What Is It?

The over-arching purpose of the Alert series is to help practitioners and parents make informed decisions about the potential effectiveness of specific instructional interventions. However, even after an intervention has been selected and implemented, decisions must be made about whether the intervention is working for a particular student in a particular setting, or whether adaptations to the intervention must be made. To that end, formative evaluation procedures can be used. Formative evaluation is the ongoing collection and use of information to evaluate the effectiveness of instructional implementations and to determine whether instructional adaptations are necessary (see reference 5). Formative evaluation can be contrasted with summative evaluation: Whereas in summative evaluation, information is gathered to judge student outcomes, in formative evaluation, it is gathered to evaluate and modify instruction.

For Whom Is It Intended?

Formative evaluation procedures are intended for use with students of all ages and in a wide range of content areas and curricula. The procedures are useful in evaluating the effectiveness of curricular innovations in a broad range of content and skill areas.

How Does It Work?

Fuchs and Deno (see reference 4) describe two general approaches to formative evaluation, each of which provides different types of information:

- Specific subskill mastery measurement (or mastery measurement) is a task-analytic approach in which a competency is broken down into subskills. These subskills usually are arranged in a hierarchical order, and student mastery of each subskill is assessed. For example, in the area of reading, decoding might be broken down into subskills that include segmenting and blending sounds, matching letters with sounds, sounding out words, sight word reading, etc. Student mastery of the first subskill is assessed until the student reaches a pre-selected criterion (e.g., 80% accuracy). Reaching that criterion signals the teacher to move on to the next subskill in the hierarchy. In IEP terms, mastery measurement focuses on short-term objectives.

- **General outcome measurement** focuses on global outcomes or desired terminal behaviors. Student progress in a general outcome measurement approach is assessed by repeatedly sampling performance on probes that represent the global outcome or desired terminal behavior. In our reading example, the global outcome is improvement in general reading proficiency. Progress toward that goal might be assessed by having students read aloud from text. In IEP terms, general outcome measurement focuses on the long-range goal.

The two general approaches to formative evaluation answer different questions (see reference 4). **Mastery measurement** answers the question, "Has the student learned the skill I have just taught?" **General outcome measurement** answers the question, "Has learning this skill in this manner led to growth and improvement in the general academic area?" In our example, mastery measurement is used to determine whether students can segment and blend sounds, whereas general outcome measurement is used to determine whether learning to segment and blend sounds leads to better reading performance.

There are many specific approaches to formative evaluation, each representing mastery or general outcome measurement to a different extent. Four prominent approaches are discussed below.

- **Curriculum-based assessment** (CBA) is the observation and recording of student performance in a local curriculum in order to gather information to make instructional decisions (see reference 12). CBA is the clearest example of a mastery measurement approach to assessment. The test materials used in CBA are developed by the teacher on the basis of a task analysis of the curriculum. Although procedures vary across CBA systems, students usually are pretested before instruction to determine which sub skills have not yet been mastered. These subskills then form the core of the curriculum. As instruction occurs, students are repeatedly measured on the selected subskills using alternative test forms Mastery of a subskill signals a move to the next skill in the hierarchy (see references 8, 10 & 12).

- **Curriculum-Based Measurement** (CBM) is a progress-monitoring system in which student performance is measured repeatedly (e.g., once or twice per week)

with test materials that represent an entire curricular domain rather than sub-components of the domain (see reference 2). CBM is the clearest example of a general outcome measurement approach. Student progress in CBM is assessed in a continuous way throughout an instructional program or academic year using measures that are valid and reliable indicators of student performance. Teachers examine the rate at which students are improving on these indicators to determine the effectiveness of their instruction. If students are progressing, instruction continues. If students are not progressing, instruction is modified.

