MAINE STATE LEGISLATURE

The following document is provided by the
LAW AND LEGISLATIVE DIGITAL LIBRARY
at the Maine State Law and Legislative Reference Library
http://legislature.maine.gov/lawlib

JOHN ELIAS BALDACCI

GOVERNOR

SUSAN A. GENDRON

COMMISSIONER

MEMORANDUM

TO: Senator Neria Douglass, Representative Glenn Cummings, and Members of the Joint Standing Committee on Education & Cultural Affairs

FROM: Patrick Phillips, Deputy Commissioner

DATE: June 18, 2003

RE: Local Assessment Systems (LAS) Guide

Attached you will find a copy of the Department's *Local Assessment Systems (LAS) Guide* for your information and review.

The LAS Guide will be used statewide at Assessment System Design Institutes. These Institutes, which have been scheduled throughout the summer, have been designed for teams of educators from each school district. Each team will comprise key players around local assessment and team members will be selected by individual school districts.

If you have any questions or would like additional information about the Assessment System Design Institutes please do not hesitate to be in touch.

LAS GUIDE

Principles and Criteria
for the Adoption of Local Assessment Systems

Prepared by the
Maine Department of Education

June 2003

STATE OF MAINE
DEPARTMENT OF EDUCATION
23 STATE HOUSE STATION
AUGUSTA, MAINE
04333-0023

JOHN ELIAS BALDACCI
GOVERNOR

SUSAN A. GENDRON
COMMISSIONER

June 10, 2003

Dear Maine Educators:

I am very pleased and proud, with the publication of the *LAS Guide*, to be able to add another significant chapter in the history of Maine's *Learning Results*. The journey toward our shared vision of high expectations for *all* Maine students has been challenging indeed. In many respects, the development of the Comprehensive Assessment System—with an unprecedented emphasis on local assessments—has resulted in a degree of learning and innovation that we could not have anticipated when the journey began. It is no small matter that the work has been widely shared and owned. Local educators, university staff, professional development service providers, and DOE staff have all contributed in major ways to this evolutionary process.

In the weeks leading up to publication, we invited local educators, the Technical and Policy Advisory Committees, and DOE staff to consider the implications for local districts—and for the system of *Learning Results*—of the Principles and Criteria contained in this guide. I wish to formally thank all those who took part in the critical review process; it has strengthened the document itself and our planning for introducing it to the field. From the draft stage to this version, the *LAS Guide* has undergone a major overhaul. A fundamental change based on early feedback was to strike a better balance in the *LAS Guide* between technical considerations and local flexibility. Throughout the document we have listed only those Principles, Criteria, and rules that would be necessary to create a truly fair, valid, and reliable system across all districts in Maine. All other matters have been left up to local discretion.

Special thanks are due to the staff of the National Center for Improvement in Educational Assessment. The Center was awarded the contract for providing technical guidance on this document, but has provided so much more. Department of Education staff deserve credit for contributions both small and large. Maine Mathematics and Science Alliance staff have also played key roles in bringing the document to completion. I am also grateful to the members of the Technical and Policy Advisory Committees and the *Learning Results* Steering Committee for their insights and guidance in ensuring we have considered all implications of the Principles and Criteria.

I will look forward with eagerness to the hard but important work that lies ahead!

Sincerely,

Susan A. Gendron
Commissioner of Education

LAS GUIDE

Principles and Criteria
for the Adoption of Local Assessment Systems

Table of Contents

Local Assessment System Guide: Principles and Criteria

Introduction

For well over a decade, many and varied standards-based models for school reform have been applied in virtually every state in the nation. Central to most of these models are efforts to strengthen the methods used to assess students' *progress* toward meeting standards and evaluating the extent to which students have in fact *attained* them. This focus has been abundantly true in Maine, as discussions about assessment have taken place in classrooms, in the Legislature and at every level of the educational establishment in between. That we are still discussing assessment so vigorously attests to both the complexity of the work and the power of the essential concept: to bring about long-term improvement of our schools, we *must* be able to determine with increasing certainty what all students know and can do.

Maine's approach to assessing our state standards, the *Learning Results*, has been to use a combination of state and local measures in building a "comprehensive assessment system." The state test, the Maine Educational Assessment (MEA), was revised in 1999 to align more closely with content standards and performance indicators established in the *Learning Results*. At the same time, schools across Maine have been aligning their local curricula, instruction, and assessments (classroom, school, and district levels) with the standards. To provide initial assistance to local districts and to clarify design parameters, the Maine Department of Education (DOE) produced a number of supportive documents: the report from the Assessment System Design Team; *Grand Ideas and Practical Work*, a collection of local assessment design efforts; *What Maine's Assessment System Should Do*, a self-assessment continuum; *Measured Measures: Technical Considerations for Developing a Local Assessment System*; the recently released Balance of Representation report; and a wide variety of tools and models that have been used in professional development events throughout the state. In addition, support for local districts has come from a number of organizations, regional partnerships, higher education institutions, and not-for-profit organizations. Support and leadership have also come from school districts themselves through presentations at statewide conferences and by regional sharing. It is a goal of this document to honor this extensive body of work and to build on it wherever possible.

Maine has employed a dynamic, evolving method in building our comprehensive assessment system. This has been due largely to the fact that there are no existing models of a system of state *and local* assessments that meet technical standards of validity and reliability. Thus, our work has been iterative, including periodic pauses for reflection and adjustment. This type of approach has produced enormous innovation in schools, districts, and among professional development agencies and organizations, but has also led to frustration and anxiety as practitioners have eagerly awaited guidance on the "big picture." Many parts of the system—the building blocks—have been developed, but with increasing urgency educators have called for guidance on the whole to which the parts contribute. This document

is intended to provide sufficient guidance on how to construct a local assessment system to serve as a plan for local action.

A Balance of Purposes:

Chapter 127, Section 4, of the Department's regulations governing Instructional Programs, Assessment, and Diploma Requirements states that the purpose of assessment is to provide high-quality information about student performance that will allow for:

- Informing teaching and learning;
- Monitoring and holding school administrative units accountable for students achieving the content standards set by the *Learning Results*; and
- Certifying students' achievement of the content standards.

To a degree, these purposes can be seen as supportive of one another, at least in the abstract. However, in practice, it can be argued that a dynamic tension exists between using assessment to inform teaching and learning (i.e., for formative purposes) and using assessment to certify achievement of standards (i.e., for summative purposes). Too much emphasis on assessment for summative purposes, for instance, could affect students' attitudes about learning and impinge upon the time available for instruction. On the other hand, too little emphasis on summative purposes could weaken school-level and district-level accountability efforts and undercut the overall fairness to all Maine students. A balanced system allows all three purposes to be achieved.

Throughout the descriptions of the Principles and Criteria set forth as Parts I and II, respectively, of this document, other underlying tensions are also suggested: between fairness to individual students and fairness to all students; between fairness to individual districts and fairness to all districts; between local flexibility and common standards across all districts; and between the need to move forward with urgency and the need to take adequate time to develop local systems at a comfortable pace. Many of these tensions can only be resolved by policy decisions, which by their nature are less exacting. The truth is that we still face unknowns in how these technical guidelines will affect schools and districts in Maine. It is important to recognize that monitoring the overall impact of state and local assessment systems—at all levels and for all purposes—will be critical to ensuring that the broad goals of Maine's *Learning Results* are met. Assessment is not an end in itself; rather, collecting and using the *evidence* of learning should not interfere with, but rather support, the *experience* of learning.

Maine's approach to assessment also rests on the belief that careful alignment of curriculum, instruction, assessment—*and* professional development—with standards will produce better results, for both formative and summative purposes. The countless hours spent by local committees to revise school and district curricula, the attention to detail by Maine teachers in developing aligned lessons and unit plans, and the focused professional development at the district level have set the stage for a truly integrated system, one that is capable of meeting *all* the purposes of the assessment system and reaching broader educational goals as well.

Building a System for Certification:
The third purpose listed above, certifying that students have met *Learning Results* standards, has proven to be a rather tall order, one that has extended and refined our understanding of the potential capability of a local assessment system. Though many studies have sought to document the validity and reliability of large-scale external tests—like the MEA— the same cannot be said for systems of *local* assessments. Much wonderful work on using assessment thoughtfully has been published, the effect of which has been to deepen our understanding of designing classroom assessments; but we have few roadmaps to help us in building valid and reliable "systems" of local assessments capable of carrying the burden of certification, that is, ensuring valid and reliable decisions at each grade span.

Measured Measures: Technical Considerations for Developing a Local Assessment System, published by the Department's Technical Advisory Committee in June 2000, offered local districts conceptual guidance for initial system design work. The purpose of *this* document, *Local Assessment System Guide: Principles and Criteria*, is to build on the ideas presented in *Measured Measures* and to provide *procedural* guidance on how to build a fully developed framework for local assessment systems that will serve, in particular, the third purpose of assessment, namely the certification of student achievement. (Note: reviewing *Measured Measures* prior to, or while, reading this document is strongly encouraged. It is available at http://www.state.me.us/education/g2000/mm.htm.) *Measured Measures* presented three technical considerations that local assessment systems must address to serve the purposes above: validity, reliability, and standard setting. These three considerations are elaborated on in this Guide, with specific Principles and Criteria for implementing a Local Assessment System consistent with Chapter 127 to allow for valid inferences and reliable decisions. However, a fundamental commitment runs throughout this Guide: to establish *only* those criteria and rules that would be *necessary* to comply with Chapter 127, leaving up to local discretion all other aspects of assessment system design and implementation. The necessary requirements are stated in this Guide as Rules. Options for local decision-making are suggested through Considerations and Examples of local assessment system plans that illustrate, wherever possible, more than one approach to meeting the criteria.

It should also be noted that the Principles and Criteria in this document have been developed to provide a level of technical rigor to meet the needs of Maine's *Learning Results*, not to comply with the recently passed federal "No Child Left Behind Act" (NCLBA). There may, in fact, need to be "connecting points" between our state's requirements and the federal law; however, at this writing the extent and nature of those connections are still emerging (see Conclusion). As work is completed on our state's NCLBA plan, the Department will provide regular updates for local districts designed to clarify any implications for local assessment system design. Though many of the goals of the NCLBA and Maine's *Learning Results* are identical (e.g., high learning expectations for *all* students), key differences exist in the accountability approaches embodied in the two laws, and these differences argue in favor of caution as we examine or develop the points of contact between them.

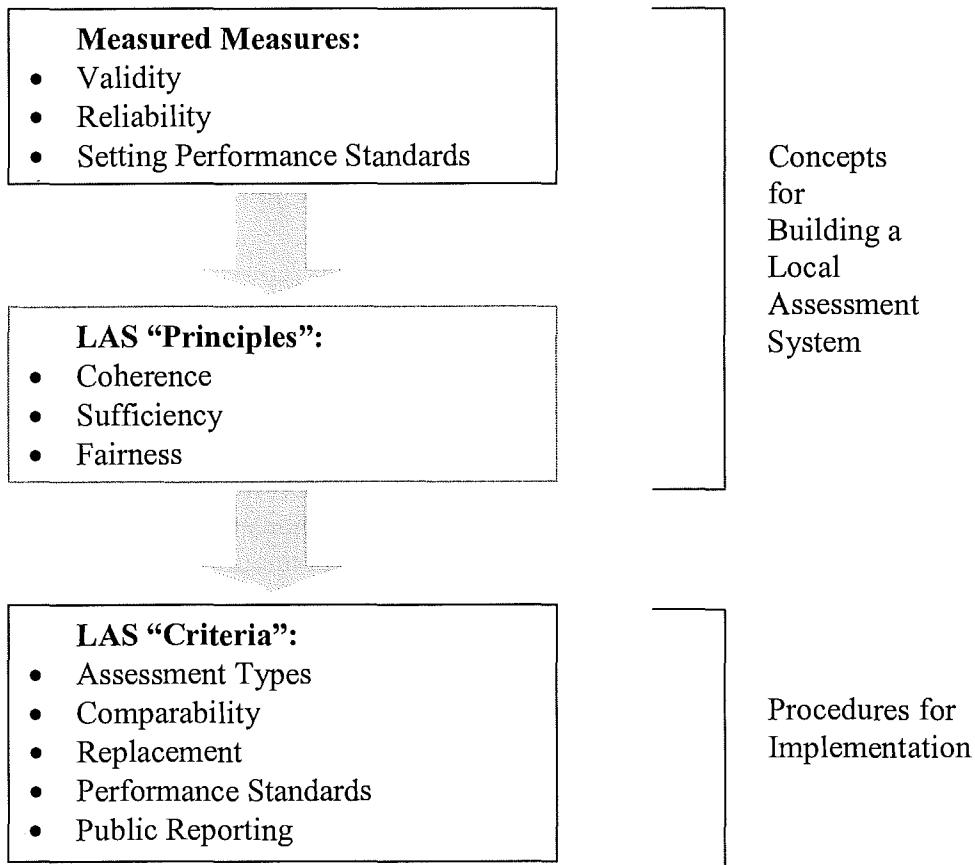# Part I: The Principles of Coherence, Sufficiency, and Fairness

The standards for Local Assessment Systems outlined in Chapter 127 and the criteria for implementing those standards contained in this document derive from three overarching principles: coherence, sufficiency, and fairness. These principles grow out of the design considerations described in *Measured Measures* namely, validity, reliability, and standard setting. To provide valid and reliable information about progress on achievement of the *Learning Results*, a local assessment system, as a component of a comprehensive assessment system, must:

- Be a *coherent* system of assessments aligned with the content standards of the *Learning Results*;
- Provide *sufficient* information to support valid and reliable inferences about achievement of the *Learning Results*; and
- Be *fair* to all students, schools, and school administrative units.

