# Table of Contents

Executive Summary.................................................................................................................   i

Introduction and Context........................................................................................................  1

1. Purposes of Assessment....................................................................................................  1

2. Connections Among Curriculum, Instruction, and Assessment.........................................  2

3. Major Assessment Components.........................................................................................  4
   A. Summative Assessments..............................................................................................  5
   B. Interim Assessments...................................................................................................  7
   C. Formative Assessments............................................................................................... 10
   D. Classroom-based Assessments.................................................................................... 14

4. Validity of Assessments and Evaluation the Assessment System....................................... 16

5. Statewide Data Information System................................................................................... 19

6. Professional Development.................................................................................................. 20

7. Summary and Discussion..................................................................................................... 21

References................................................................................................................................ 24
Executive Summary

The state of Washington is undertaking the development of a comprehensive statewide assessment system as called for in legislative act RCW 28A.300.041. The system is designed to provide information for accountability purposes and information that supports and enhances students’ learning. It will connect and integrate several assessment components and focus all efforts on well-defined learning outcomes and goals. The major issues and components of such a system are briefly summarized below.

Purposes of Assessment

The comprehensive assessment system is embedded in and is an integral part of a broader educational effort, the purpose of which is to enhance students’ learning. Assessment is used to help achieve a goal and the goal is to enable all students to be at least proficient on the state content standards. The summative and interim assessment components will focus on assessment of learning; the formative and to some extent the interim assessment components will focus on assessment for learning. The system is designed to be comprehensive and inclusive, allowing for the valid assessment of all students.

Connections Among Curriculum, Instruction, and Assessment

Educational systems are comprised of curriculum, instruction, and assessment. Curriculum refers to content standards and describes intended learning outcomes, the knowledge and skills we expect all students to develop. Instruction refers to the vast array of activities teachers use and experiences they arrange for students through which students can make progress toward becoming proficient on the content standards. Assessment takes many forms that measure and describe students’ status, growth, and accomplishments with respect to the content standards.

An effective educational system connects all three major components by aligning instruction to curriculum standards and by using assessments to determine students’ status and progress on the standards. Clearly described and appropriately developed content standards are the focus of instruction and assessment and considerable care and effort should be taken in developing, organizing, and articulating these standards.

Components of a Comprehensive Assessment System

A comprehensive assessment system includes summative, formative, interim, and classroom assessments, all of which must be aligned with and embody the intended learning outcomes. Classroom assessments include summative, interim, and formative assessments.

Summative assessments are given at the end of an instructional sequence to determine students’ status relative to content and performance standards, serve an important accountability purpose, and concerted efforts should be made to extract instructional relevant information from these assessments.

Interim assessments are given periodically at the end relative small units of instruction, provide a “snap shot” of students’ progress, and serve instruction, evaluation, and prediction purposes.
Formative assessments are given during the process of instruction to help students and teachers understand students’ current learning relative to the content standards. They are used to identify students’ strengths as well as their learning deficiencies or gaps, and should provide guidance for responsive instructional actions. Claims that an assessment or assessment system is formative should be reviewed critically and evaluated against evidence that the assessment or system has the characteristics that research shows are found in effective formative assessments.

Classroom assessments include a wide range of activities and procedures that teachers use to obtain information and provide feedback about students’ progress on important learning outcomes that are not easily assessed with fixed format externally developed assessments. Teachers routinely use classroom assessments that are summative, interim, and formative assessments. A serious weakness of many classroom assessments is that they are not properly aligned with the content standards they are used to assess.

Validity of Assessments and Evaluating the Assessment Systems

The quality of an assessment system is determined by the validity of the system’s assessment components. Validity analyses should focus on assessment content, student response processes, internal structures of the assessments, assessments’ relationship to other measures, and the consequences of testing. Specific plans for validity studies are an integral part of the assessment development process and reports of validity studies should be included in all technical reports documenting the development of the assessments.

Evidence of validity for all students must be collected and presented as part of the validity portfolio. The assessment system and all its components must be inclusive and designed to eliminate any barriers to students’ access. An inclusive assessment system must serve students with a wide range of achievement levels, students with diverse cultural and linguistic backgrounds, and students with a wide range of special needs.

A plan for evaluating the assessment system as a whole must be developed and implemented to provide a basis for reviewing the value and impact of the entire endeavor.

Formats and Platforms for Delivery Assessments

Assessment systems should migrate toward the use of computer-based and web-oriented administration and scoring wherever and whenever appropriate. Interim assessments and summative assessments lend themselves to computer administration since they are designed to measure a particular set of learning outcomes. Certain approaches to formative assessment can also be delivered and supported with a computer-based system. There are a number of advantages to computerized adaptive testing but this approach to assessment requires considerable item development efforts. The value of a computerized adaptive assessment system is limited by the depth, quantity, and quality of the aligned items available for use by the system.

Statewide Data Information System

The comprehensive assessment system involves coordinating information of many different types from many different sources. Various types of information about students, their
backgrounds, and educational experiences must be collected easily and stored efficiently and securely. The use of such information must be authorized and monitored carefully but the educational value requires that information be accessed and presented in a number of forms to serve the needs of teachers, other educators, and other appropriate users. The information system is a critical infrastructure feature of a comprehensive assessment system and its final development and full implementation should parallel the development of other components of the assessment system.

**Professional Development**

Developing and implementing a comprehensive assessment system that provides accountability information and enhances students’ learning will require substantial professional development of preservice and inservice teachers and school leaders if the system is to be effective. Ultimately, what teachers do in their classrooms is what creates student learning and an assessment system will only be as effective as the ability teachers have to use it. In support of enhancing students’ learning, professional development should emphasize assessment for instruction, assessment as an instructional tool, and recognize the need to help teachers develop instructional strategies for responding to learning deficiencies and gaps revealed by the assessments. All professional development activities should stress the need to develop and use assessments that are inclusive and valid for all students.
Envisioning a State Educational System Improving Learning Through a Comprehensive Assessment System

Introduction

Helping all students become at least proficient on state content standards has become the focused mission of schools, schools districts, and state departments of education. Educational agencies and institutions are involved in a wide range of demanding, complex, and engaging activities in pursuit of this goal. It is important to recognize, however, that all activities designed to support education in various ways are means to an end and the effectiveness of various educational endeavors can be evaluated by examining the answer to the question, “Has this educational program or policy led to increased student learning?”

The purpose of this paper is to describe the key elements of an effective comprehensive assessment system. The major sections of this paper are organized around these elements:

1. Purposes of assessment;
2. Connections among curriculum, instruction, and assessment;
3. Major Assessment Components: summative, interim and formative assessment;
4. Validity of assessments and evaluation of the assessment system;
5. Statewide Data Information System;
6. Professional development; and
7. Summary and discussion.