- **Portfolio and performance assessment** both rely on identification by the teacher of broad-based "authentic" tasks deemed necessary for students to succeed in the "real" world. **Portfolio assessment** is the collection of student work demonstrating what a student has done and, by inference, what a student can do. In portfolio assessment, performance is evaluated on the basis of an ongoing collection of student works that are judged by the teacher to be important indicators of the outcomes of learning activities (see reference 7). **Performance assessment** emphasizes the use of a direct measure of student performance in real or simulated situations rather than the indirect measure usually obtained by traditional paper-pencil tests (see references 1, 3, & 14). The frequency with which portfolio and performance assessments are collected, and the manner in which they are used to inform instruction is determined by the teacher. Portfolio and performance assessments include components of both mastery measurement and general outcome measurement. Both involve breaking the curricular domain into subskills, but because these subskills represent tasks necessary for the student to succeed in the "real" world, they often rep resent year-end goals rather than short-term objectives.

How Adequate Is The Research Knowledge Base?

In a review of the research on formative evaluation, Fuchs and Fuchs identified components of formative evaluation that contribute to its effectiveness in promoting student achievement. Two of these components are rules for data use and graphing (see reference 5).

- **Data use** is the analysis of students' data at regular intervals to determine the effectiveness of instruction and to determine whether instructional changes are necessary. When teachers use specified rules to analyze and interpret formative evaluation data, as opposed to using teacher judgment alone, student achievement

gains are greater. An example of a rule that might be used to guide the practitioner in responding to the data would be: When a student's performance falls below the goal line on 3 consecutive days, change the instruction.

- **Graphing** the data, as opposed to merely recording the data, also leads to greater student achievement gains. Graphs seem to facilitate more accurate and frequent analysis of the data by teachers and to provide more useful feedback to students.

The extent to which each of the four formative evaluation procedures includes data-use rules and graphing procedures varies, as does the empirical support for the effectiveness of each approach.

- In a CBA approach, the data-use rule generally is to continue instructing on a selected subtask until the student has mastered the skill. Once the skill is mastered, the student moves on to a new skill and the measurement material changes to reflect that new skill. Student performance over time usually is graphed in CBA. Since performance is repeatedly tested within each subskill until a level of proficiency is reached, the graph represents progress within a subskill. When the next subskill is selected, a new curriculum-based assessment is begun, and a new graph is developed. Few studies have been conducted to examine the technical characteristics of CBA as a progress-monitoring procedure (see reference 11).

- In a CBM approach, data-use rules are provided so that instructional decisions can be made by comparing student progress at time 1 to that at time 2, or by comparing progress for an individual student to a long-range goal set for that student. Student progress over time is displayed graphically, and represents student performance on the global outcome measure over time. A large research base supports both the technical adequacy of the measures used in CBM, and the positive effects on student achievement associated with the use of CBM (see references 2 & 6).

- Neither portfolio nor performance assessments include specific data-use rules. Teachers are left on their own to determine ways to use the data to inform their instruction. Portfolio and performance assessments also do not systematically make use of graphing procedures. With respect to the research base supporting portfolio and performance assessments, concerns have been raised about the technical adequacy of the measures developed for both systems (see references 3 & 14), and little research has been conducted to examine the effects of the use of either system on student achievement.

Comparison of Formative Evaluation Systems

	Type of Approach	Timing	Measures	Data Utilization Rule	Graphing
CBA	Mastery Measurement	Throughout instruction to check mastery of objectives	Teacher made; different for each subskill	Decision making based on proficiency level	Within subskills
CBM	General Outcome Measurement	Continuous	Selected from research-validated reliable measures	Decision making based on rate of student progress	On equivalent forms representing the general outcome measures
Portfolio and Performance Assessment	Mastery General Outcome Measurement	Ongoing collection of products representing curriculum goals	Teacher-or student-selected assignments	None	None

How Practical Is It?

Each type of formative evaluation procedure requires extra time and effort on the part of the teacher. All formative evaluation procedures, by definition, require ongoing data collection, and all require some development of measurement materials and procedures.

CBA, performance, and portfolio assessments all require a considerable amount of on-site resource development. Performance and portfolio assessments require the identification of desired long-range goals and tasks that reflect performance on those goals. CBA requires the development of a hierarchy of skills and measures to assess those skills. In addition to curriculum development, portfolio and performance assessments require the development of data use rules and graphing procedures if these components are to be implemented.