While these principles are not in themselves guidelines for local action, they do form a conceptual foundation that will help the reader interpret and understand the criteria outlined later in this document (see Figure 1).

A Framework: From Conceptual to Procedural

Figure # 1

Measured Measures:
- Validity
- Reliability
- Setting Performance Standards

Concepts for Building a Local Assessment System

LAS "Principles":
- Coherence
- Sufficiency
- Fairness

LAS "Criteria":
- Assessment Types
- Comparability
- Replacement
- Performance Standards
- Public Reporting

Procedures for Implementation

Coherence

Coherence refers to the extent to which the individual assessments that make up the Local Assessment System function in an integrated and balanced fashion within that system to capture both the depth of understanding (or cognitive complexity) and breadth of content knowledge and skills in each content area as defined by Maine's *Learning Results*. Establishing validity—or the match with the specific language of standards in each content area—is at the heart of building a truly coherent system. In addition, coherence refers to the manner in which the components of a local assessment system provide a complete picture of student learning, the extent to which the assessments are **representative** of both the discipline and of the students' skills and knowledge. A coherent local assessment system will include a variety of assessments designed for specific purposes (e.g., informing instruction, certifying performance, and program evaluation.)

Finally, coherence refers to the manner in which a local assessment system is designed to measure student progress on the content standards of the *Learning Results* across time – that is, across grade spans PK-12, as well as across grade levels within grade spans. The system of Maine's *Learning Results* is organized, within each content area, around a system of outcomes with the outcomes increasing in depth and complexity across grades levels PK-12. A coherent local assessment system should be designed from that same broad perspective, not as a set of discrete pieces of information.

Sufficiency

Sufficiency refers to the amount of evidence needed in a local assessment system to allow valid and reliable inferences about achievement of the standards of Maine's *Learning Results* by an individual student, school, or school administrative unit. There must be enough assessment opportunities and enough *variety* in assessment methods to produce consistent or reliable indications of student performance in the content area. As noted in the Introduction, the need for sufficiency in *assessment* must be balanced with the need for sufficiency in *instruction*. Instruction and assessment must work in concert so that the number of assessments administered to an individual student does not detract from necessary instructional opportunities.

The models of assessment systems that accompany this document provide examples of systems designed to meet both aspects of sufficiency when providing information for certification of student performance. As will be seen in those models, and throughout the discussion in this document, concerns about sufficiency will have a major impact on the organization of local assessment systems. The Local Assessment System must be able to provide valid and reliable results about student performance in the content area as a whole (Chapter 127, Section 4.02C10). Chapter 127, states: "This does not require the assessment of each performance indicator specified in [Maine's *Learning Results*]."(Section 4.02C7). The outstanding question, therefore, is: What level of assessment is necessary to produce sufficient information about student performance in the content area as a whole?

The Department of Education has introduced the concept of content "clusters," or groupings of closely related content standards within a content area, to establish an intermediate level of organization of the *Learning Results* so that the answer to the sufficiency question might remain within the parameters of "practical and manageable."

The use of content clusters will be illustrated and discussed in the Criteria section of this document, and in the models of assessment systems that accompany this document. Though both the *Learning Results* and Chapter 127 speak to the significance of providing evidence at the content standard level, the Rules for sufficiency permit "aggregating up" to the level of content clusters for reporting purposes. **Districts may choose, however, to report at the content standard level if local decisions about other aspects of system design would make this level of reporting advantageous to the district.**

Fairness

In *Measured Measures*, fairness is discussed in terms of opportunity to learn, develop, and demonstrate their knowledge and skills, as well as fairness with respect to gender, ethnicity, socioeconomic status, and disability (pp. 15-16). First and foremost, a local assessment system—and the curriculum and instructional systems from which the assessments are derived—must ensure adequate opportunities for students to learn the standards on which they will be assessed. As the accompanying case studies amply demonstrate, building a local assessment system begins with ensuring that all courses or instructional offerings are aligned with the standards. Building an assessment system that includes a purpose as significant as certification—including graduation from high school—requires not only that students

have fair opportunities to learn the standards but that they and their parents are informed and knowledgeable about this component of the system. More so than with the other principles of assessment systems, fairness is indeed a matter for public understanding.

Organizational Framework of Maine's
***Learning Results* (MLR)**

Figure # 2

Guiding Principles:
• Characteristics of a well-educated person

Content Areas:
• MLR levels at which student achievement is certified

Content Clusters:
• Level added to provide school level data for reporting purposes and to keep necessary amount of assessment evidence "practical and manageable."

Content Standards:
• MLR level that is assessed through a sampling of performance indicators

Performance Indicators:
• MLR level that can be measured to provide evidence of achievement of content standards (not all)

Measured Measures also implicitly discusses an aspect of fairness in its discussion of standard setting:

> Where assessment information is intended to serve as part of a comprehensive system, classroom performance standards must be consistent with grade level standards and with school-wide standards. Standard setting must be aggregated up; this is, in the classroom a teacher may make an independent decision but it should be consistent with standards set at the grade level. Teachers at each grade level should talk with each other and come to consensus about what the standard should be across the school. Schools come to consensus across the district, and districts across the state. (p. 28)

(Note: the Department will be conducting a statewide standard setting process to help local districts calibrate their local standard setting processes.)

It is this latter interpretation of fairness—holding all students, schools, and school administrative units to the same standards—that applies directly to the discussion about performance standards for local assessment systems provided in this document.

To support valid and reliable conclusions about student performance within a classroom, school, and school administrative unit, it is essential that all students be measured against the same standards—content standards and performance standards. Further, to support valid and reliable conclusions about school administrative unit performance as part of a statewide comprehensive assessment system, as well as to determine statewide performance (Chapter 127, Section 4.02E), it is essential that all school administrative units be measured against the same content standards and performance standards statewide.

Part II: Criteria

This section of the document presents, and provides rationales for, five Criteria that contribute to the development of a local assessment system that will conform to the principles of coherence, sufficiency, and fairness. In particular, the Criteria are designed to help local educators design, develop, and adopt local assessment systems that are aligned with Maine's *Learning Results* and that support valid and reliable conclusions about student performance in each of the content areas measured.

Organization

The five Criteria presented in this section are:

Criterion 1: Assessment Types – Selection and Distribution
Criterion 2: Comparability
Criterion 3: Replacement
Criterion 4: Performance Standards
Criterion 5: Public Reporting

The order of presentation of the Criteria is intended to begin the initial design of a local assessment system with the selection of individual assessments (Criterion 1), and the administration of assessments (Criteria 2 and 3), then move to the application of performance standards (Criterion 4), and conclude with the aggregation and reporting of assessment results (Criterion 5).

Each Criterion is introduced with a brief description and a list of Rules summarizing the responsibilities local educators have in implementing the Criteria. Local flexibility is provided within the Rules and is addressed in the "Considerations" section intended to support thoughtful implementation of the Criteria. One or more brief examples follow each Criterion for purposes of illustrating important points, followed by a detailed description of the background and rationale for the inclusion of the particular Criterion in this document.

Throughout the Criteria, the assumption is implicit that all assessments must meet the requirements of Chapter 127, Section 4.02D. Each assessment in the system included *for the purpose of certification* must demonstrate validity and reliability. While the Criteria provide guidance for steps that will contribute to the validity and reliability of the system, the system itself relies on the technical quality of individual assessments.

At their core, both the Principles and Criteria involve a combination of complex technical issues and practical issues related to the design, development, and adoption of a local assessment system. As is usually the case with complex issues, there is no single correct and easy-to-implement solution to these issues. It is often necessary to strike a balance among conflicting demands to arrive at the best available solution.

The Criteria, Rules, Considerations, and Examples provide a summary of the key technical issues that will support and inform local educators' decisions.

Criterion 1: Assessment Types - Selection and Distribution

The term "Assessment Type" is used to classify methods of assessment that measure content and skills within a content area at a specified breadth and depth. The distribution of assessment types across a content area is the key to determining whether the system of assessment is coherent, and whether it is sufficient to support valid and reliable conclusions about student performance in the content area as identified by Maine's *Learning Results*. Each school administrative unit will identify a combination of specific assessments for certification that is representative of the character, depth and breadth of the content area.

RULES
To fulfill the requirements of this criterion, each school administrative unit must:

- Assess each content standard for each grade span and content area.

- Include a minimum of 8-12 assessments for each grade span and content area for certification.

- Select a sample of performance indicators based on a method of prioritization. Use Maine's Balance of Representation results (available from the Department) or another method - e.g., the *Measured Measures* procedure or another strategy - to identify high priority or "core" ideas.

- Measure each school level reporting category (content cluster) 5 times.*

- Use a variety of assessment types for each school level reporting category (content cluster)*. Select types based on Form and Function (provided in the *Form and Function Regional Workshop Information* – posted on the DOE website), considering the requirements of selected indicator(s), the developmental levels of students, and the purpose of the assessment (certification).

*Districts must report at the content cluster level at a minimum. They may report at the content standard level if they choose by assessing each content standard a minimum of 5 times.

CONSIDERATIONS
To thoughtfully implement these rules, each school administrative unit should consider:

- Including assessments of both the PK-2 and the 3-4 performance indicators when selecting the 8-12 assessments for the PK-4 grade span.

- Including fewer assessments (within the minimum of 8-12) to accommodate the self-contained classroom structure (a single teacher being responsible for all or many of the content areas).

- Including more assessments at the high school level because of graduation requirements.

EXAMPLE for Criterion 1: Assessment Types – Selection and Distribution

The K-4 Science Committee for SAU #007 met to discuss assessment selection. They had already identified one or more high priority performance indicators for each content standard in their content area and appropriate assessments for each one. Now they are considering the collection of assessments to ensure that it meets all of the requirements of Criterion 1: Assessment Types – Selection and Distribution. The set includes 10 assessments with each content cluster measured at least 5 times (see table below). They have selected a variety of assessment types and are now reviewing to ensure that each assessment conforms to the guidelines of Form and Function.

LAS Example – Science & Technology (K-4)	Types below represent suggested science assessment models - see discipline-specific examples packet		Life Sciences Cluster			Physical Sciences Cluster			Earth and Space Sciences Cluster			Nature and Implications of Science Cluster			
Assessment	Source of assessment	Assessment Type	A	B	C	E	H	I	D	F	G	J	K	L	M
Life Cycle Book(PK-2)	MAP	Structured Response	4						4				4	4	
Melts in the Sun (PK-2)	LAD	Scientific Investigation					4					4	4		
Insects and Me (PK-2)	LAD	Research Project			4								4	4	
Energy Everywhere (PK-2)	LAD	Bundle					4 4								
Food Web (3-4)	LAD	Bundle		4 4											
Science Around Us (3-4)	Local	Scientific Critique				4				4 4				4	4
Soils (3-4)	LAD	Exhibition Assessment (with ELA)								4		4	4	4	
Moving Massive Things (3-4)	Local	Scientific Investigation						4				4		4	
Plot Study (3-4)	MAP	Scientific Investigation	4									4	4		
Earth and Its Moon (3-4)	LAD	Bundle									4				
Total # of Assessments=10		**6 different assessment types**	5 measures of Cluster 1			5 measures of Cluster 2			5 measures of Cluster 3			15 measures of Cluster 4			

Please Note: This illustrative example is intended to clarify Criterion 1: Assessment Types – Selection and Distribution. It should not be considered exemplary and should not be construed as discouraging the selection of locally developed assessments.