This paper is written at the request of the Washington Office of Superintendent of Public Instruction (OSPI) to elaborate on the legislative mandates described in Senate Bill 5414 also described in RCW 28A.300.041. The legislation calls for the redesign of the statewide student assessment system and the development of a comprehensive and integrated statewide assessment system. The paper also serves as one of the resource documents for the Formative Assessment Symposium sponsored by the OSPI for Washington educational leaders to be held February 1 and 2, 2010.

1. Purposes of Assessment

Statewide assessment systems are designed to serve two major purposes. These are instructional/educational support and educational accountability. Current research and discussions of assessment approaches have lead to a distinction between these two purposes as assessment for learning and assessment of learning (Brookhart, 2009; Stiggins, 2005). Assessment for learning includes assessment activities which aid teachers and others in improving instructional practices and thus enhance students’ learning. Formative assessment is the most common form of assessment for learning. Black and Wiliam (1998a, 1998b) cite strong evidence that properly planned and carefully implemented formative assessment increases student learning.

Assessments of learning include district and statewide end-of-year and end-of-course summative assessments. State assessment programs have relied heavily on these assessments to
provide information for educational accountability purposes and they have functioned reasonably well. Summative assessments, however, have not been helpful in providing instructionally useful information for supporting the instructional efforts of teachers and other educational leaders for several reasons (Perie et al., 2010). Most critically, end-of-year summative assessments are not designed to provide fine-grained feedback focused on teachers’ instruction and students’ learning during the educational process. On the contrary, by design, statewide assessments are given after instruction has been completed. Furthermore, in most states, the programs assess a sample of content objectives from a wide range of objectives but not to the depth and level of detail that would be most useful to teachers and students. The move toward end-of-course examinations can improve this situation somewhat because they assess a range of content objectives that is more narrowly focused. This allows for some additional depth in assessment and a closer connection to classroom instruction.

The distinction between assessment for learning and assessment of learning may be clear conceptually but in practice many assessments serve both purposes, for different users, and at different times. Data from formative assessments for learning can be used for that purpose at the classroom level but may also be aggregated at the school, district, and state levels to provide evaluative information of learning. Conversely, statewide end-of-year assessments of learning are used for that purpose but may also provide useful information for revising curricular emphases and instructional approaches. A comprehensive assessment system should have a combination of assessments both of and for learning and remain flexible in their use to take full advantage of the information they provide.

A comprehensive and inclusive statewide student assessment system must be designed and implemented to serve all students. This includes students from a wide range of achievement levels, students from diverse cultural and linguistic backgrounds, and students with a variety of special needs that may require modifications or accommodations in content standards, instruction, and assessment. All steps in planning, developing, and implementing the statewide assessment systems must reflect the mission to serve all students.

2. Connections among Curriculum, Instruction, and Assessment

A comprehensive assessment system must be planned, developed, and implemented as part of a broader educational system comprised of three major components. These are curriculum, instruction, and assessment (Popham, 2004; Pellegrino, 2006). They can be briefly defined as follows for the current discussion.

- **Curriculum** refers to the goals and objectives of education and describes what we want students to learn as a result of their school experiences.
- **Instruction** is used in the broad sense to include the wide range of teacher activities, strategies, teaching/learning materials, and experiences teachers provide and arrange for students to facilitate their learning progress toward the curricular goals.
- **Assessment** is used in the broad sense to include the full array of procedures and strategies teachers and others use to collect information about students and to make inferences about their progress on the curriculum goals and objectives. Assessments
take many forms, serve many purposes, and can be developed, applied, and usefully interpreted by teachers and students.

**Describing Curriculum Goals and Objectives for Instruction and Assessment**

The statements of curriculum goals and objectives provide the focus for coordinating and connecting with instruction and assessment. Mandated by law, these expectations are contained in state content standards and school district curricula and define and describe educational intentions. The content standards or curricular expectations should be well developed, clearly articulated, and fully understood by teachers and students. In *Knowing What Students Know*, Pellegrino et al. (2001) argue that learning expectations should be described in terms of learning progressions (Masters and Forster, 2001) that are embedded in and reflect a conceptually sound model of students’ cognitive development and processing. Content objective statements can represent comprehensive learning but should also include enough specifics to provide guidance for teachers and students about what students are expected to know and do.

Popham (2009) warns that the current approach used in most states to describe curriculum standards suffers from three major problems. First, there are far too many standards; second, the standards are too narrow and specific; and third, the relative importance or priority of the standards is not described. In the chapter “Too many curricular targets,” Popham advocates an approach that would employ a relatively small number of curriculum standards, described at granular level that is useful to teachers but still provides guidance for assessment, and which reflect carefully chosen curricular priorities. The development or revisions of Washington curricular standards should carefully consider the guidance in this chapter in future activities.

**Aligning Curriculum, Assessment, and Instruction**

Although it is convenient to describe curriculum, instruction and assessment (C-I-A) as separate components, they need to be seamlessly connected and integrated for an educational system to be effective (Popham, 2004). In practice, curriculum, assessment, and instruction are highly inter-related and in some cases are difficult to distinguish in effective classrooms.

**Alignment** is the term used to describe the appropriate inter-relationships among curriculum, instruction and assessment and proper alignment is essential for an educational system to be effective in enhancing students’ learning. Some of the dynamics of C-I-A alignment and related issues include the following features.

- **Curriculum goals and content objectives provide the learning targets for teachers, students, and parents/community.** They help teachers develop or select appropriate instructional activities and materials and enable students to understand what is expected of them. Teachers and students must be clear about what is expected of them so they can focus teaching and learning efforts on the same targets. Research shows that the students’ learning is greatly facilitated when they and their teachers clearly understand curricular goals and objectives (Black and William, 1998a).

- **Examples of appropriate instructional strategies and materials should be provided in the curriculum document** to clarify the meaning of the content standards. Such examples provide explicit connections between curriculum and instruction in a form that helps
teachers translate the content standards into instructional approaches and also provide models of appropriate instruction.

- **Examples of appropriate assessment strategies should be provided in the curriculum document** to connect the assessments to the curriculum goals and objectives. The examples also give teachers and students concrete illustration of what they are expected to know and be able to do in terms of how they will be assessed. Teachers can use such examples in developing their own classroom assessment, which will be better aligned to the content standards than they might be without the examples.

- **Frameworks and blueprints for developing coherent assessments should be aligned with the targeted curriculum objectives.** In other words, all assessments in a comprehensive system should provide information about students’ status with respect to the content standards and objectives.

- **Explaining the criteria used to evaluate students’ work is an important part of the assessment-curriculum connection.** Shepard (2006) points out that for students to fully understand what is expected of them the descriptions of learning targets should include an elaboration of the criteria that will be used to evaluate their work. A familiar example of this approach involves making sure students understand the rubric used in scoring their writing.

