen, B. Z., Rankin, S. C., & Suhor, C. (1988). *Dimensions of thinking*. Alexandria, VA: Association for Supervision and Curriculum Development.

McKay, A. (1977). The Alberta studies of teaching: A quinquereme in search of some sailors. *CSSE News, 3*, 14–17.

Miller, M. D., Linn, R. L., & Gronlund, N. E. (2008). *Measurement and assessment in teaching* (10th ed.). Upper Saddle River, NJ: Prentice-Hall.

Morine-Dershimer, G., & Vallance, E. (1975, November). *A study of teaching and pupil perceptions of classroom interaction* (Technical Report 75-11-6). San Francisco: Beginning Teacher Evaluation Study, Far West Regional Laboratory for Educational Research and Development.

National Center on Response to Intervention. (2010). *Progress monitoring tools chart*. Washington, DC: Author. Retrieved October 7, 2010, from http://www.rti4success.org/chart/progressMonitoring/PMToolsChart_04-20-10a.pdf

National Council of Teachers of Mathematics. (1995). *Assessment standards for school mathematics*. Reston, VA: Author.

National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development. Retrieved October 7, 2010, from http://www.nichd.nih.gov/publications/nrp/upload/smallbook_pdf.pdf

No Child Left Behind Act of 2001, Pub. L. No. 107-110, § 101, 115 Stat. 1439 (2002). Retrieved October 7, 2010, from http://www.ed.gov/policy/elsec/leg/esea02/107-110.pdf

Noell, G. H., Gansle, K. A., Witt, J. C., Whitmarsh, E. L., Freeland, J. T., LaFluer, L. H., et al. (1998). Effects of contingent reward and instruction on oral reading performance at differing levels of passage difficulty. *Journal of Applied Behavior Analysis, 31*, 659–663.

O'Connor, R. E., White, A., & Swanson, H. L. (2007). Repeated reading versus continuous reading: Influences on reading fluency and comprehension. *Exceptional Children, 74*, 31–46.

Perie, M., Marion, S., & Gong, B. (2007). *A framework for considering interim assessments*. Dover, NH: National Center for the Improvement of Educational Assessment. Retrieved October 7, 2010, from http://www.nciea.org/publications/ConsideringInterimAssess_MAP07.pdf

Peterson, P. L., & Clark, C. M. (1978). Teachers' reports of their cognitive processes during teaching. *American Educational Research Journal, 15*, 555–565.

Popham, W. J. (2010). *Classroom assessment: What teachers need to know* (6th ed.). Upper Saddle River, NJ: Prentice-Hall.

Powell-Smith, K. A., & Ball, P. L. (2008). Best practices in reintegration and special education exit decisions. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology, V* (pp. 263–280). Bethesda, MD: National Association of School Psychologists.

Race to the Top Fund, 74 Fed. Reg. 59,836 (Nov. 18, 2009). Retrieved October 7, 2010, from http://www.federalregister.gov/articles/2009/11/18/E9-27427/overview-information-race-to-the-top-fund-notice-inviting-applications-for-new-awards-for-fiscal

Reschly, D. J., & Wood-Garnett, S. (2009). T*eacher preparation for response to intervention in middle and high schools* (TQ Research & Policy Brief). Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved October 7, 2010, from http://www.tqsource.org/publications/September2009Brief.pdf

Roy, P., & Hord, S. M. (2004). Innovation configurations chart a measured course toward change. *Journal of Staff Development, 25*, 54–58.

Salvia, J., Ysseldyke, J., & Bolt, S. (2010). *Assessment: In special and inclusive education* (11th ed.). Belmont, CA: Wadsworth.

Shapiro, E. S., & Kratochwill, T. R. (Eds.). (2000). *Behavioral assessment in schools: Theory, research, and clinical foundations* (2nd ed.). New York: Guilford.

Shavelson, R. J., & Stern, P. (1981). Research on teachers' pedagogical thoughts, judgments, decisions, and behavior. *Review of Educational Research, 51*, 455–498.

Shepard, L. A., Hammerness, K., Darling-Hammond, L., & Rust, F. (with Snowden, J. B., Gordon, E., Gutierrez, C., & Pacheo, A.). (2006). Assessment. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 275–326). San Francisco: Jossey-Bass.