Criterion 1: Assessment Types – Selection and Distribution
Background and Rationale

Form Follows Function
Measured Measures

As noted earlier, the term "Assessment Type" is used to classify methods of assessment that measure content and skills within a content area at a specified breadth and depth. Breadth refers to the range of performance indicators assessed (within or across content standards and clusters). Depth refers to the level of cognitive demand indicated by the performance indicators assessed. A coherent and sufficient local assessment system will consist of a collection of different assessment types that is representative of the content standards and their related performance indicators contained in Maine's system of *Learning Results*.

The Selection of Assessment Types

Particular methods of assessment are better suited than others for assessing certain content and skills. *Measured Measures* discussed methods of assessment in terms of the type of response elicited from students. Other references may define them somewhat differently. In any case, a range of types such as those shown below must be part of any system.

Range of Assessment Types

Selected Response	Constructed Response	Performance Tasks	Personal Communication

As explained in *Measured Measures*, some types of assessment are better suited for accommodating the cognitive and behavioral demands of particular performance indicators than others. For example, a selected-response item (e.g., multiple-choice, true-false, matching) might be a more efficient tool than a performance assessment for measuring a student's ability to identify the purpose of the text or to recall factual information (*Measured Measures*, p. 9-11). Other considerations concerning the relationship between assessment format and the who, what, and why of the assessment are provided in *Form and Function Regional Workshop Information*. The requirement of variety within the local assessment system is directly related to the consideration of form serving function.

Because of their precision in defining methods of assessment in terms of the content of the content standards and performance indicators assessed, many assessment types are intrinsic to particular content areas. For example, the demands of the science content standards necessitate an assessment type such as a "scientific investigation". Similarly, the demands of the English language arts content standards inevitably result in an assessment type such as "reading fluency". The particular relationship between the content of a discipline and assessment type is the most important consideration in selecting the combination of assessments to make up the Local Assessment System. To facilitate this process, the Department will prepare Content Area- Specific Assessment Type Definitions and Specifications for each of the five content areas included in the Local Assessment System. (Visual and Performing Arts, Career Preparation and Modern and Classical Languages will be developed on a later time line.)

Standardized norm-referenced tests and other commercial assessments (norm-referenced or criterion-referenced) can play an important role in a local assessment system, although they are not necessary for certification. The loosely defined class of assessments referred to as standardized tests (norm-

referenced achievement tests like the TerraNova, ITBS, Stanford-10 or other commercial assessments) are not classified as separate assessment types in and of themselves. The variation in the content and complexity among those instruments does not permit identification of the standardized test as an assessment type. Each standardized assessment *item* considered for inclusion in the system must be reviewed to ensure alignment with Maine's *Learning Results*.

All students are required to participate in the Maine Educational Assessment through standard administration, administration with accommodations, or through an alternate assessment. For the MEA, the alternate option is the Personalized Alternate Assessment Portfolio (PAAP). None of the avenues of participation for the MEA are considered an assessment type. In addition, neither the Maine Educational Assessment (MEA) nor the Personalized Alternate Assessment Portfolio (PAAP) are considered an assessment type. The Policy Advisory Committee will provide alternative ways to use information obtained from the MEA in conjunction with information obtained from other selected assessment types. Within the Local Assessment System, the same three avenues for participation must be available to students as are available on the MEA. The Department will provide a model that districts may choose to use to meet the alternate assessment requirement.

The Distribution of Assessment Types

> *A collection of assessments does not entail a system*
> *any more than a pile of bricks entails a house.*
> *Measured Measures*

As noted above, the distribution of assessment types across a content area is the key to determining whether the system of assessment is coherent, and whether it is sufficient to support valid and reliable conclusions about student performance in the content area as identified by Maine's *Learning Results*. The fundamental question to address regarding the distribution of assessment types is whether or not the content and cognitive demand required by the collection of assessments is **representative** of the content area as identified by the content standards and their related performance indicators contained in Maine's *Learning Results*. Assuming that the local curriculum is aligned with the *Learning Results* (i.e., students have a full and fair opportunity to learn the concepts, knowledge, and skills articulated in Maine's *Learning Results*), and that the local assessment system is aligned with the local curriculum, then it follows that the distribution of assessments embedded in the curriculum must be aligned with Maine's *Learning Results*.

A second, but equally important, question regarding the distribution of assessment types is whether the collection of assessments is sufficient to support the level of reporting required. Each local assessment system must: a) support valid and reliable conclusions about student performance in the content area as a whole (i.e., students have a full and fair opportunity to learn the concepts, knowledge, and skills articulated in Maine's *Learning Results*); and b) support the reporting of school level performance data on each content cluster. In order to collect sufficient evidence to support school and school administrative unit level reporting at the content cluster level, a minimum of five measures within each cluster is necessary. This does not mean five separate assessment tasks (see Example on page 3). This provides the students five separate opportunities to demonstrate achievement, allows for depth and breadth, and generates a total of 20 (commonly considered an adequate sample for analysis) potential score points (see Example on page 3).

To support valid and reliable conclusions about student performance in the content area as a whole requires a minimum of 8-12 assessments distributed across a variety of assessment types. This is shown in the accompanying Example for Criterion 1, which includes 10 assessments.

Criterion 2: Comparability

Comparability means that similar judgments of student performance in the content area (e.g., Meets the Standard v. Partially Meets the Standard) reflect similar levels of proficiency in terms of content and skills as identified for Maine's *Learning Results*.

RULES
To fulfill the requirements of this criterion, each school administrative unit must:

- Include a minimum of 75% "common" assessments in each grade span and content area (i.e., the same assessment is administered to each student within the school administrative unit at a time when it is instructionally appropriate for the individual student or group of students). (See Figure 3.)

- Follow rules for selection and distribution of types when including classroom assessments (up to 25%) for certification. Classroom assessments *included for certification purposes* must meet the requirements for validity and reliability as described in *Measured Measures*. (See Figure 3.)

- Administer and score assessments at each grade span and at each of the following levels: classroom, school, school administrative unit; and administer state assessments.

- Apply state-established performance standards based on Criterion 4.

- Measure all students against the same standards, providing appropriate accommodations or alternate assessments as necessary.

- Place scores for each performance indicator on a 4-point scale (1-Does Not Meet, 2-Partially Meets, 3-Meets, 4-Exceeds). Scores from assessments that are aligned with performance indicators can be converted to a 4-point scale if they have been scored using another rubric (see the Example for this Rule, below).

- Compare results from classroom, common, state and national assessments (when available) for consistency – similar percentages of students achieving at each level of proficiency.

EXAMPLE for Criterion 2: Comparability

Converting a 6-point rubric to the required performance levels for an individual task:

Performance Levels	DNM	PM	PM	M	M	E
4 Point Rubric Score	1	2	2	3	3	4
6 Point Rubric Score	1	2	3	4	5	6

This example illustrates the conversion of points on a single assessment task. In order to determine performance levels for the content area and make a judgment regarding certification, all scores from the 8-12 assessments must be considered.

Levels of Assessment

Figure # 3
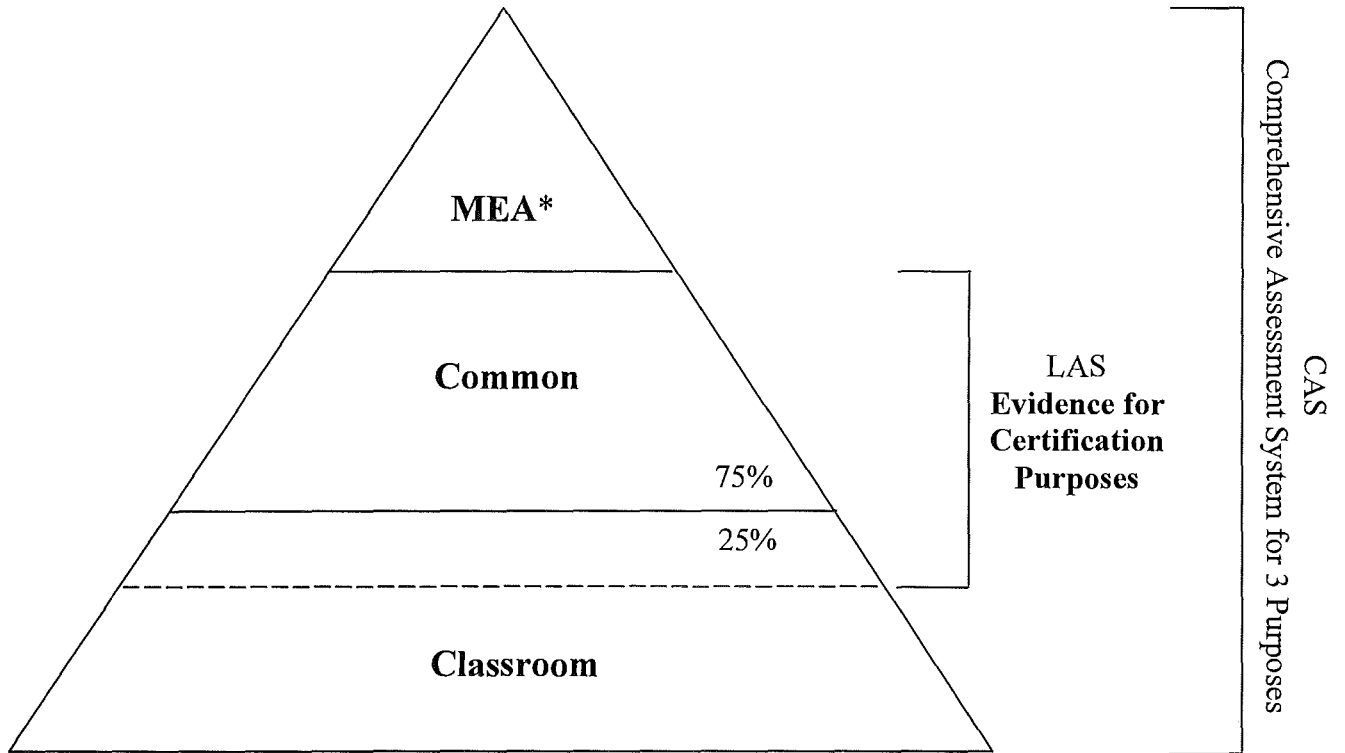
Note: Though Common and Classroom assessments are depicted separately above, common assessments should be integrated into instructional units and administered at an instructionally appropriate time.

* And standardized tests for districts that choose to include them.

Criterion 2: Comparability
Background and Rationale

> *The adoption of common standards and an*
> *accompanying mix of measures which assess learning*
> *is widely regarded as the most important next step in*
> *improving the quality of public education for all*
> *students.*   Measured Measures

Comparability means that similar judgments of student performance in the content area (e.g., Meets the Standard v. Partially Meets the Standard) reflect similar levels of proficiency in terms of content and skills as identified for Maine's *Learning Results*. To produce a local assessment system that is fair for all students, schools, and school administrative units, comparability must apply to the assessment results across students within schools, across schools within school administrative units, and across school administrative units within the State. In short, comparability means that all students must be measured against the same content standard and performance standards.

As discussed previously in this document (see discussion of Fairness in Part 1), the underpinning of comparability is based on the state's adoption of a common set of learning standards for all Maine students (i.e., Maine's *Learning Results*) and a common set of performance standards (See Criterion 4). The performance of all Maine students must be based on those common content standards and measured against the common performance standards.

Comparability within a Local Assessment System

Within a school administrative unit, all students are measured against the same content standards and performance standards identified by the *Learning Results*. The most direct way to ensure that all students are measured against the same standards is to administer the same set of assessments to all students within the school administrative unit. If all students are administered the same set of assessments, one important source of variation in student results (i.e., different assessments) is eliminated.

As a starting point, the Department requires that a minimum of 75% of the assessments in the local assessment system included *for certification of student performance* be common assessments – that is identical assessments administered to all students within the school administrative unit at a time when it is instructionally appropriate for the individual student or group of students. This proportion is set intentionally high to ensure that, even with replacement, the majority of the assessments will be common (see Criterion 3). To provide statistical stability to the local assessment system, there must be a set of assessments that have been administered to all students (i.e., common assessments).