### 3. Major Assessment Components

Perie, Marion, and Gong (2010) provide a very useful model for classifying various components of a comprehensive assessment system that includes summative, interim, and formative assessments. The distinctions among summative, interim, and formative assessments are based primarily on when the assessments are given relative to the instructional sequence and how the results of the assessments are used. These three types of assessments cannot be distinguished simply by examining the assessments themselves. Each of these assessment approaches will be discussed with a description of the basic approach, key features, delivery systems, and issues and challenges.

The final component of the assessment system is classroom-based assessment, which can be summative, interim, or formative. Classroom based assessments are discussed in detail in the preceding descriptions of summative, formative and interim assessments. Classroom-based assessments are described in terms of their different purposes, development, uses, professional development demands, and assessment quality. While it is convenient to describe the components of the comprehensive statewide assessment system separately, it is useful to keep in mind that the effectiveness of the system depends on the interconnections of the parts. The impact of the system on students’ learning is greatly enhanced when all components are working together and focused shared goals and aspirations.
A. Summative Assessments

Summative assessments are designed to be primarily assessments of learning. They are given after a period of instruction and are used to evaluate students’ performance relative to a set of content standards for which a set of performance categories (e.g., A, B, C, etc.) or achievement levels (e.g., Below Basic, Basic, Proficient, Advanced, etc.) have been defined. Tests administered at the end of a unit of instruction or semester for the purposes of grading are summative assessments.

The most well known form of summative assessments is statewide end-of-year tests. These tests are relatively short (generally 40- to 60 points worth of questions) compared to the large number of state content standards they are used to measure. They lack instructionally useful depth and detail in assessing any particular area of content because they are spread so thinly across so many standards. State assessments are usually given as late in the school year as possible to allow for maximum instructional time. This generally means, however, that results are not returned on a schedule that allows them to be used to modify or adapt instruction for the students who have taken the tests. Perie et al. (2010) describe the intended role and actual impact of these statewide assessments when they state: “While many had hoped that these once-a-year tests would provide instructionally useful information, educators and others know that this is not occurring.”

An approach to extracting some instructionally useful information from summative assessments is described by Ryan (1986) and Ryan and Rowls (1988). In this approach,

- data from multiple annual assessments were combined to increase the range and depth of the content objectives assessed;
- comparisons were made for students across different performance levels (roughly analogous to Below Basic, Basic, Proficient and Advanced) in terms of the average proportion of items answered correctly, the content characteristics of students’ choices of wrong answers, and the substantive characteristics of items that maximally differentiated students across the four performance levels;
- actual test items were reviewed along with the data to develop descriptions of students’ strengths and weakness and learning gaps between higher and lower achieving students;
- results were reported using content focused narrative with tables of mean proportions and ranges when appropriate;
- for the reading tests, prototypical reading passages and cloned test items were developed to illustrate the major problems students were having;
- a team of teachers and other content specialists elaborated on the interpretation; and
- curriculum and instructional materials were produced to aid teachers in dealing with the observed weaknesses.

Students’ learning challenges revealed through the re-analysis of multi-year summative assessment data are systemic in that they are observed across the state and persistent in that they are based on data from multiple years. Information from the re-analyses of summative data
seems especially valuable in a comprehensive assessment system because it provides a focus for targeting interim and formative assessments. Because these assessments are sensitive to systemic and persistent learning gaps, they can help teachers identify students with common learning challenges. Effective instructional materials and interventions can be developed to respond to these learning gaps and then well-targeted interim and formative assessments would be used to identifying the individual students who need help.

This discussion is not presented to propose a particular procedure but is used to illustrate the potential for using of summative assessment data to support instruction and learning. The additional cost for combining, organizing, and reanalyzing data from multiple years is almost trivial compared to the considerable cost of developing, administering, scoring, and reporting annual assessments. It seems worthwhile to examine ways to reanalyze summative data given the potential educational benefits and the modest additional cost compared to the considerable initial expenditures.

Developing and Delivering Summative Assessments

The state will continue to develop and administer the summative assessments working with their contractors and advisors as needed. As with all assessments, the design, development, and implementation of summative assessments should reflect the needs of the diverse subpopulations of students who will be taking the assessments. The current system employs a paper-and-pencil system and this approach will be changed to a computer administered or computer adaptive system as rapidly as resources allow. A number of issues must be considered in moving to a fixed-form computer administered or computerized adaptive testing (CAT) system.

- **Hardware and software capacity** for schools, districts, and the state must be evaluated carefully and meet technical standard to assure that no students are advantaged or disadvantaged by problems with technical support.

- **Item and test security problems** for computer delivered assessments can be unusual and challenging and are often related to the technical resources of schools and districts. Groups of students taking computer-based assessment at different times, for examples, can create some security risks. Written guidelines and policies for maintaining proper security must be developed and implemented as part of the transition to a computer-based assessment system.

- **Item development demands increase dramatically** when computerized assessment systems are used and an assessment system is only as good as the items used on the assessments. Computerized adaptive assessment, in particular, requires a large number of items. The “bells and whistles” of CAT are irrelevant if the fidelity with which the items measure the intended learning targets cannot be demonstrated. The potential benefits of computerized assessment system must be carefully evaluated against the cost of developing and maintaining adequate and appropriate item banks.

Summary

A comprehensive and inclusive assessment system, accessible to a diverse group of learners, must continue to use end-of-course or end-of-year summative assessments to provide an accounting of educational progress and achievement and to respond appropriately to state
and federal requirements. Summative assessments as have been typically employed will continue and, when appropriate, end-of-course examination will be employed. Given the considerable expense paid for the summative assessments, strategies should be developed and implemented to extract instructionally relevant information from summative assessments. Summative assessment should move to a computer-based delivery system as soon as practical but caution and care must be taken in the transition from paper-and-pencil delivery.

B. Interim Assessments

Interim assessments have some features of both summative and formative assessments. Salient features of interim assessments include:

- assess a well-defined and somewhat narrow portion of the curriculum or instructional sequence;
- are given periodically, more often than end of year or end of course assessments but less often than formative assessments;
- are group-based, not individualized;
- provide information about students’ progress relative to learning goals and objectives; and
- yield results that can be combined from classrooms to the school level, schools to district level and district to state level.

Interim assessments are sometimes referred to as benchmark or quarterly assessments and are occasionally marketed as diagnostic or formative assessments. Perie et al. (2007, 2009) describe three major purposes and uses of interim assessments: instructional feedback, evaluating and monitoring the impact of curriculum and instruction, and predicting students’ performance on some future criterion. The intended purpose of an interim assessment informs and influences the design, procedures for administration, and use of the assessment results.