Shinn, M. R. (2008). Best practices in using curriculum-based measurement in a problem-solving model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology, V* (pp. 243–262). Bethesda, MD: National Association of School Psychologists.

Stecker, P. M., & Fuchs, L. s. (2000). Effecting superior achievement using curriculum-based measurement: The importance of individual progress monitoring. *Learning Disability Research and Practice, 15*, 128–134.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Torgesen, J. K., & Miller, D. H. (2009). *Assessments to guide adolescent literacy instruction*. Portsmouth, NH: Center on Instruction.

VanDerHeyden, A. M., & Witt, J. C. (2008). Best practices in can't do/won't do assessment. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology, V* (pp. 131–139). Bethesda, MD: National Association of School Psychologists.

Wesson, C. L. (1991). Curriculum-based measurement and two models of follow-up consultation. *Exceptional Children, 57*, 246–257.

Wideen, M., Mayer-Smith, J., & Moon, B. (1998). A critical analysis of the research on learning to teach: Making the case for an ecological perspective no inquiry. *Review of Educational Research, 68*, 130–178.

Zeichner, K. M., & Gore, J. (1990). Teacher socialization. In W. R. Houston, M. Haberman, J. P. Sikula, T. J. Buttery, E. Guyton, & Association of Teacher Educators (Eds.), *Handbook of research on teacher education* (2nd ed., pp. 525–547). New York: Macmillan.

APPENDIX. Innovation Configuration for Linking Assessment and Instruction in Teacher Preparation and Professional Development: Essential Components for Teacher Competency

Essential Components	Degree of Implementation					
Instructions: *Place an X under the appropriate variation implementation score for each course syllabus that meets each criterion specified from 0 to 4. Score and rate each item separately.* ***Descriptors and/or examples are bulleted below each of the components.***	**Code = 0**	**Code = 1**	**Code = 2**	**Code = 3**	**Code = 4**	**Rating**
	No evidence that the concept is included in the class syllabus.	Syllabus **mentions** content related to the concept.	Syllabus mentions the concept, **with required readings** on the topic.	Syllabus mentions the concept; requires readings; and has **an assignment, project, or test** on the topic.	Syllabus mentions the concept; requires readings; has an assignment, project, or test; **and has supervised practice** related to the topic through fieldwork.	*Rate each item as the number of the highest variation receiving an X under it.*
Fundamentals of Assessment Reliability—definition and types commonly used to judge educational assessments (e.g., test-retest, interrater) • Validity—definition and types commonly used to judge educational assessments (e.g., criterion-related, content) • Use and interpretation of score scales (e.g., percentiles, standard scores, systematic observation metrics) • Legal provisions of assessment (e.g., ESEA, IDEA) • Issues of cultural and linguistic bias and fairness • Accommodations and modifications for students with disabilities or English learners • Types of educational decisions for which assessment data can be collected to help in decision making (i.e., screening, progress, broad and targeted diagnostic, and outcome)						

Essential Components	Degree of Implementation					
Instructions: *Place an X under the appropriate variation implementation score for each course syllabus that meets each criterion specified from 0 to 4. Score and rate each item separately.* ***Descriptors and/or examples are bulleted below each of the components.***	**Code = 0**	**Code = 1**	**Code = 2**	**Code = 3**	**Code = 4**	**Rating**
	No evidence that the concept is included in the class syllabus.	Syllabus **mentions** content related to the concept.	Syllabus mentions the concept, **with required readings** on the topic.	Syllabus mentions the concept; requires readings; and has **an assignment, project, or test** on the topic.	Syllabus mentions the concept; requires readings; has an assignment, project, or test; **and has supervised practice** related to the topic through fieldwork.	*Rate each item as the number of the highest variation receiving an X under it.*
Standards for Comparison of Performance • Norm referenced (i.e., comparison to age- or grade-similar peers) • Criterion referenced (i.e., comparison to empirically derived level of proficiency) • Ispative standards (i.e., comparison to prior performance)						
Considerations for Decision Making • Usage of and differentiation between the terms *assessment* and *evaluation* • Comparisons of inside and outside purposes of decision making • Definitions and comparisons between summative and formative evaluation (and possibly interim assessments) • Use of structured decision-making frameworks (e.g., curriculum-based evaluation)						