However, for reasons such as variations in instructional programs for particular groups of students, the desire to promote assessment development skills among teachers, or the need to offer replacement assessments to students with weak performance (see Criterion 3), local educators may wish to allow variation in the collection of assessments administered to students within the school administration unit. Therefore, a school administrative unit may choose to have up to 25% of classroom assessments within their collection of assessments. The need for comparability must be balanced with the need for flexibility within the Local Assessment System.

Comparable assessments include the use of appropriate accommodations for students with an Individualized Education Plan (IEP) or 504 Plan and students with limited English proficiency. Assessments that are administered to students with appropriate accommodations or in alternate forms are considered common assessments.

Comparability across Local Assessment Systems

The system of Maine's *Learning Results* was adopted to provide common standards for all Maine students. It follows logically that there should be comparability in the performance standards (see Criterion 4) against which the performance of all Maine students is measured through local assessment systems. Furthermore, in Chapter 127, the Standards for Assessments (Section 4.02D) and presentation of data (Section 4.02E) address the need for local assessment systems to provide data that can be used to provide information about statewide performance.

Criterion 3: Replacement

Replacement is the means by which the assessment system provides the opportunity for a student with low performance to demonstrate an acceptable level of proficiency for certification in a content area through another assessment.

RULES
To fulfill the requirements of this criterion, each school administrative unit must:

- Provide the opportunity for a student who has not met the standard across the 8-12 assessments for a content area, at a grade span, to replace low performance on a specific assessment with a higher result from an another assessment.

- Replace an assessment with another assessment that ensures a variety of types, reflects the belief that form must serve function, and maintains distribution across the content area (each content standard must be measured).

- Determine how many times, and when, a student is allowed to replace each assessment across the system and under what circumstances, based on the purpose of replacement. (See Bullet 1 above.)

CONSIDERATIONS
To thoughtfully implement these rules, each school administrative unit should consider:

- Choosing replacements that address content as close to the content of the original assessment as possible, and at least within the same cluster.

- Developing a process for scheduling and administering original assessments that minimizes the need for replacement assessments while maintaining the efficiency of the system.

- Determining the level of instruction/remediation and ongoing assessment required before allowing a replacement opportunity.

- Determining which assessments measure broad content or process skills (such as a research project or health plan) and may be repeated, and which assessments must be replaced with a completely different assessment. It is not the intent of replacement that students continually retake the same assessment. (See Bullet 3 in the above Rules.)

EXAMPLE for Criterion 3: Replacement

Note: This sample replacement policy is adapted from the Squirrel High School Case Study (see Case Studies Document).

SAU #007 has established the following **policy** regarding replacement for certifying a student for graduation. This policy will be reviewed at the end of the **first** year of implementation. Adjustments will be made, as needed, that respond to balancing concerns about fairness and resources.

1) Students will be provided the opportunity to replace a weaker performance with a stronger performance at the end of every school year if the total number of aggregated points at the end of the third marking period of a given year indicates that the student is "not on track" for graduation.
2) SAU #007 will inform all students and their parents/guardian if the student is "not on track" for graduation at the end of the third marking period of each school year. The notification will indicate the content area, the cluster(s), and the standard(s) in which the performance is low.

Yearly 3rd Quarter Review (Shaded area)

Mathematics	1st Quarter	2nd Quarter	3rd Quarter	4th Quarter
Integrated 1				
Integrated 2				
Integrated 3				

3) The student must receive additional instruction before s/he takes a replacement assessment. (Note: It is the student's responsibility to arrange for additional instructional time with the classroom teacher.)
4) A student may have only two replacement opportunities per assessment, and only if the student is in danger of not meeting certification requirements for graduation.

Replacement of assessments must maintain the distribution of assessment types across the grade span for the content area and follow Form and Function.

Criterion 3: Replacement
Background and Rationale

Inferences are more defensible when students have
multiple opportunities to demonstrate proficiency.
Measured Measures

The goal that all Maine students achieve the standards contained in Maine's *Learning Results* and the consideration of the technical standards required to be met in using assessments for high-stakes decisions involving individual students lead to this conclusion: each school administrative unit must provide, within locally set limits, the opportunity for a student who does not "Meet the Standard" on the Local Assessment System for a content area at a grade span, to replace low performance on a specific assessment with a higher result from an additional assessment. This is true at all grade spans but particularly true at the high school grade span (9-12), where graduation will be linked to the certification of student performance through the Local Assessment System.

To the extent that individual assessments within the collection of assessments for certification are administered at a time when it is instructionally appropriate for the individual student, the likelihood of low performance and the need for replacement assessment should be minimized (see Criterion 2).

Consistent with Criterion 1 (Comparability), the replacement of an assessment with another assessment must ensure distribution across the content standards of the content area, maintain a variety of types, and reflect the belief that form must serve function as adopted by the school administrative unit. That is, to maintain the distribution, an assessment must be replaced with another assessment that addresses content that is as close as possible to the content of the original content of the assessment and is, at least, within the same cluster. Within that structure, there is ample opportunity for replacement assessments to reflect individual learning and assessment styles of students in terms of both presentation and response.

Criterion 4: Performance Standards

Performance standards are the level of student performance across and within the Local Assessment System that provide sufficient evidence that the student "Meets the Standard" in the content area as a whole.

The proposed performance standards below will be validated and adjusted as necessary through a statewide standard setting procedure based on an examination of student work. This procedure will establish the consistency of the performance standards when applied to different collections of student work. The standard setting exercise will also seek to validate the alternative (Pattern of Performance) set of performance standards compared with the recommended set (and the percent of points earned).
Note: Student performance data recorded in terms of a 1-4 point scale can be prepared for the application of either set of performance standards.

RULES
To fulfill the requirements of this criterion, each school administrative unit must:

- Apply either the **recommended** or the **alternative** set of performance standards identified below to certify at each grade span and content area.

Recommended: "Percent of Points Earned"
Students must earn the specified percentage of all possible points for the content area and no fewer than the specified percentage of the available points for any content cluster to achieve each level of performance.

Performance Level	All Points	No Cluster Lower Than
1. Does Not Meet	0-37.4%	
2. Partially Meets	37.5-62.4%	
3. Meets	62.5-87.4%	37.5%
4. Exceeds	87.5-100%	62.5%

Alternative: "Pattern of Performance"
Students must establish a pattern of performance with a specified modal score (mode is the most frequently occurring score) for the content area and no less than the specified mode for any content cluster to achieve each level of performance.

Performance Level	Mode* Across All Assessments	No Cluster Mode Lower Than
1. Does Not Meet	1	1
2. Partially Meets	2	1
3. Meets	3	2
4. Exceeds	4	3

*For bimodal performance (equal number of two different scores), use the mean (average of all scores) of the modal scores.

CONSIDERATION
To thoughtfully implement these rules, each School Administrative Unit should consider:

- Reviewing performance on individual assessments and assessment types at the end of each grade span to analyze and address students' strengths and weaknesses vis-à-vis the performance standards.

Note: These performance standards procedures and their application will be further refined. (See italicized and bolded section on page 13.)

EXAMPLE for Criterion 4: Performance Standards

At the end of her third year of high school science, Alice reviewed her collection of certification assessments to see if she was meeting the standard. If not, she figured she would need to take an additional science course during her senior year. Her scores were as follows:

Content Cluster	Life Science	Physical Science	Earth and Space Science	Scientific Tools of Inquiry and Habits of Mind	ACROSS ASSESSMENTS
Scores	3, 3, 3, 3, 3	2, 3, 2, 2, 4	3, 3, 2, 3, 1	3, 4, 1, 2, 3	
Per Cent of Points	75%	65%	60%	65%	66%
Mode	3	2	3	3	3

Happily, Alice meets standards by demonstrating achievement of 66% (at or above 62.5%) of possible points across assessments and no lower than 60% (above 37.5%) of possible points in any cluster. In the alternative "Pattern of Performance" model, Alice also meets standards with an overall mode of "3" and no cluster model lower than "2".

Criterion 4: Performance Standards
Background and Rationale

How good is good enough?
Measured Measures

What level of student performance across the collection of assessments provides sufficient evidence that the student meets the standards of proficiency for Maine's *Learning Results* in each of the content areas?

The performance standards defined in Criterion 4 will be applied to all students, schools, and school administrative units across the State to ensure fairness (holding everyone and every system against the same standards) and to enable the aggregation of local assessment system results across school administrative units to determine statewide performance (Chapter 127, 4.02D5). Educators already will have established performance standards for each assessment included in the Local Assessment System to indicate the level of a student's achievement of the particular performance indicator(s) measured (Chapter 127, 4.02D5).

The determination of whether a student performance "Meets the Standard" for the content area will be based on the student's body of work across the content area. The definition of performance at the level of proficiency that "Meets the Standard" requires student performance that demonstrates consistent knowledge of content and skills. Both sets of performance standards described above are based on a system where scores for all performance indicators assessed have been placed on a 4 point scale indicating the level of student performance on that indicator (i.e., 1 = Does Not Meet, 2 = Partially Meets, 3 = Meets, 4 = Exceeds) as described in Criterion 2: Comparability.

The recommended method of determining consistency is to review the mean or average of student performance across the content area as a whole, as well as student performance within each content cluster. By this method, a student is determined to meet the performance standard for the content area as a whole by: a) scoring at least 62.5% of the possible points across all of the assessments; **and** b) attaining at least 37.5% of the possible points within each content cluster.

Note that a student who "Meets the Standard" on every assessment will earn 75% of the possible points. The percentages above are based on a calculation of an average of 2.5 points divided by the 4 possible points. Requiring a mean of 3 out of 4 points would establish a performance standard of near-perfect performance. A mean of 2.5 (or 62.5% of all points) allows for a balance of "Partially Meets" and "Meets the Standard" scores.

The alternative method of determining consistency is to examine a Pattern of Performance that a student establishes across the content area as a whole, as well as his or her performance within each content cluster. In this model, a student is determined to meet the performance standard for the content area as a whole by: a) scoring a mode of 3 across all assessments; **and** b) attaining a mode of no lower than 2 on any content cluster. This system is based on the pattern of performance most frequently demonstrated by the student. The mode represents the score or level of achievement the student most often received.   This system also requires a student to demonstrate proficiency across all of the content standards in a content area, and no lower than partial proficiency (a mode of 2) in any content cluster.

Local educators should select the model that is most consistent with the philosophy and design of their assessment system. As data become available from local assessment systems over the course of the next several years, the Department will conduct validation studies to compare results provided by each of the models and will provide additional guidance regarding performance standards. In the meantime, assessments scored or converted to a 4-point scale can use either set of performance standards.

Criterion 5: Public Reporting

"Public Reporting" refers to the processes and procedures used to record and aggregate student performance information related to the Local Assessment System in order to support annual reporting at each grade span for certification of student achievement of the *Learning Results* and school level information about performance at the content cluster level (at a minimum) as required by Chapter 127.

RULES
To fulfill the requirements of this criterion, each school administrative unit must:

- Report annually on the percentage of all students at each performance level for the 5 content areas at the end of each grade span (4th, 8th, and 12th grade).

- Report school level information annually for the 5 content areas at the end of each grade span (4th, 8th, and 12th grade) at a level of greater specificity than content area. At a minimum, the reporting must be at the content cluster level. (See Considerations below, second bullet.)

- Aggregate and disaggregate as necessary to report at the two levels noted above and for identified subgroups.

- Ensure that reporting of results for individual students or groups of students maintains the confidentiality of individual students.

CONSIDERATIONS
To thoughtfully implement these rules, each school administrative unit should consider:

- Ensuring that scores on all assessments reflect student achievement of Maine's *Learning Results* and are not influenced by extraneous factors (e.g., effort, neatness)

- Balancing the desirability and utility of specificity in reporting school level performance information against the resources available to accomplish the corresponding amount of assessment necessary.

- Developing a communication strategy to clarify the levels of reporting, e.g., public reporting, report cards.

EXAMPLE for Criterion 5: Public Reporting

The whole SAU #007 community eagerly opened the morning paper to read about student performance. They expected to read about the percentage of students in each performance level and the identified areas of strength across the content clusters.