Timely Assessment Information for Instruction

Interim assessments designed for instructional purposes can be used to identify students and groups of students who may not be making adequate progress throughout the school year. However, in a summary of interim assessments, the authors of *Overview of Selected State Assessment Programs* (Great Lakes West Comprehensive Center 2009) point out that, “… unless teachers are able to use this information from these assessments immediately at the individual student level, the assessments will have little real-time instructional value” (p.10). This report also describes a conflict many teachers face when information from interim assessments reveals gaps in their students’ learning. When confronted with information about a learning gap, teachers need to decide if they can devote the time and effort needed to reteach the material or move on to new material that has not yet been covered. This circumstance is especially problematic if the un-mastered material is not prerequisite to learning the new material.

Interim assessment information is valuable in identifying learning problems of individuals and groups of students. This information can be used to continuously improve the impact of teachers’ initial instruction, and thus, over time, reduce learning gaps. In addition, interim
assessment information can also be used to specify the targets of formative assessments so that these real-time assessments will be sharply focused in those areas where student have been most challenged.

**Interim Assessment Information for Evaluation and Monitoring**

Interim assessments may not be ideal for informing real-time instructional decisions; however, they can provide very useful information for monitoring or evaluating curriculum, teachers’ instruction, and students’ learning. Information from interim assessment can be aggregated and reported at the school and district levels to inform principals and central staff, respectively. Such information can be used to review the relative effectiveness of various programs, identify school and district-level learning deficiencies, manage students’ progress, and target professional development. Interim assessments focus on more specific and narrower portions of content domains than end of year summative assessments and thus provide content-focused feedback that is more fine grained and likely to be more useful for subsequent instructional planning. Interim assessment data aggregated at the state level can serve many of the same instructional support purposes as those at the school and district levels. In addition, interim assessment information can be especially useful in identifying the content and skills to target in the development of formative assessments.

**Interim Assessment for Prediction**

Interim assessments provide data that can be used to predict students’ subsequent performance on other achievement measures and outcomes. The most obvious application is the use of interim assessments to predict how well students are likely to do on end-of-course or end-of-year summative assessments. Data collected over time can be used to develop prediction models that will serve as “early warning” systems for students whose progress may not be promising of eventual success. Predictions of performance on other external examination, such as NAEP, can also be developed, as can predictions of success in future academic or career related endeavors. Prediction models can serve a variety of educational and policy purposes and data from interim assessments and other assessments can be used in developing statistical models for predictions.

Problems with prediction models based on interim assessment results can occur and raise a variety of issues that need to be addressed before the prediction results are interpreted.

- *There should be some logical connection and explanation for why the interim assessment would be a useful predictor of the criterion measure.* Performance on an interim assessment is logically connected to performance on a summative assessment because they both measure content sampled from the same set of content standards. Thus it is reasonable to expect performance on an interim assessment to predict summative assessment performance. When no logical connection can be made, however, then the basis of the prediction is likely to be aptitude and general ability and not knowledge and skills that might be sensitive to and altered by instruction. In general, statistical predictions are not advisable unless there is a logical connection between the two measures and the validity of the criterion measure has been established.
• **The assumptions of the statistical model must be examined and tested.** Statistical predictions with selective subsets of students frequently use data from a restricted range and challenge the statistical assumption that data are normally distributed.

• **Effective instruction based on interim (and formative) assessment information can undermine the predictive power of these assessments.** Students whose performance on interim or formative assessments indicates learning deficiencies are predicted to have relatively low achievement on summative assessments. If effective instruction is applied in response to formative assessment information and students recover from their deficiencies, then they will outperform their predicted level of achievement. In this situation, the prediction will not be very accurate but the outcome is highly desirable.

**Developing and Delivering Interim Assessments**

Interim assessments are content-focused and designed to measure progress along a learning progression. Interim assessments are developed by teachers, districts, and the state. Input from experienced teachers and content experts is essential in developing valid interim assessments at any level. In addition to classroom and district interim assessments, the state should be responsible for constructing and distributing a set of interim assessments and the data information system to support the assessments. Interim assessment can be administered as fixed forms that are sensitive to selected or defined ranges content objectives along the learning sequence. As such, they would provide information about all students on a standard set of items measuring a specific portion of the learning progression.

Interim assessments can also be administered using a computerized adaptive system. Such a system would provide a more accurate measurement of students along a broader range of the learning progression since relatively low- and high-achieving students would take items appropriate for their estimated achievement levels. As mentioned earlier, computerized adaptive systems require extensive item development efforts and a large number of validated items. It would be prudent to conduct a careful evaluation of the cost of such a system relative to the potential educational value.

The data information system designed for interim assessments should have the capacity to support and integrate school and district developed interim assessments as well as the state produced assessments. The data information system can be used to collect, score, and analyze data locally or centrally. An interactive computerized system should be developed to enable schools and districts to extract customized score reports from the central data base to support their instructional efforts.

**Summary**

A comprehensive and inclusive assessment system that is accessible to a diverse population of learners can use interim assessment information for a number of valuable purposes. Clearly, interim assessment information can be used to plan and modify instruction over time through focused professional development efforts and thus improve the impact of initial instruction. Improving the impact of teachers’ initial instruction is essential and can reduce the demands placed on interim and formative assessments. Interim assessments can be
used to reveal learning gaps for which sharply focused formative assessments can be developed and used by teachers in real-time instruction. Predicting students’ likely performance on important future activities is a valued application of interim assessment data and, with caution and care, procedures can be explored to develop and apply appropriate prediction models. Interim assessments can be delivered, used, and results reported via electronic resources.

C. Formative Assessments

Many definitions and models for formative assessment are provided in the assessment literature (Sadler, 1989; Black and Wiliam, 1998a, 1998b, Popham, 2008; Perie et al., 2010) and these all share certain common elements. Popham’s definition seems to include most features of other definitions and is a useful focal point for discussion: “Formative assessment is a planned process in which assessment-elicited evidence of students’ status is used by teachers to adjust their ongoing instructional procedures or by students to adjust their current learning tactics” (p.6).

By way of elaboration, Popham points out that this definition reflects several critical attributes of formative assessment:

- formative assessment is a process, not a particular test;
- the process is planned not ad hoc;
- assessments are used to elicit evidence regarding a student’s status relative to a particular skill or body of knowledge;
- teachers respond to the assessment information by adjusting their ongoing instructional activities; and
- students can respond by adjusting what they are doing in their efforts to learn.

Some authors have an additional requirement that for an assessment to be considered formative there must be evidence that the use of the assessment has actually led to increased learning (Brookhart, 2010; Nichols et al., 2010). Others say that looking at assessment consequences is certainly an important activity but is not an essential part of the definition of validity (Popham, 1997; Mehrens, 1997). When data for evaluating the impact of a formative assessment system on students’ learning are available, it seems prudent to examine such information.