Essential Components	Degree of Implementation					
	Code = 0	**Code = 1**	**Code = 2**	**Code = 3**	**Code = 4**	**Rating**
Instructions: *Place an X under the appropriate variation implementation score for each course syllabus that meets each criterion specified from 0 to 4. Score and rate each item separately.* ***Descriptors and/or examples are bulleted below each of the components.***	**No evidence** that the concept is included in the class syllabus.	Syllabus **mentions** content related to the concept.	Syllabus mentions the concept, **with required readings** on the topic.	Syllabus mentions the concept; requires readings; and has **an assignment, project, or test** on the topic.	Syllabus mentions the concept; requires readings; has an assignment, project, or test; **and has supervised practice** related to the topic through fieldwork.	*Rate each item as the number of the highest variation receiving an X under it.*
Assessment Procedures • Review of prior records • Interview with relevant individuals • Observation of performance in appropriate settings • Administration and interpretation of test results • Selection of assessment procedures that provide the information needed to make instructional decisions						
Identification of Content to Teach • Consideration of and focus on broad areas (e.g., reading, mathematics) • Consideration and focus on specific subskills, important prerequisites, or related skills (e.g., phonological segmenting, understanding of place value) • Addressing alignment of assessment or instruction on different forms of knowledge (i.e., facts, concepts, strategies) • Working within the instructional hierarchy: accuracy, fluency, generalization, adaptation • Consideration of difficulties arising from skill deficits or performance deficits • Alignment with or writing of goals and objectives • Analysis of student work						

Essential Components	Degree of Implementation					
Instructions: *Place an X under the appropriate variation implementation score for each course syllabus that meets each criterion specified from 0 to 4. Score and rate each item separately.* ***Descriptors and/or examples are bulleted below each of the components.***	**Code = 0**	**Code = 1**	**Code = 2**	**Code = 3**	**Code = 4**	**Rating**
	No evidence that the concept is included in the class syllabus.	Syllabus **mentions** content related to the concept.	Syllabus mentions the concept, **with required readings** on the topic.	Syllabus mentions the concept; requires readings; and has **an assignment, project, or test** on the topic.	Syllabus mentions the concept; requires readings; has an assignment, project, or test; **and has supervised practice** related to the topic through fieldwork.	*Rate each item as the number of the highest variation receiving an X under it.*
Identification of Student Response • Characteristics of good formative measures • Ease of administration and scoring • Representation of performance with graphs • Aggregation of data to make individual or small-group, classwide, and schoolwide or districtwide decisions • Standards for comparison of performance • Norms • Benchmarks • Prior progress						
Column Totals						

ABOUT THE NATIONAL COMPREHENSIVE CENTER FOR TEACHER QUALITY

The National Comprehensive Center for Teacher Quality (TQ Center) was created to serve as the national resource to which the regional comprehensive centers, states, and other education stakeholders turn for strengthening the quality of teaching—especially in high-poverty, low-performing, and hard-to-staff schools—and for finding guidance in addressing specific needs, thereby ensuring that highly qualified teachers are serving students with special needs.

The TQ Center is funded by the U.S. Department of Education and is a collaborative effort of ETS, Learning Point Associates, and Vanderbilt University. Integral to the TQ Center's charge is the provision of timely and relevant resources to build the capacity of regional comprehensive centers and states to effectively implement state policy and practice by ensuring that all teachers meet the federal teacher requirements of the current provisions of the Elementary and Secondary Education Act (ESEA), as reauthorized by the No Child Left Behind Act.

The TQ Center is part of the U.S. Department of Education's Comprehensive Centers program, which includes 16 regional comprehensive centers that provide technical assistance to states within a specified boundary and five content centers that provide expert assistance to benefit states and districts nationwide on key issues related to current provisions of ESEA.

NATIONAL COMPREHENSIVE CENTER FOR TEACHER QUALITY

1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
877.322.8700 | 202.223.6690

www.tqsource.org

This work was originally produced in whole or in part by the National Comprehensive Center for Teacher Quality with funds from the U.S. Department of Education under cooperative agreement number S283B050051. The content does not necessarily reflect the position or policy of the Department of Education, nor does mention or visual representation of trade names, commercial products, or organizations imply endorsement by the federal government.

The National Comprehensive Center for Teacher Quality is a collaborative effort of ETS, Learning Point Associates, and Vanderbilt University.

4953_12/10

LEARNING POINT Associates®
An Affiliate of American Institutes for Research®