TPW COURIER
Area Schools Report on Student Performance

English Language Arts	% of Students who Do Not Meet	% of Students who Partially Meet	% of Students who Meet	% of Students who Exceed	Area of Strength	
PK-4	7%	25%	62%	6%	Cluster 2 Writing & Speaking	73% of possible points
5-8	4%	18%	70%	8%	Cluster 1 Reading & Viewing	67% of possible points

Criterion 5: Public Reporting
Background and Rationale

Individual assessments do not exist in isolation.
Measured Measures

"Public Reporting" refers to the processes and procedures used to record and aggregate student performance information related to the Local Assessment System. Within the Local Assessment System, reporting of student performance for certification occurs at the content area level. Reporting of school level performance occurs at the content cluster level as discussed below.

Reporting Student Performance by Content Area
In order to report student performance across students in the school, the school administrative unit and particular subgroups, content area performance standard results must be aggregated. The process will be straightforward, requiring only the counting of students and calculating the percentage of students in each performance level category (Does Not Meet the Standard, Partially Meets the Standard, Meets the Standard and Exceeds the Standard).

Reporting School and School Administrative Unit Performance by Content Cluster
In order to report school or school administrative unit and particular subgroups information across students in the school, content cluster performance standard results must be aggregated. This will be done according to future guidelines.

A complete discussion of the reporting of results from local assessment systems is beyond the scope of this document. The Department will provide additional guidance in this area at a later time.

CONCLUSION

Areas for Further Investigation and Work to be Accomplished:
Though this document is intended to clarify the next phase of work on local assessment systems, a number of critical technical and policy considerations must be addressed in the coming months:

- It is not clear whether each district in the state has the capacity to complete work on its Local Assessment System within the timelines required in Chapter 127. While this document has been developed on the assumption that building such a system is indeed feasible, it must be remembered that nowhere in the country does such a fully developed system exist. The Department will be gathering additional survey data from districts attending the summer assessment system design institutes, both to inform policy making and to develop appropriate services and strategies to support districts at their individual levels of need.

- It is not clear what student interventions and support systems will be required to ensure that all students meet *Learning Results* standards as the basis for graduation. We will have assessment data to support informed decision making at the local level, but it is not at all apparent whether other aspects of local systems (e.g., extended learning times, personal learning plans, etc.) will be developed on the same timeline. The Department will place a high priority on identifying best practices from around the state and the nation and on strengthening internal communication and dissemination of information on this critical issue.

- The Department is committed to conducting a field study during the 2003-04 school year that would examine all aspects of the implementation and impact of the Comprehensive Assessment System (state and local). Through the field study, the Department will be able to validate the application of these Principles and Criteria across a range of school districts in all regions of the state. In addition, at a more detailed level, the field study will suggest areas where refinements and adjustments could be made to state Criteria, models, case studies used to illustrate the Criteria, and Local Assessment System plans. The commitment to this study reflects the Department's overall commitment to conduct business as a learning organization, and to continue to learn as we go. The Department is also committed to developing additional tools and models to support the evolving focus of this work, including the possible publication of a *Measured Measures II,* if the need exists. Also, the Balance of Representation report, distributed at this spring's Curriculum Coordinators' Conference, will be disseminated widely this summer.

- All students are expected to achieve the standards of Maine's *Learning Results*, including those who attend the vocational-technical schools. Years of effort have gone into linking and aligning the programs at these schools to the *Learning Results*. It is not clear how Maine's vocational-technical schools will be affected by these Principles and Criteria. Remaining questions are whether the vocational-technical school assessments can be fully utilized for certification, or whether students will have a full opportunity to learn the standards that are assessed. It is also not clear how effectively sending schools will be able to capture the

assessment data from vocational-technical schools in support of certification and graduation decisions.

- Additional discussion and guidance is required to help local districts address the issue of diploma requirements for special needs students. At present, Chapter 127 allows for a diploma to be awarded only on the basis of meeting *Learning Results* standards. However, it is not clear how special education students' needs will be met by such a uniform diploma requirement.

- It is not clear how effective the new Maine Educational Data Management Systems (MEDMS) will be in providing a record-keeping template for all local systems. If all systems were identical, problems would be minimal. However, it remains to be seen whether the state system will be flexible enough to allow local district discretion regarding key aspects of their assessment system designs. The Department is committed, however, to developing a record-keeping tool that will have maximum utility for local districts.

- The initial phase of implementation for Local Assessment Systems addresses the content areas of English Language Arts, Mathematics, Science and Technology, Social Studies, and Health and Physical Education. The remaining three content areas—Career Preparation, Modern and Classical Languages, and Visual and Performing Arts—have been tied to Essential Programs and Services (EPS). Future action on developing Local Assessment Systems will be, of necessity, linked to legislative action on EPS. Based on critical actions by the Legislature, the Department will adjust professional development programming, as well as print and on-line resources, to support local work in these three areas.

- As noted above, these Principles and Criteria are being developed at a level of rigor necessary for certifying achievement of Maine's *Learning Results*, not to comply with the NCLB. One area of impact on local curriculum, instruction, and assessment will be the federal requirement to develop grade level expectations (GLEs) in Reading and Math for grades 3-8. The Department will be developing these GLEs through an open, broadly based process during the late summer and early fall of 2003. The Department will also develop an assessment framework for the GLEs that will be designed to have minimal negative effects on local assessment systems. Data on how students are doing against these GLEs must be reported to the United States Department of Education at the end of the 2005-06 school year; this does allow adequate time to phase in both GLEs and the assessment framework. The process of developing the GLEs will begin in the summer of 2003 and it will include ample opportunities for input from local educators. Once the GLEs are completed—including review of draft GLEs by the field—they will be disseminated to all Maine school districts to ensure that local curricula and instructional practices can be reviewed in light of these additional learning expectations. At present, the Maine DOE is carefully reviewing our NCLB accountability plan to be certain that whatever assessments are used to measure achievement of the new GLEs will be consistent with our broad goals for local assessment systems and of Maine's *Learning Results* in general.

A Shared Enterprise:
This document has benefited greatly from the feedback of local educators during the months preceding publication. The Department wishes to thank all those who participated in focus groups or who submitted comments in writing or through e-mail. The development of Maine's Local Assessment System model has been, and will continue to be, a shared endeavor. As such, the Department of Education will continue to welcome input on ways to strengthen this document and to support local districts in their efforts to build local systems in accordance with these guidelines.

The Commissioner of Education's key advisory groups—the Policy Advisory Committee (PAC), Technical Advisory Committee (TAC), and Learning Results Steering Committee (LRSC)—on matters related to assessment and *Learning Results* implementation will continue to monitor all aspects of this long-term effort to build a standards-based system in Maine. As we move forward to ensure Maine students are prepared for life in the 21st Century, we will do so through a system of *shared* accountability—with all stakeholders playing a supportive role as is consistent with the best traditions of Maine life and is, in reality, the only way to reach the vision of Maine's *Learning Results*: high expectations for *all* Maine students.

Appendices

❖ **Department of Education (DOE) Clusters**

❖ **Assessment Contacts**

Proposed DOE Clusters

English Language Arts
1. Reading and Viewing
- A. Process of Reading
- B. Literature and Culture
- D. Informational Texts

2. Writing and Speaking
- E. Process of Writing and Speaking
- F. Standard English Conventions
- G. Stylistic and Rhetorical Aspects of Writing and Speaking

3. Integrated Literacy
- C. Language and Images
- H. Research-Related Writing and Speaking

Health and Physical Education
1. Health Knowledge
- A. Health Concepts
- B. Health Information, Services, and Products
- D. Influences on Health

2. Health Skills
- C. Health Promotion and Risk Reduction
- E. Communication Skills
- F. Decision-Making and Goal Setting

3. Physical Education Knowledge and Skills
- A. Physical Fitness
- B. Motor Skills
- C. Personal and Social Interactions

Mathematics
1. Numbers and Operations
- A. Numbers and Number Sense
- B. Computation
- I. Discrete Mathematics

2. Shape and Size
- E. Geometry
- F. Measurement

3. Mathematical Decision Making
- C. Data Analysis and Statistics
- D. Probability
- J. Mathematical Reasoning

4. Patterns
- G. Patterns, Relations, and Functions
- H. Algebra Concepts
- K. Mathematical Communication

Science and Technology
5. Life Sciences
- A. Classifying Life Forms
- B. Ecology
- C. Cells

6. Physical Sciences
- E. Structure of Matter
- H. Energy
- I. Motion

7. Earth and Space Sciences
- D. Continuity and Change
- F. The Earth
- G. The Universe

8. Nature and Implications of Science
- J. Inquiry & Problem Solving
- K. Scientific Reasoning
- L. Communication
- M. Implications of Science & Technology

Social Studies
1. Economics
- A. Personal and Consumer Economics
- B. Economic Systems of the United States
- C. Comparative Systems
- D. International Trade and Global Interdependence

2. Geography
- A. Skills and Tools
- B. Human Interaction With Environments

3. History
- A. Chronology
- B. Historical Knowledge, Concepts, and Patterns
- C. Historical Inquiry, Analysis, and Interpretation

4. Civics and Government
- A. Rights, Responsibilities, and Participation
- B. Purposes and Types of Government
- C. Fundamental Principles of Government and Constitutions
- D. International Relations

Assessment Contacts

Name	Content Area	E Mail Address	Phone Number
Baker, Mona	Alternate Assessment	Mona.Baker@maine.gov	624-6825
Bachinski, Melanie	Asst- Alternate Assessment	Melanie.Bachinski@maine.gov	624-6678
Byers-Small, Beth	MMSA-Math & Science	bsmall@mmsa.org	794-6986
Cottrell, Marsha	DE Local Assessment	Marsha.Cottrell@maine.gov	624-6783
Cote, Linda	DE Health& Phys. Ed	Linda.Cote@maine.gov	624-6686
Demers, Anne	DE Eng Lang Arts	Anne.Demers@maine.gov	624-6615
Doiron, Diana	English Language Arts	Diana.Doiron@maine.gov	624-6823
Downing, Lora	Career Preparation	Lora.Downing@maine.gov	624-6740
Hatch, Denice	DE Local Assessment	Denice.Hatch@maine.gov	624-6786
Hupp, Dan	Mathematics	Dan.Hupp@maine.gov	624-6827
Johnston, Tad	Mathematics	Tad.Johnston@maine.gov	624-6829
Keller, Tom	Science	Tom.Keller@maine.gov	624-6828
McIntyre, Debra	MMSA-Math & Science	dmcintyre@mmsa.org	338-4201
Manter, Connie	Social Studies	Connie.Manter@maine.gov	624-6824
Maxcy, Brud	MEA	Brud.Maxcy@maine.gov	624-6774
McCue, John	Technology Assistance	John.McCue@maine.gov	624-6790
Monthey, Wanda	Gifted/Talented AP	Wanda.Monthey@maine.gov	624-6831
Morgan, Grace	Health Education	Grace.Morgan@maine.gov	624-6695
Parkin, Linda	DE Alternate Assessment	Linda.Parkin@maine.gov	624-6782
Phillips, Patrick	Deputy Commissioner	Patrick.Phillips@maine.gov	624-6606
Pooler, Jennifer	Project Assistant	Jennifer.Pooler@maine.gov	624-6638
Reutershan, Don	Modern & Classical Lang	Don.Reutershan@maine.gov	624-6826
Rolfe, Pam	Local Assessment Coordinator	Pam.Rolfe@maine.gov	624-6785
Rosenblum, Jill	MMSA- Assessment	jrosenblum@mmsa.org	287-6644
Seaberg, Valerie	Education Team Leader	Valerie.Seaberg@maine.gov	624-6834
Smith, Susan	Assessment- State & Local	Susan.Smith@maine.gov	624-6775
Stivers, John	Career Preparation	John.Stivers@maine.gov	624-6745
Swan, Stephanie	Health Education	Stephanie.Swan@maine.gov	624-6697
Wilbur, Katherine	Health Education	Katherine.Wilbur@maine.gov	624-6696

DE indicates Distinguished Educator

LAS GUIDE

Principles and Criteria
for the Adoption of Local Assessment Systems:

CASE STUDIES

Prepared by the
Maine Department of Education

June 2003

Moose and Squirrel High Schools
Two Case Studies

to illustrate ideas, procedures, & requirements contained in Local Assessment System (LAS) Guide:
Principles and Criteria for the Adoption of Local Assessment Systems

The two case studies that follow are intended for instructional purposes. They are hypothetical, and the high schools that they discuss are fictitious. The case studies are designed to illustrate the application of the Principles, Criteria, Rules, and Considerations defined in the Local Assessment System (LAS) Guide: Principles and Criteria for the Adoption of Local Assessment Systems. Importantly, they demonstrate the fact that there will be a variety of viable Local Assessment Systems. The two examples provided in Moose and Squirrel are different, and yet each, in theory, meets the requirements established for local systems. The hope is that using these cases as companion documents, the Principles and Criteria will be illuminated and better understood.