Central to all formative assessment approaches is the need for clarity in understanding and communicating what students are supposed to be learning so that the formative assessment can be designed or selected to shed light on progress and gaps in the attainment of the desired learning goals. Shepard (2006) describes the formative assessment model of Sadler (1989) in which Sadler points out “…that it is insufficient for teacher merely to give feedback about whether answers are right or wrong. Instead, to facilitate learning, it is equally important that feedback be explicitly linked to clear performance standards and that students be provided with strategies for improvement.”

While certainly recognizing the primary role of teachers in using formative assessment information, many authors also describe the value of formative feedback from student self-
assessment and students’ adaptive or self-regulated learning based on formative assessment information (Popham, 2008; Black and Wiliam, 1998a). Thus, part of the teachers’ task is to teach students how to use assessment information themselves to determine where they are in their learning progression, what challenges impede their learning, and what strategies can help them address their needs.

The nature of the feedback students receive from formative assessments is vital to the efficacy of the assessment. Research (Black and Wiliam, 1998a) indicates that feedback is most effective when

• teachers and students share a common, clear and accurate understanding of the learning goal;
• the feedback and information focuses on the learning goal or target; and
• the feedback focuses on the students’ work on problems or tasks in the assessment.

Feedback is least effective when it

• is gratuitously positive;
• references students’ growth; and
• is normative relative to other students or performance levels.

Developing and Delivering Formative Assessments

Formative assessment is essentially a classroom activity and can take many forms. These include systematic questions and probes, observations of student work and performance, student written works and other products, student responses to carefully structured tasks designed to provide insights into students’ cognitive processes, and even student answers to and discussion of selected response questions. They may include complex intricate problem solving tasks that students work on over a period of time, tasks that require important higher order cognitive processes, outcomes of students’ efforts in production and performing arts, and cognitive processing skills such as fluency and automaticity that can only be evaluated as students demonstrate them. Formative assessment includes procedures and materials for students’ self assessment and peer assessment. These approaches have been shown to be helpful in enhancing students’ learning but are often neglected.

A useful component of a comprehensive assessment system would be a compendium of classroom-based formative assessment materials and practices. This could be developed from a survey of current practices around the state as well as from a review of research and materials on best assessment practices. The purpose for such a collection would be to provide teachers with exemplary models of classroom-based formative assessments.

In a comprehensive assessment system, formative assessments can be developed at the school, school district, and state levels. The participation of teachers in developing formative assessments is essential regardless of whether the assessments are produced at the local or state levels. It is likely that targeted professional development would be needed to help teachers in developing effective formative assessments. Preservice and inservice professional development for constructing formative assessments should emphasize:
• approaches for constructing and communicating detailed and rich descriptions of learning targets, including the criteria used to evaluate students’ performance, products and work;

• strategies for helping students understand these learning expectations;

• techniques and procedures for selecting and/or constructing assessments prompts, guides, and other appropriate materials;

• strategies and procedures that will help teachers identify barriers in presentations, assessment formats, or response modes that prevent some students from demonstrating their knowledge and skills;

• procedures and techniques for eliciting and recording students’ performance;

• mechanisms, strategies, techniques and effective approaches for providing feedback; and

• the need to develop specific instructional responses to learning deficiencies revealed by the assessment.

State-developed and distributed formative assessments are a valuable part of a comprehensive assessment system. These assessments would target learning outcomes that data from interim and summative assessments show to be problem areas. Formative assessments would also be constructed based on a review of state standards by teachers and other content experts who could identify areas that students would likely find challenging. State produced formative assessments could be distributed and used in either hard copy or electronic formats. The choice of approach would depend on school resources and also on the content and developmental level of the students. Tasks used for formative assessment could be web-based, could include directions for use, and also contain (or have links to sites that offer) instructional strategies aimed at learning gaps the assessment is designed to detect. Not all data from formative assessment would necessarily be recorded and aggregated to the school, district, or state level. Data could be collected on a sampling basis, sampling students and different assessments, to monitor the use of the assessments and students’ progress.

Formative assessments that use selected-response or short answer questions can be easily supported with a computer-based system but the instructional value of these assessments must be carefully evaluated. Assessments with tasks containing elements that students can manipulate on a screen can also be computer-based. Computer-based formative assessments for more complex tasks are more difficult to develop and manage because their instructional impact depends on rapid scoring and feedback to students. Nevertheless, approaches to providing online formative assessments of higher level skills should be explored.

Computer adaptive formative assessment systems can provide rapid information about students’ responses to various questions but this does not mean that the information will necessarily help teachers. Feedback from a formative assessment system must be focused on specific learning targets and the feedback should relate specifically to the students work and attempts to solve the problems. Simply reporting a score on a formative assessment does not usually provide the feedback needed to assist teachers and students.
The challenges for teachers using formative assessment results are important and need to be dealt with if instruction is to have an impact on student learning. The literature and research supporting the use of formative assessment does not generally include a discussion of the challenges teachers face in deciding when, how often, and for whom such assessments will be used. However, school improvement facilitators have found that the major challenges faced by teachers are the problems of deciding how much time can be devoted to the use of formative assessments and when will they find time for the remediation or the re-teaching that formative assessments indicate are needed. In a relative large class, how do teachers decide when to move on in the instructional sequence when some students might benefit from the remediation or re-teaching suggested by formative assessment but many more students seem ready to advance to new material? These challenges are especially problematic when many students might be able to learn the new material without necessarily having mastered the particular content taught in the earlier curriculum sequence. These and other implementation challenges are serious practical problems that must be addressed in professional development activities. Strategies for constantly improving the effectiveness of teachers’ initial instruction are a major focus for improving student learning in the long run. In addition, teachers need to learn how to use different but appropriate strategies for remediation instruction.

Summary

A comprehensive and inclusive assessment system will include formative assessments as an integral part of a system that uses assessment for learning. Formative assessments that are effective in increasing students learning focus on learning goals and performance expectations that are clear to students and teachers and are embedded in a cognitive model of how students learn. Effective formative assessments use a variety of techniques to elicit evidence of student’s learning, provide feedback that focuses attention on the assessment task at hand, and provide students and teachers guidance as to what to do next to improve learning and to acquire proficiency on the targeted content and skill.

One part of a formative assessment component can be a compendium of classroom “best practices.” It is not easy to collect and codify examples of classroom-based formative assessments but a comprehensive assessment system should focus attention on these, help cultivate their development, facilitate procedure for collecting and distributing them in a variety of formats and platforms, and inject them when possible and appropriate into professional development programs. A major focus of the formative assessments component are state developed assessments or assessment systems. These formative assessments can be made available to schools on-line and the capacity to support various types of fixed-form or adaptive assessments should be advanced. Computer-based formative assessments can be part of an integrated electronic information system as long as they are built from banks of item that are adequately aligned to the state content standards.