Although CBM does not require curriculum development or the development of data-use rules or graphing procedures, it does require teacher development of alternative assessment probes that are representative of the desired general outcome. In addition, teachers must graph student progress and use the graph to make decisions regarding the effectiveness of instruction.

How Effective Is It?

Of the four formative evaluation procedures described, only CBM and CBA include the two components of effective formative evaluation—data-use rules and graphing. In terms of validity and reliability of the measures, and effects on student achievement, CBM has the strongest empirical data base, although research on CBM has been conducted primarily at the elementary school level. Even though CBM provides valid and reliable information

regarding student progress, it does not provide information regarding how to change instruction when students are not progressing. CBA, portfolio, and performance assessments, on the other hand, do provide such instructional information, but less is known about their reliability and validity for measuring student progress. Several authors have suggested ways to combine CBM with other formative evaluation procedures to create a valid and reliable measurement system (CBM) that is informed by a rich instructional data source (CBA, portfolio or performance assessments) (see references 3, 4, 10, & 13).

What Questions Remain?

Questions remain about the impact of CBA, portfolio, and performance assessments on student achievement, and about the validity and reliability of the specific measures developed by individual teachers in implementing these systems. With respect to CBM, questions remain regarding its use with students in early education and secondary-level education, and in areas other than reading, written expression, spelling, and mathematics. Only recently have extensions of CBM to other age levels and to other content areas been made (see reference 9).

How Do I Learn More?

For more information on the different approaches to formative evaluation see:

- Jones, C.J. (1998). Curriculum-based assessment: The easy way. Springfield, IL: Charles C Thomas.

- Lewin, L, & Shoemaker, B. J. (1998). Great performances: Creating classroom-based assessment tasks. Alexandria, VA: Association for Supervision and Curriculum Development.

- Shinn, M.R. (Ed.) (1998). Advanced applications of curriculum- based measurement. NY: The Guilford Press.

- Shinn, M.R. (Ed.) (1989). Curriculum-based measurement: Assessing special children. NY: The Guilford Press.

- Wesson, C. L. & King, R. P. (1996). Portfolio assessment and special education students. Exceptional Children, 28, 44-48.

- Valencia, S.W. (Ed.) (1994). Authentic reading assessment: Practices and possibilities. Newark, DL: International Reading Association.

GO FOR IT

*A*bout the Authors

This Alert issue was written by **Dr. Christine Espin, Dr. Jongho Shin,** and **Todd Busch,** in collaboration with the DLD/DR Current Practice Alerts Editorial Committee. Christine Espin is an Associate Professor, Jongho Shin a Postdoctoral Research Associate, and Todd Busch a doctoral student in the Department of Educational Psychology at the University of Minnesota. Dr. Espin coordinates the Learning Disabilities Teacher Certification Program in Special Education at the University. The authors wish to thank Addison Stone, Naomi Zigmond, Stanley Deno, and the DLD/DR board for their helpful comments on this issue, and wish to acknowledge the Netherlands Institute for Advanced Study in the Humanities and Social Sciences for its support of this project.

*A*bout the Alert Series

The Current Practice Alert Series is a joint publication of Division for Learning Disabilities and Division for Research within the Council for Exceptional Children. The series is intended to provide an authoritative resource concerning the effectiveness of current practices intended for individuals with specific learning disabilities. Each Alert issue will focus on a single practice or family of practices that is widely used or discussed in the LD field. The Alert will describe the target practice and provide a critical overview of the existing data regarding its effectiveness for individuals with learning disabilities. Practices judged by the Alert Editorial Committee to be well-validated and reliably implementable are featured under the rubric of Go For It. Those practices judged to have insufficient evidence of effectiveness are featured as Exercise Caution. For more information about the Alert series and a cumulative list of past Alert topics, visit the Alerts page on the CEC/DLD website: **http://www.cec.sped.org/dv-menu.htm**

Target practices for future issues: Accommodations for High-Stakes Assessments, Mnemonic Instruction, Class-wide Peer Tutoring, Co-teaching.