The case studies should be considered as examples showing the range of possibilities allowed by the Local Assessment System (LAS) Guide: Principles and Criteria for the Adoption of Local Assessment Systems. They should not, however, be construed as exemplary or as perfect models to be exactly imitated. Nor should any instructional or assessment philosophy articulated by either the Moose or Squirrel Schools be considered "the way" according to Maine's Department of Education.

Again, the DOE anticipates that these case studies, and others that may follow, will be useful to those seeking to understand and internalize the Principles and Criteria described in the Local Assessment System (LAS) Guide: Principles and Criteria for the Adoption of Local Assessment Systems.

Local Assessment System
Case Study I:

Certification in Mathematics

Moose High School

Local Assessment System Case Study I - Certification in Mathematics, Moose High School

Moose High School[1] is a 200-student high school; the only high school in Jackson District. Moose High School faculty has a long commitment to developing and using standard-based assessments in their mathematics classes. As a school they have become deeply familiar with Maine's *Learning Results*, *Measured Measures* and different forms of assessment.

They are meeting to develop the 9 - 12 mathematics component of the district's Local Assessment System (LAS) as required under state law and rules. They have studied the Local Assessment System (LAS) Guide: Principles and Criteria for the Adoption of Local Assessment Systems and associated documents issued by the Department of Education, and are making decisions about the system of assessments that needs to be put in place for certification.

As they began this process, they identified the values that have been implicit in the work at Moose High School and the Jackson School District regarding teaching, learning, and assessment of mathematics. They felt that these values should be reflected in both the instruction that students receive and in the design of the LAS.

Identified Values about mathematics curriculum, instruction, and assessment.

1) All courses and course paths should provide students with a full and fair opportunity to learn the concepts, knowledge, and skills in Maine's *Learning Results*, and all students should have a full and fair opportunity to demonstrate their knowledge and skills in relation to Maine's *Learning Results*.
2) While mastery of skills is important, it is not enough. Courses at Moose High School have incorporated a mix of skill and concept development, and application of mathematics through problems and large projects. *All* courses at Moose High School include important aspects of consumer applications and other applications of mathematics.
3) There are multiple purposes for assessment. Therefore, the set of assessments used for certification will be a subset of all the assessment that students will experience in the Moose High Schools Mathematics Program. The department decided to use smaller assessments like the assessment type "bundle" for ongoing classroom assessment, but to rely on other assessments for determining certification.
4) All courses and course paths should provide students with the opportunity to learn, practice and develop the content of Maine's *Learning Results* and be assessed at multiple levels from formative to summative. Information about student progress on standards and related performance indicators should be provided through observations from classroom discussions, quizzes, tests, and projects.
5) Assessment should be embedded in the instructional program.
6) Moose High School's LAS needed to address the multiple purposes of assessment.

---

[1] Fictional school created for this example to illustrate a possible way to deal with the decisions a school/district must make in creating a coherent system, sufficient, and fair local assessment system that will meet the requirements for high school certification.

a) To inform teaching and learning regarding student achievement of Maine's *Learning Results*;
b) To monitor and hold schools accountable for student achievement of the standards; and
c) To certify achievement of standards.

Moose High School's faculty and administrators discussed how these values were consistent with the intent of the MLRs in mathematics, and with the values for assessment as articulated in *Measured Measures* and *Local Assessment System (LAS) Guide: Principles and Criteria for the Adoption of Local Assessment Systems.*

Jackson School District Local Assessment System: Moose High School

The students in the Jackson School District (K –12) participate in a variety of assessments that encompass a variety of purposes. These assessments include ongoing assessments, which are used primarily for classroom purposes, periodic assessments, which are used both for classroom purposes and from which certification decisions will be made, and cyclical assessments, which provide primarily external, standardized (norm-referenced) views of student performance.

Graphic 1:Relationship between Assessment and Purposes – **Jackson School District**

Includes assessments that meet the three purposes of a LAS

Assessment	**Purposes**
Ongoing assessment of MLRs and other local curriculum targets with formative and summative classroom assessments	❑ Feedback for students on their progress toward acquiring skills and knowledge expected in MLRs and other local targets ❑ Grades
Periodic Assessment: Projects and structured response aligned with curriculum used for certification.	❑ Feedback for students on their progress toward acquiring skills and knowledge expected in MLRs and other local targets ❑ Grades ❑ Certification
Cyclical Assessment: MEA Standardized achievement Test	❑ To monitor and hold schools accountable for student achievement of Maine's *Learning Results* ❑ To compare student performance to national norms.

Performance results on all formal assessments are included in grades that students receive at Moose High School. Assessments for certification are a subset of all the assessments that students experience. Maine's *Learning Results* are a subset of the achievement targets for which Moose High School students are responsible.

Graphic 2:Relationship of Assessment for Certification Purposes to All Assessments (Except cyclical)

All assessments (Except cyclical)

All assessments (except cyclical) used for...
❑ Ongoing Instructional Feedback
❑ Grades
A subset of assessment used for...
❑ Certification decisions.

Assessments for certification

Coherence: Curriculum aligned with Maine's *Learning Results*
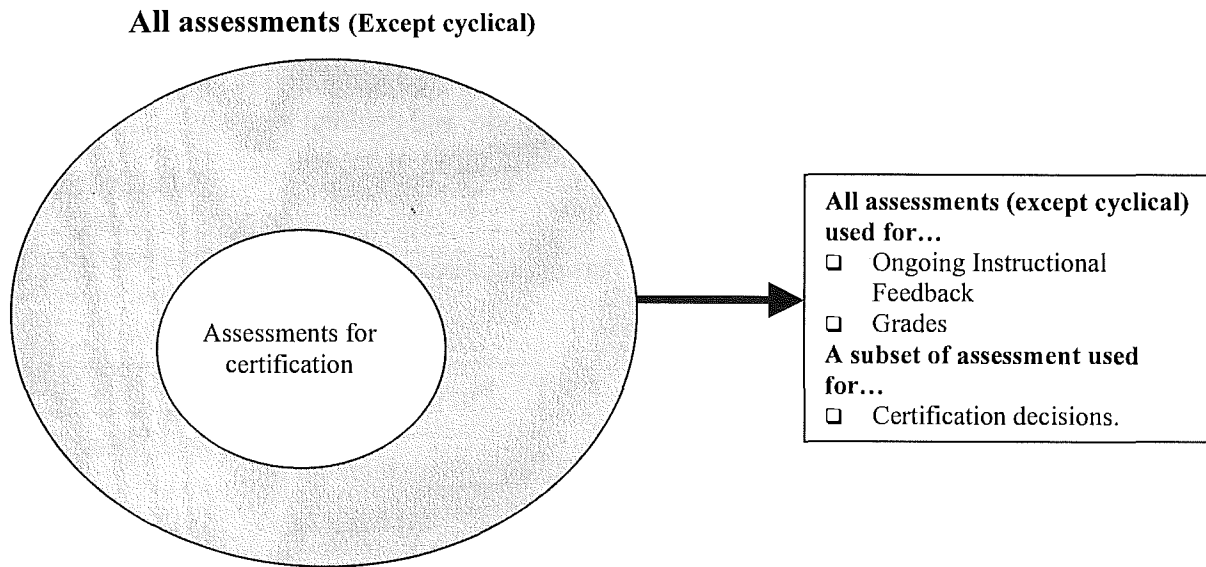
Alignment to Maine's *Leaning Results* Work to Date – To assure that all students had a full and fair opportunity to learn the mathematics concepts, knowledge, and skills articulated in Maine's *Learning Results* at Moose High School, the mathematics faculty completed the following.

1) Moose High School raised the requirement for high school mathematics from 2 mathematics courses to 3;
2) Moose High School reviewed all potential pathways that students could take in mathematics at Moose High School, and then aligned all courses and pathways with Maine's *Learning Results* to assure every student had the opportunity to learn, at minimum, the expectations in Maine's *Learning Results*.

While over 30% of the students at Moose High School complete four years of mathematics, the three paths that students can take to complete the three-year requirement are listed below. An analysis of each of the courses was conducted to assure that students had a full and fair opportunity to learn. Where gaps existed, the courses were modified.

Pathway 1: Algebra I, Geometry, Algebra II

Pathway 2: Integrated Mathematics I, II, and III

Pathway 3: Applied Mathematics I, II, III

To make choices about the assessments to be used in Local Assessment System (LAS) Guide: Principles and Criteria for Implementing LAS, the mathematics department understood what they needed to consider. They reviewed the principles and criteria provided by the MDOE. As a result they identified 5 steps that they needed to complete in order to develop their LAS in mathematics.

1) Review the MDOE Balance of Representation data and decide to either use the data to help prioritize standards and related performance indicators that should be assessed, or make other prioritization decisions.
2) Decide on the assessment types and their distribution that best samples the content clusters, standards, and the related performance indicators consistent with the prioritization.
3) Develop, adapt, or adopt assessments to be used as a part of the assessment system that fulfill the assessment types and distributions.
4) Make decisions about which assessments should be common to all students, and which assessment should not be common.
5) Decide in which courses students would complete the assessments.

Prioritization of Clusters, Standards, and Performance Indicators

Graphic 3 illustrates the relationship between the clusters, standards, and performance indicators for certification purposes. Table 1 provides the data from the statewide Balance of Representation study. The department members understood that the certification decision was at the content area level, not the content cluster or content standard level. They also understood that in order to achieve adequate representation across mathematics to make a reliable and valid decision for certification the assessments that were selected would have to sample every standard at least once, and every content cluster multiple times (at least 5 times). Moose High School's plan assesses each content cluster 5 or more times. It was also important to them to consider Form and Function for all assessments. (Note: This does NOT mean that there are 5 or more separate assessments for each content cluster.)

Graphic 3: Relationship between content area, clusters, and standards for Certification Decision and sufficiency to obtain reliable and valid results

| Mathematics | | Decision for Certification |

Cluster

| Number | Shape and Size | Mathematical Decision Making | Patterns |

Clusters assessed multiple times (at least 5) by a range of assessment types

Standards

A B I   E F   C D J   G H K

Each standard assessed within each cluster at least once for certification purposes

← **Performance Indicators** →

Measurable outcomes: Level at which performance is measured

Prioritization: Prioritized clusters, standards, and performance indicators.

Table 1: Statewide Balance of Representation (BoR) in Mathematics Grades 9 -12

Mathematics										
Number (21%)			Shape and Size (18%)		Mathematical Decision Making (22%)			Patterns (39%)		
A- 7%	B - 6%	I - 9%	E - 11%	F - 7%	C-12%	D - 6%	J - 3%	G - 15%	H - 15%	K - 9%
39%	32%	29%	62%	38%	52%	23%	25%	38%	40%	22%
1 – 50%	1 – 70%	1 – 17%	1 – 29%	1 – 50%	1 – 21%	1 – 51%	1 – 100%	1 – 27%	1 – 27%	1 – 50%
2 – 50%	2 – 30%	2 – 30%	2 – 36%	2 – 50%	2 – 24%	2- 49%		2 – 29%	2- 19%	2 – 50%
		3 – 13%	3- 35%		3 – 20%			3 – 26%	3 – 27%	
		4 – 40%			4 – 21%			4 – 18%	4 – 27%	
					5 – 13%					

BoR across standards

Case Studies I & II
Moose & Squirrel High Schools
Maine Department of Education   June 2003

Prioritization ensures that the "Breadth" of the discipline is assessed

The department reviewed these data and used them to make decisions for prioritizing clusters, standards, and performance indicators for assessment purposes. With the exception of perfor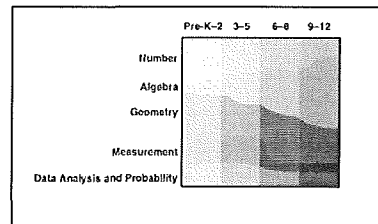mance indicators for standard B.2, I.1, and I.3 the faculty committee decided that the differences at the performance indicators level were not significant. Therefore, all performance indicators are "fair game" for assessment purposes. They also reviewed the Balance of Representation across the standards and found the differences were also not significant. Finally, they reviewed the Balance of Representation across clusters and made the following observations.