Formative assessments are only as valuable as the capacity that teachers have for using them effectively. Resources should be allocated to appropriate professional development to support teachers in their efforts to use assessment to enhance students’ learning.
D. Classroom-based Assessments

Classroom-based assessments are an important component of the Washington statewide comprehensive assessment system. The legislation that created the statewide student assessment redesign (RCW 28A.300.041) finds that a “… component of the assessment system should include classroom-based assessments, which may be formative, summative or both. Depending on their use, classroom-based assessments should have the same design elements and characteristics described in this section for formative and summative assessments.”

The major design elements of formative and summative assessments have already been described in Sections A (page 5) and C (page 10), respectively, and are not repeated here in detail. The subsection “Developing and Delivering Formative Assessments” (page 11) describes formative assessment as essentially a classroom activity, offers numerous examples of formative assessments as types of classroom assessments, and elaborates on the features of formative assessments that make them effective in supporting and advancing students’ learning. The information and suggestions offered in Section C, Formative Assessment, applies directly to classroom-based assessments.

The most serious challenge in teachers’ use of classroom-based assessments is the degree to which such assessments are properly aligned with the content standards. Teachers use a wide variety of assessments in their classrooms but many are chosen based on availability and ease of use. Classroom assessment must be carefully evaluated in terms of their alignment to the content standards if they are to be an effective component of a comprehensive assessment system. This alignment check should include evaluating whether the cognitive demands of the classroom assessments match the same level of cognitive demand expressed in the standards and reflected in various summative assessments.

Classroom-based Summative, Interim, and Formative Assessment

Teachers’ classroom-based assessments routinely include summative, interim, and formative assessments and teachers commonly use all three types of assessments in their everyday teaching practices. End-of-chapter or unit tests are used as interim and summative assessment but in classroom practice these assessment can also be used formatively. Teachers use all manner of informal and formal formative assessments through such strategies as classroom questions, discussions, observations, quizzes, test, discussions, reviews of students’ classroom work and home assignments, demonstration projects, and lab reports. Most student-teacher interactions provide teachers with some insight into the student’s skills, knowledge levels, and needs. Classroom assessments used primarily for grading purposes are summative assessments while classroom assessments used to support and facilitate learning are formative assessments.

Developing Classroom Assessments

Teachers may develop various types of assessments for classroom use or they may select them from text books or other vendor supplied materials. Many districts supply teachers with assessments designed for classroom use and the statewide assessment system may make various types of assessments available to teachers. Teachers may also use assessment materials that
have already been developed but modify them to ensure that they are aligned with the content standards. Since teachers do not do the initial development, they have more time to focus on refining the assessments to better match their students’ characteristics and needs. As the comprehensive assessment system develops, it will be important to clarify how teachers will be expected to use state supplied assessments. State-supplied assessments may be used as examples for teachers to model, as drafts for teachers to modify, or as prescribed assessments designed to be used “as is” by teachers.

**Using Information from Classroom Assessments**

The results of classroom assessments have traditionally been used by teachers in their classrooms. Teachers use classroom-based summative assessment for grading and formative assessments to better understanding students’ current knowledge and skill level in order to adapt their instruction and improve students’ learning. District-supplied assessments are often used for classroom purposes and information from district assessments are often collected across classrooms to the school level and across schools to the district level. Information aggregated by classrooms and schools are used to inform broader curricular and policy issues. The development of the statewide comprehensive assessment system will need to clarify and describe in detail any plans to collect and aggregate information from classroom-based assessments. The plan should indicate how the information will be stored and retrieved, who will have access, and how the information will be used and reported.

**Professional Development**

Developing, implementing, and interpreting information from classroom assessments requires skill and experience that many classroom teachers have not yet developed. Thus, the effective development and use of classroom assessments will require a considerable professional development effort. Details of the assessment and instruction-related competencies of teachers and district consultants are described in a subsequent section devoted a discussion of professional development efforts that are needed for an effective comprehensive assessment system (page 20). Professional development of classroom assessment competencies should focus on assessment for learning, the support of instruction, and assessment approaches that embody the best features of formative assessment described earlier. Considerable emphasis should be placed on helping teachers learn to properly interpret the results of their classroom assessments.

**Technical Quality of Classroom Assessments**

Classroom assessments should be valid, fair, and free of bias. In many cases, traditional technical criteria for evaluating assessments are not appropriate for evaluating classroom assessments. Classroom assessments used for formative assessment purposes should be reviewed to see if they embody the best features of effective formative assessment as described in Section C on Formative Assessment. Some classroom-based formative assessments are as much instructional practices as assessment procedures and as such can be monitored as part of instructional supervision efforts. A subsequent section examines the validity of classroom assessments (page ).
Summary

Classroom-based assessments are an important component of the Washington statewide comprehensive assessment system. Teachers routinely use summative, interim, and formative assessments in their classrooms. Because classroom-based formative assessments are most closely connected to instruction and learning, they have the greatest potential for improving student learning. Classroom assessments can be developed, selected, or adapted from exiting assessments by teachers. Information from classroom assessments are used by teachers but could be aggregated across schools and school districts. Any plans to collect and aggregate classroom assessment information need to be described and reviewed in detail. Classroom assessments, like all components of the assessment system, will require considerable preservice and inservice professional development. The criteria for evaluating the quality of classroom assessments should also include that informal classroom assessments are tightly embedded in classroom discourse instructional strategies.

4. Validity of Assessments and Evaluation of the Assessment System

The quality of an assessment system depends on many factors. Of central importance is the validity of the assessments used, but the professional development of the instructional staff, materials developed for the system, and the system’s electronic information system combine to create the overall impact of the system. Described in this section are considerations for examining the validity of the assessments and some general consideration for evaluating the system as a whole.

Validity: Evaluating Assessment Quality

The value of an assessment system is limited by the quality of the assessments that make up the system and the critical feature of all assessments is their validity. The concept of validity has been developed and refined over time and Sireci (2009) summarizes the most current thinking on validity by pointing out that

- validity is not an inherent property of a test itself;
- validity is actually a property of the interpretations or actions that are made on the basis of test scores; and
- validity must be evaluated with respect to the purposes of the test and how the test is used.

Kane (2009) elaborates on procedures for validating the interpretations and uses of tests scores and Schafer et al. (2009) provide a very useful analysis of validity evidence as it has been developed and submitted by states to support compliance with NCLB requirements.