1) **Across Clusters: The data indicate that the Patterns cluster should be assessed more** than the other clusters.
2) **Number Cluster:** In general the faculty interpreted the differences between standards A, B, and I as insignificant.
3) **Shape and Size:** Although the BoR for Number, Shape and Size, and Mathematical Decision Making would lead one to believe that assessment should be similar across these clusters, the faculty reviewed national literature that supported more curricular focus on geometry at high school than Number and Mathematical Decision Making.

Local flexibility applied to decide method for prioritization.	The *PSSM* Content Standards Distribution[2] for curriculum	Emphasis purposes.

4) **Mathematical Decisions Making:** When compared with standards D and J standard C should be sampled more heavily.
5) **Patterns:** Emphasis should be placed on standards G and H.

The Jackson School District has adopted the broader definition of reasoning and communication articulated in *Principles of Standards for School Mathematics, 2000*, than standards J and K in Maine's *Learning Results*. As a result, the faculty has embedded standards J and K in each of the mathematics Cluster.

"Reasoning and proof are not special activities reserved for special times or special topics in the curriculum, but should be natural, ongoing part of classroom discussions, no matter what the topic is being studied."
(PSSM 2000, page 342).[3]

Opportunities for Assessing the "Depth" of the Learning Results:

To capture the depth of the mathematics discipline the department members reviewed the mathematics standards and related performance indicators in Maine's *Learning* Results. Since the department members

> **Assessing Depth:** Identified standards and related performance indicators that are best assessed through a large project or investigation.

---

[2] *Principles and Standards for School Mathematics (PSSM)*, NCTM, April 2000.

[3] Ibid.

valued the application of mathematics in larger projects, more realistic of the type of experiences students would have upon graduation, they included projects as an important assessment type.

The first choice they made was related to standard C. While they felt that they could assess each of the performance indicators with small assessments like bundles, they decided that type of assessment would be better as a part of regular classroom assessment in preparation for a larger investigation in which they would be required to address C: 1 – 5, and J. They decided that all students would complete a statistical study during their high school experience. See Appendix A for the Item Specification the department made for Statistical Studies.

Standards E and F offered the opportunity for larger projects as well. They decided that students at Moose High School would complete 1 measurement project and 1 concept project during their high school experience.

The Measurement Project for Pathway 3 included the development of a scale model in which students had to meet certain specifications. The measurement project for Pathway 1 would provide an analysis of a series of formulas to illustrate the relationships among different two and three-dimensional figures. Pathway 2 could choose either the development of a model or the derivation of formulas projects. Although these projects treat different indicators, the faculty are determined that the level of rigor be similar and depth of mathematical understanding be comparable for the projects.

Distribution of Assessment Types

Assessment Types and Distribution: Assessment types and distribution identified

Moose High School faculty and administration decided that the distribution across assessment types would be as follows.

Table 2: Distribution of Assessment Types for Certification

Assessment Type	Number of Assessments
Bundles	None
Structured Response	10
Statistics Study	1
Projects	3
Total Number of Assessments	14

MHS Blueprint for 9 - 12 Math

Because Moose High School has not developed any common district Structured Response assessments, the faculty decided to use the MAP and LAD Assessments available from the MDOE for the Structured Response assessments. However, the faculty is committed to developing the projects locally. Teachers have had experience in the development of projects and have some examples with student work. Appendix A contains draft item specifications for the Projects.

Table 3: Distribution of Assessments (Note: **Bolded** assessments are common to all students. Unbolded assessments are specific to the mathematics pathway. Fours (4) in table indicate the highest level of points possible from each rubric aligned to each performance indicator assessed. "P" indicates that the type of assessment has the potential to sample given performance indicators.) All assessments selected meet the Standards for Assessment (Chapter 127, Section 4.02) with the exception of the projects under development. A plan is put in place to assure that the new assessments meet the Standards for Assessment.

> MHS Blueprint for 9 - 12 Math

*P denotes that students choose a performance indicator from one of these standards for their

		Mathematics																
		Number					Size and Shape				Mathematical Decision Making				Patterns			
		A	B	I	J	K	E	F	J	K	C	D	J	K	G	H	J	K
Total Number of Assessments in Cluster	Type of Assessment	4					6				3				6			
Statistical Study I	Statistics Study										16		4					
Concept Development Project I	Project	P*	P*	P*		4												
Concept Development Project II	Project														P*	P*		4
Measurement Project	Project							4	4	4								
Builders are Us!	SR - Concept						8											
The Softball Set up	SR -						8											
Wire Triangle	SR															8		
The Number Line	SR -	4	4		4	4												
Buying a Jet Ski	SR -														8	4		
Bagels and Donuts	SR -														4	4	4	
Surprise	SR -			4	4	4												
The Deep End	SR -						4	4	4									
Community Growth	SR -			4											4	4		
Probability Booth	SR -											4	4	4				
		36+					40				32				44+			

project.

> **Sufficiency and Reporting:** Each content cluster for Moose High School is assessed at least 5 times. (Minimum requirement for reporting information at the school level is 5 times.)
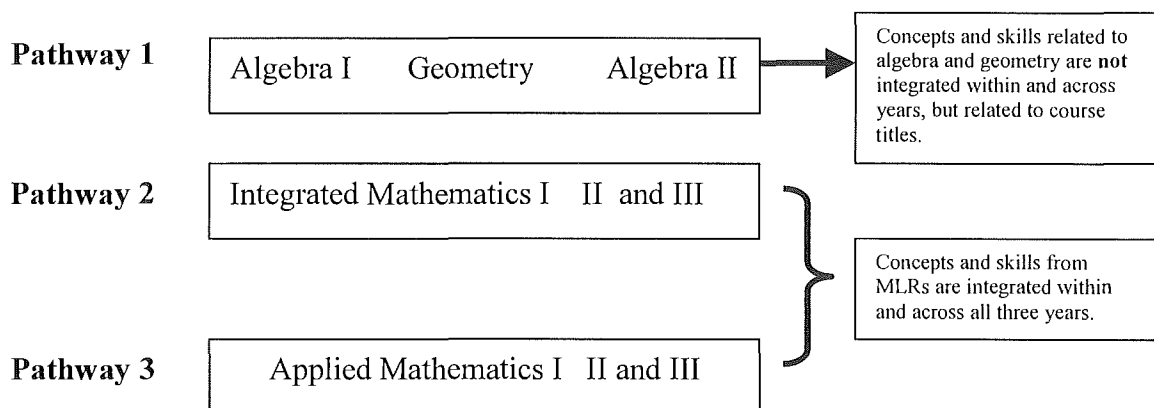
Review

The team reviewed their decision to determine how closely their distribution of projects and constructed response questions met the BoR. The faculty decided that they are satisfied that they have sampled the breadth and depth of clusters and the related standards and performance indicators. The faculty is also satisfied with the degree to which the Balance of Representation (# of assessments) is consistent with their decision to sample Patterns and Shape and Size more heavily than the other content clusters.

Embedding Assessments into Instructional Program

As identified earlier, an important value at Moose High School is embedding assessments into the instructional program. Therefore, the next step for the Moose High School mathematics faculty was to review each mathematics Pathway and embed the assessment activities appropriately.

> **Method of embedding assessments in the instructional program**

Graphic 4: Relationship of Mathematics Pathways and Placement of Assessments

| Pathway 1 | Algebra I   Geometry   Algebra II | Concepts and skills related to algebra and geometry are **not** integrated within and across years, but related to course titles. |

| Pathway 2 | Integrated Mathematics I   II and III | Concepts and skills from MLRs are integrated within and across all three years. |

| Pathway 3 | Applied Mathematics I   II and III | |

Tables 4, 5, and 6 indicate the distribution of the assessments in each of the pathways.

Table 4: Pathway 1 – Distribution of Assessments for Certification into the Instructional Program (common assessments **bold**)

	Integrated I	Integrated II	Integrated III
Statistical Study I	X		
Concept Development Project I		X	
Concept Development Project II			X
Measurement Project	X		
Builders are Us!		X	
The Softball Set up	X		
Wire Triangle			X
The Number Line	X		
Buying a Jet Ski	X		
Bagels and Donuts		X	
Surprise			X
The Deep End		X	
Community Growth			X
Probability Booth	X		

Table 5: Pathway 2 – Distribution of Assessments for Certification in Instructional Program

	Algebra I	Geometry	Algebra II
Statistical Study I	X		
Concept Development Project I		X	
Concept Development Project II			X
Measurement Project		X	
Builders are Us!		X	
The Softball Set up		X	
Wire Triangle			X
The Number Line	X		
Buying a Jet Ski	X		
Bagels and Donuts			X
Surprise			X
The Deep End		X	
Community Growth			X
Probability Booth	X		

Table 6: Pathway 3 – Distribution of Assessments for Certification into the Instructional Program

	Applied I	Applied II	Applied III
Statistical Study I			X
Concept Development Project I	X		
Concept Development Project I			X
Measurement Project	X		
Builders are Us!		X	
The Softball Set up			X
Wire Triangle			X
The Number Line	X		
Buying a Jet Ski		X	
Bagels and Donuts		X	
Surprise			X
The Deep End		x	
Community Growth			X
Probability Booth	x		

Comparability, Replacement, Performance Standards, and Public Reporting

A district wide panel will decide how Jackson School District will implement the criteria for Comparability, Replacement, Performance Standards , and Public Reporting. Members of the Moose High School mathematics faculty will serve on that working group to assure that the unique features of mathematics are represented in the district wide decision-making.

After considerable discussion the mathematics faculty identified the following values as they relate to Comparability, Replacement, Performance Standards, and Public Reporting.

Comparability

The mathematics faculty made the following decisions relating to comparability.

1) All students must be included in local assessments through any combination of standard administration, administration with accommodations or alternate assessments. The distribution and assessment types are the same for all[4] Moose High School students and common assessments taken with accommodations are considered common assessments.
2) 11/14 (79%) of the assessments are common assessments.
3) Decisions have been made about which assessments are repeatable and which must be replaced. (See Replacement #4.)
4) All assessments used for certification have at least 20% of the student papers for each teacher double scored. If the exact scorer agreement is less than 70% on the double scored paper, then all papers will be doubled scored. Disagreements between scores will be resolved with a third scoring.
5) All assessments for certification are selected, adopted or developed to meet the Standards for Assessment in Chapter 127.

Replacement

Moose High School mathematics faculty agrees that students should be allowed to replace weaker performances with stronger performances. However, they believe it is a joint responsibility of the faculty and students to assure that students have the prerequisite skills to be successful on assessments administered for certification purposes. They have decided the following.

1) If students have had the opportunity to learn the prerequisite skills, but have not met the standard for Certification, or are **not** on schedule for meeting the standard for Certification, then they should be able to replace a weaker performance with a stronger performance.
2) Students should *not* have the opportunity **to replace every** weak performance with a stronger performance. Replacement opportunities should be provided *only* if a student is in jeopardy of not graduating or is not on schedule to graduate. This determination is made by the student and her or his advisor at the end of each year.

---

[4] Students with an Individualized Education Plan (IEP) or 504 Plan, and students with limited English proficiency are provided with appropriate accommodations identified in individual educational plans. Decisions about students requiring an alternate assessment will be made by the Special Education Department.

3) Students should not receive more than 2 opportunities to replace a single assessment within a cluster.
4) Replacement of a Statistical Study can be achieved by addressing a new question or claim. Because of the large size of the "Statistical Study" the faculty have decided to design replacement items around scenarios that would let students replace one indicator without redoing the entire study. When more than one indicator is not met students should do a new study to demonstrate their achievement in a valid way. Replacement of a Conceptual Project can be achieved by studying another concept within the same content cluster. Structured response questions are to be replaced by new structured response questions within the same content standard.
5) Students cannot replace an assessment unless they can verify that they have received additional instruction.
6) It is the student's responsibility to schedule additional instructional time either with the classroom instructor or through the Academic Intervention Center at Moose High School.

Performance Standard

Moose High School has adopted the MDOE performance standards Alternative Method.

Alternative: "Pattern of Performance"

Students must establish a pattern of performance with a specified modal score for the content area and no less than the specified mode for any content cluster to achieve each level of performance.

Performance Level	Mode* Across All Assessments	No Cluster Mode Lower Than
1. Does Not Meet	1	1
2. Partially Meets	2	1
3. Meets	3	2
4. Exceeds	4	3

*For bimodal performance, use the mean of the modal scores.

Reporting

Student information is readily available to students and parents on each assessment so they can track progress toward meeting the requirements for Certification. Moose High School faculty provides students regular feedback on their progress toward acquiring the knowledge, concepts, and skills in MLRs and other Moose High School targets as a part of the grading system.