These writers all describe the five sources of validity evidence used as a validity framework in Standards for Educational and Psychological Testing (AERA, et al., 1999). These are evidence based on:

1. test content;
2. student response processes;
3. internal structure;
4. assessment’s relationship with other variables; and
5. consequences of testing.

In their chapter, Schafer et al. (2009) provide an excellent elaboration of the details for each source of evidence with clear information about what evidence a state might collect in support of the validity argument. In addition, the authors review the response of the USDE to states’ NCLB submissions and summarize the additional evidence USDE has required states to supply in support of their validity case. The summary is quite detailed and is organized into the five-category validity framework. The peer review feedback from five states is also summarized into the five category framework.

As plans for the comprehensive system are developed, it is critical that they include specific plans for collecting validity evidence for each component of the system. It would be useful to structure the validity framework using the five categories of validity evidence as described above in advancing the validity portfolio for the comprehensive assessment system.

Assessment Validity for All Students

The validity case for a comprehensive assessment system must carefully consider and plan for the inclusion of all students. The assessment system plan must include assessments that are valid for (1) students who vary widely on their current level of achievement; (2) students who come from culturally and linguistically diverse backgrounds; and 3) students with special cognitive and physical needs.

The components of the assessment system must be designed to provide information for both relatively low- and relatively high-achieving students. Assessment information for these students must support valid inferences about what they know and can do and lead to appropriate prescriptions for instructional interventions. Interventions for relatively low-achieving students might be remedial while interventions for relatively high-achieving students would build on their strengths and challenge them to advance further.

The assessment system must also include procedures for evaluating the validity of the various assessment components used for students who come to school from widely varying cultural and linguistic backgrounds. The validity of assessments for a culturally and linguistically diverse student populations is especially complex when some assessment components emphasize support for instruction and others emphasize evaluation of instruction. Valid assessments for instruction must enable teachers and others to identify students whose performance is limited by cultural and language barriers. Such students may need language instruction before they can benefit from instruction in the subject area. Assessments of various types, especially diagnostic formative assessments, must provide information about students’ familiarity and fluency in the formal or academic type of language and language protocols used in schools. Students’ success in schools is influenced by their academic language skills in English and formative assessments are perhaps the best tool for monitoring these skills.

*Students with special needs* refers to a diverse subgroup of the general student population. This group ranges from students with mild learning disabilities to students with severe cognitive and physical disabilities. Special needs students may be assessed with the
general assessments, with assessments based on modified content standards, and with modifications or accommodations to the assessments. The full range of students with disabilities should be carefully described and the challenges in validly assessing these students addressed in planning, developing, and implementing the comprehensive assessment system.

Building a comprehensive and inclusive statewide assessment system is a daunting task. But the value of a comprehensive and coordinated system, as opposed to a single assessment or isolated assessment component, is that it provides several opportunities and in most cases redundancies in assessing all students. Kitterlin-Geller (2008) offers useful advice for assessing all students:

. . . we cannot focus on creating one test that fits all students. Instead, we need to develop a flexible testing system [italics added] that maintains the integrity of the tested construct while supporting the needs of the diverse student population. By recognizing the limitation of the current assessment model and conceptualizing how we test students with special needs, we can alleviate the current problem of unprincipled test changes for students with disabilities and English learners that have unmediated effects on test interpretations.

Validity of Classroom Assessments

The validity of some classroom assessments becomes difficult to describe when well designed assessment activities become indistinguishable from effective instructional activities. Blurring the sharp distinction between instruction and assessment may be highly desirable and in many cases may advance students’ learning very effectively. Such assessments cannot be judged against typical measurement criteria since they are instructional activities as well as assessment activities. The review of these types of classroom instruction/assessment activities might more properly be considered part of the responsibility of those charged with instructional supervision and be monitored and evaluated as instructional innovations.

A common and very serious problem with some classroom assessment is the lack of alignment between the items or tasks and the content objectives they assess. Teachers’ assessments may not measure the objective at the cognitive level specified in the objective, that is, the assessments can often be at a lower level of learning than expected. In such cases, the teachers’ formative assessment may indicate that students are doing well, but students’ performance on subsequent summative assessments may indicate that they have not achieved the objective at the expected level.

Evaluating the Assessment System

It is essential to examine the quality and value of the assessment system as a whole in addition to examining the validity of individual components of the assessment system. An evaluation design should be developed for the overall assessment system and implemented to determine whether or to what extent the system has had the impact it was designed and funded to achieve. As with all educational interventions, the assessment system should be judged based on how well it achieves its goals, not merely in terms of the degree to which it was implemented on schedule as designed. The evaluation must include a component to
collect and evaluate information about unintended outcomes, some of which might bolster the validity case and some of which might detract from the overall educational effort.

Summary

The complexities and demands for validity across a multi-component assessment system designed for a diverse student population are critical and considerable. Careful attention and concerted efforts must be devoted to the issue of validity for all students with personnel, materials, and financial resources focused on reviewing current best practices and infusing the most valid approaches into developing and maintaining the assessment system. In addition, a systematic evaluation of the overall impact of the assessment system is needed to account for the massive investment in the enterprise.

5. Statewide Data Information System

A comprehensive assessment system must be supported by an equally comprehensive system for collecting, processing, storing, retrieving, and reporting information. The system should be highly integrated, connecting all components of the system, and easily accessible to the different audiences who might need and are authorized to have access at different levels. A brief description of some of the key features of such a system includes the following.

- Links to instructional and assessment resources aligned with state and local content standards that can be used for professional development activities.
- Links connecting assessment information to relevant content standards and instructional resources to which the assessments are aligned.
- An item and test storage and retrieval engine that schools and districts can use to develop and store items and construct their own assessments. The system would enable users to score their own assessments, store results, and produce a variety of score reports.
- The capacity to deliver, score, analyze, warehouse, and retrieve information from local and state formative, interim, and summative assessments.
- Flexible score reporting capacity designed so users can customize reports depending on their level of access and the intended use of the reports.
- The capacity to display test results in a variety of tabular and graphical formats that are easily understood by all users.
- Tiered secure access so students, parents, the general public, teachers, school officials, and others can access information for which they are authorized including links to information in different languages about students’ performance.
- Links to resources to help all users properly interpret test results.
- A comprehensive set of connections linking all student data sources (e.g., assessment, demographic, course, academic intervention exposure, and attendance information,
etc.) so that information about students can be accessed, aggregated, or disaggregated.

- The system will provide up-to-date descriptions of students’ current learning status and learning history, measures of growth across time, and support studies into factors that may influence students’ learning.
- Links to strategies and resources which help schools and districts use assessment and other data in their yearly improvement planning process.
- Links to reports and other resources related to technical aspects of the assessments

6. Professional Development

A revised and comprehensive assessment system has major implications for teacher preservice and inservice professional development. Assessment for learning implies professional development in instructional methods that incorporates assessments in teaching materials, in instructional strategies, and in helping teachers learn to adapt their instruction based on assessment information. Inservice professional development would integrate assessments for learning into ongoing workshops and collaborations with other teachers and staff developers dealing with teaching techniques in particular content areas at certain grade levels. The need for more general professional development in assessment for preservice and inservice teachers is described by Popham (2009) in a chapter entitled, “Abysmal Assessment Literacy.”

The professional development that teachers will need to fully utilize the comprehensive assessment system described in this paper include, but are not limited to, the following assessment and related instruction competencies for teachers and district consultants:

- Learning to align instruction and assessments at the cognitive level and scope of the curriculum objectives;
- Developing and communicating detailed and rich descriptions of learning targets, including the criteria used to evaluate students’ performance, products, work and progress toward proficiency;
- Helping students understand the learning expectations;
- Developing a variety of high quality, aligned classroom assessments such as multiple-choice, matching, and short answer items, performance tasks, problem sets, writing prompts, projects, portfolios, instructionally embedded assessments, etc., and the scoring rubrics that are also aligned to the task and curriculum objectives;
- Developing assessments that are sensitive to the needs of diverse learners including relatively high- and low-achieving students, students from culturally and linguistically diverse backgrounds, and students with a wide range of special education needs;
- Organizing, analyzing and displaying data from assessments;
- Interpreting assessment results including, criterion-referenced and norm-referenced test scores as well as aggregated and disaggregated data;
• Developing effective techniques and approaches for providing feedback to students;
• Using assessment results to identify appropriate content and instructional strategies to improve initial instruction and for remediation;
• Developing plans and strategies to extend instructional time for students when they need it;
• Teaching students to understand and use assessment results to monitor and improve their own learning;
• Incorporating assessment results appropriately in grading students; and
• Communicating assessment information effectively to parents and other educational stakeholders.

At a broader level, a comprehensive professional development program should recognize and work within the national Standards for Teacher Competence in Educational Assessment of Students. These national standards were developed by the American Federation of teachers, National Council on Measurement in Education, and the National Education Association (1990). The seven major principles of these standards state that teachers should be skilled in:

1. choosing assessment methods appropriate for instructional decisions;
2. developing assessment methods appropriate for instructional decisions;
3. administering, scoring and interpreting the results of both externally-produced and teacher-produced assessment methods;
4. using assessment results when making decisions about individual students, planning teaching, developing curriculum, and school improvement;
5. developing valid pupil grading procedures which use pupil assessments; and
6. communicating assessment results to students, parents, other lay audiences, and other educators; and
7. recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information.

7. Summary and Discussion

A brief synopsis of the major sections of this paper is contained in the Executive Summary, pages i to ii. Below is a brief summary and discussion of major issues that must be considered as the Washington comprehensive assessment system is planned, developed, and implemented.

Curriculum, Instruction, and Assessment Connections

The assessment system must serve as a mechanism for coordinating and aligning curriculum, instruction, and assessment. All aspects of the system must be aligned with the
curriculum standards and it is especially critical that the validity of the assessments be evaluated by checking their curriculum alignment. Revision to content standards will occur from time to time during the life of the program and such revisions should be reflected in changes to the assessments. Teachers and others may not be as familiar with the state content standards as needed to support a comprehensive assessment system, especially when standards are revised, and professional development activities should be used to address this issue.

**Summative and Interim Assessments**

Summative, interim, and formative assessments all play important roles in supporting assessment for learning and assessment of learning. Summative and interim assessments serve several purposes including documenting students’ learning status and progress, providing information for revising instructional strategies and materials, and focusing formative assessments on learning targets that students find especially challenging. There is useful instructionally relevant information in data from summative assessments and serious efforts should be made to develop this information.

**Formative Assessments**

Formative assessments are commonly used at the classroom level although they can be developed and supplied to schools from the school district and the state. Formative assessments are most useful in helping students when teachers use them as part of the instructional process. In many cases, the best formative assessments will be spontaneously employed by teachers at timely teachable moments. Formative assessment may be formal or informal and can take many forms. The creative and effective use of formative assessments should not be hindered by fixed schedules and prescribed procedures for their use. Research shows that formative assessments benefit students most when teachers and students understand the learning objectives they are designed to assess, when feedback from teachers focuses on the learning expectations and the actual work or processes the students are using, and when the feedback helps students understand what they need to do to improve their own learning.

**Classroom-based Assessments**

Classroom-based assessments are used routinely by teachers and include summative, interim, and formative assessments. Teachers can develop or select assessments for classroom use or modify existing assessments so they are closely aligned to the content standards. Classroom-based assessments are the assessments most closely connected to instruction and learning. They have the most promise for enhancing students’ learning if they embody the features of formative assessment best practices. Most teachers have not yet developed all the skills they need to use classroom assessments most effectively and a substantial professional development effort will be needed to support teachers in the development, use, and interpretation of classroom assessments.

**Validity**

Validity is an attribute of the use and interpretation of assessment information and not a characteristic of the assessment itself. Alignment to district and state content standards is
the central feature of the validity argument since assessment information will generally be used to make inference about students’ knowledge or growth with respect to these standards. Evidence must be developed to show that the assessment system validly measures students across the full range of achievement levels and students from different cultural and linguist backgrounds.

**Data Information System**

An integrated comprehensive assessment system requires a similarly comprehensive data information system to support it. If people cannot store, retrieve, and use assessment information combined with information from other sources then collecting the information is an exercise with little educational value. The value of an assessment system is that it provides information that teachers and others can use to increase students’ learning.

The use of computerized adaptive testing has considerable promise as part of a comprehensive assessment system. Such systems place additional burden on school and district resources and centralized information processing but these technical issues seem to be manageable over time. Adaptive testing is more problematic for item development. A large number of test items must be developed, with items fully aligned to content standards, in order to make computer adaptive assessment worthwhile.

**Professional Development**

The state comprehensive assessment system will require a considerable professional development effort. Professional development for preservice and inservice teachers is needed to improve the general level of assessment literacy, to equip teachers with the skills needed to effectively develop and use formative assessments, and to enable teachers to link appropriate instructional action to the results of the formative assessments.

Research cited in the paper points out that teachers are better able to identify what students do not know than they are able to identify what instructional steps to take next to address the students’ learning deficiencies. Professional development should focus strongly on helping teacher learn strategies and procedures for developing appropriate instructional interventions targeted to students’ learning weaknesses as identified by various assessments. In this regard, workshops and training on formative assessment should emphasize procedures for developing instructional responses aimed at the learning deficiencies the assessments reveal.
References


Great Lakes Comprehensive Center (January, 2009). Overview of selected state assessment systems. [Paper], Naperville: IL: Author