Appendix A: Item Specifications for Assessment Types for Mathematics

Moose High School Projects	Moose High School	Structure	Points Possible
General Description: There are 2 types of projects at Moose High School. ❑ **Concept Development Project** ❑ **Measurement Project** **Replacement:** Projects should be replaced with another project within the same content cluster, or same standard. If the project is the only assessment of a standard then the replacement should be within the same standard in order to maintain the distribution.	**Concept Development Projects** assess Number and Patterns Clusters K2. **Measurement Projects** assess F.2, J, and K in the Shape and Size Cluster,	**Students Interaction:** Students may select the topic or the teacher may assign the topic or project to be completed. Students work independently. **Structure:** Common guidelines are being developed for the three types of projects. **Concept Development Project:** Students identify or are provided with a concept to study. There are two aspects to this project: 1) Fully explaining the concept using models and other representations on a poster and in a written report. 2) Providing examples of how the concept is applied. **Measurement Project:** There are two types of Measurement Projects: 1) **3 – D Model:** Students are asked to make a three-dimensional scale model of an object. The volume and surface area of the object are to be calculated. 2) **Formula derivation:** Students are to demonstrate how surface and volume formulas are derived using models, diagrams, and explanations.	**Points:** **Concept Development:** 8 – 16 points **Measurement Projects:** 12 points

Appendix A: Item Specifications for Assessment Types for Mathematics

Assessment Type Structured Response	Recommended Cluster/Standards/Performance Indicators Assessed	Recommended Structure/Format/Setting	Example MAP or LAD Assessments
General Description: A structured response assessment is defined by students being provided a set of guiding questions and/or formats in which to respond to a topic or problem. To respond to this assessment type, the student does not have to make decisions about the questions that need addressing, or the format in which to respond. The cognitive demand of the question increases across the assessment. A structured response should include both content and process demand either by assessing J and K or by the implied cognitive demand of the performance indicators assessed. **Replacement:** A structured response should be replaced with another structured response that assesses the same standard.	Reasoning (Mathematical Decision Making) and Communication (Patterns) are standards that are features of most structured responses. (The models should deal with distribution.) Computation should be counted or scored only in structured responses where an extensive amount of computation is required.	**Student interaction:** Students respond to a given prompt(s). The work method is prescribed as to which standard will be demonstrated within the response **Structure:** A prompt or set of prompts that describes a problem situation and asks for a student response with a clear expectation of what is expected of students. The item prompts may be scored on process (Reasoning and/or Communication) and one to three content performance indicators. **Interaction of process and content:** Most structured response questions require students to demonstrate achievement in both process and content either through the direct assessment of performance indicators J or K, or through the cognitive demand implied in the performance indicator assessed. **Setting:** Structured response items are treated as a whole. The items usually take more than one class period to complete and may require extended time.	**Examples** Buying a Jet Ski **Points:** Depending on the number of dimensions tested a structured response ranges from 8 (2 dimensions) to 16 (4 dimensions) points

Appendix A: Item Specifications for Assessment Types for Mathematics

Moose High School Statistical Study	Moose High School	Structure	Points Possible
Statistical Study – In a statistical study at the high school level students identify a question or issue to address, design the data collection tools, collect the data, organize and appropriately display the data to address the question or issues. Students analyze the data and draw conclusions based upon the findings. **Replacement:** A statistical study should be replaced by another statistical study that addresses a different topic or question.	**Statistical Studies** assess Standards C and K in the Mathematical Decision Making Cluster. • At grade 9 the Statistical Study will assess performance on 4 performance indicators from standard C depending on the project and J.1.	**Students Interaction:** Students may select the topic or the teacher may assign the topic or project to be completed. Students work independently. **Statistical Study:** Students identify or are provided with a question/claim, devise a study to evaluate the claim/question, collect appropriate data to study the claim/question, appropriately represent data collected, and draw conclusions. The product is a written summary of the question/ claim that was studied. **Setting:** This project would be developed over an extended period of time.	**Points:** **Statistical Study:** 20 points

Local Assessment System
# Case Study II:

Certification in Mathematics

Squirrel High School

Local Assessment System Case Study II: Certification in Mathematics, Squirrel High School

Background: Opportunity to Learn

In 1998 the East Coast School District (K –12) mathematics committee made a number of important curriculum decisions that have a direct impact on the development of the Local Assessment System. The East Coast School District (ECSD) was concerned at that time that the existing curriculum – both paper and "enacted[5]"– was not strongly aligned with Maine's *Learning Results* adopted in 1997. The decision to conduct a study was directly related to assuring that *ECSD students have a full and fair opportunity to learn* the concepts, knowledge, and skills articulated in Maine's *Learning Results.*

The district conducted a yearlong study that involved a gap analysis between the paper curriculum, the materials used by teachers – both textbook and supplemental, and the "enacted" curriculum. The study revealed a large gap between the paper and "enacted" curriculum from classroom to classroom, and even larger gaps between East Coast School District curriculum and expectations in Maine *Learning Results.*
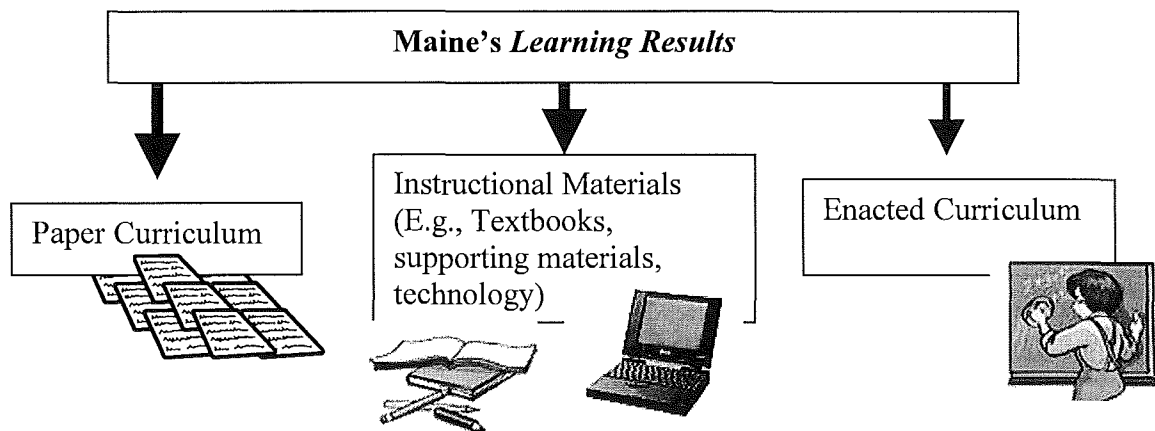

As a result of the study, the district mathematics committee made the following recommendations to the East Coast School District's Board. The ECSD Board subsequently adopted the recommendations.

---

[5] "Enacted" curriculum refers to the curriculum that students received through the instructional program.

1) The district mathematics committee would assure that the expectations in Maine's *Learning Results* were articulated in the local paper curriculum.
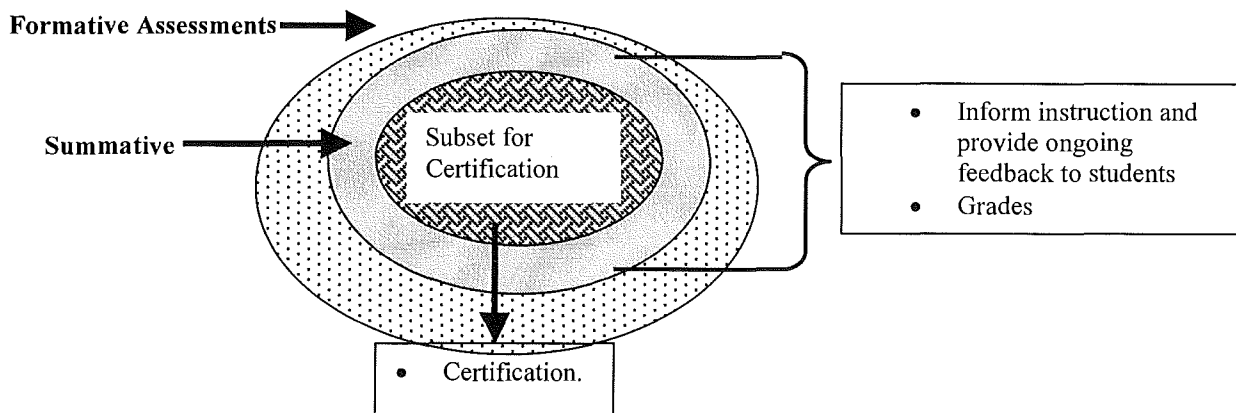
East Coast District Mathematics Curriculum
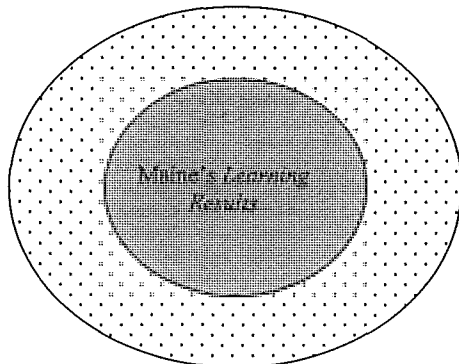
2) The district mathematics committee would recommend mathematics curriculum materials (textbooks and supporting materials).
3) Each school would recommend curriculum materials (textbooks) that were aligned with the expectations in Maine's *Learning Results* and the ECSD curriculum.
4) A system would be put in place to implement new curriculum materials.
5) A system would be put in place to assure that the "enacted" curriculum was consistent with the paper curriculum and instructional materials.
6) At the high school level, ECSD students would be required to complete 3 years of the same mathematics curriculum. *(Note: The high school mathematics faculty decided to implement an integrated[6] mathematics program.)*
7) The committee would put into place a 5-year implementation plan that included the development of the Local Assessment System.

Background: Assessment

In 2001 the District Mathematics Assessment Committee convened to start work on developing assessments – both formative and summative - that would provide information about student learning in relationship to the local curriculum and Maine's *Learning Results*. At that time the high school committee decided the following:
1) Formative assessments would be developed to support continuous feedback on concepts, skills, and knowledge in the ECSD curriculum and Maine's *Learning Results*.

2) Summative assessments would inform instruction, and would be used for grading as well. The thinking at this point was to assume that some subset of the summative assessments and projects would be used for certification purposes and graduation.

---

[6] In an integrated mathematics curriculum related geometry, measurement, algebra, statistics, and probability concepts and skills are interspersed in each year of instruction for all three years, instead of, for example, focusing algebra concepts in a single course.

Formative Assessments →

Summative →

Subset for Certification

- Inform instruction and provide ongoing feedback to students
- Grades

- Certification.

3) To help focus both the curriculum and assessment in relationship to Maine's *Learning Results* the committee conducted a Balance of Representation Study as outlined in *Measured Measures*. (The BoR study is more fully explained below.)

The Present

Squirrel High School[7] is one of two high schools in East Coast School District. Faculty members from both high schools are meeting to develop their Local Assessment System (LAS) as required by state law and rules. Faculty members have studied the "LAS Guide: Principles and Criteria" and associated documents issued by the Department of Education, and are making decisions about the system of assessments that needs to be put in place for certification.

The goal of the meeting is to decide on the assessments and the distribution of the assessments to use for certification at each of the high schools.

The faculty members reviewed the "LAS Guide: Principles and Criteria" provided by the MDOE. From the guidelines, 5 steps were identified for the group to complete in order to develop the ECSD high school LAS for mathematics.

6) Prioritize performance indicators to be assessed for certification.
7) Decide on the assessment types and their distribution that best samples the content clusters, standards, and the related performance indicators consistent with the prioritization process, and Form and Function.
8) Develop, adapt, or adopt assessments to be used as a part of the assessment system that assures the distribution of assessment types.
9) Make decisions about which assessments should be common to all students in the district.
10) Recommend policy for replacement assessments considering comparability issues.

---

[7] Fictional school created for this example to illustrate a possible way to deal with the decision a school/district must make in creating a coherent system, sufficient, and fair local assessment system that will meet the requirements for high school certification

Prioritizing Performance Indicators

The first step in this prioritization was to review the findings from the MDOE Balance of Representation Study conducted in the spring of 2002.

Table 1: Statewide Balance of Representation (BoR) in Mathematics Grades 9 -12

Mathematics										
Number (21%)			Shape and Size (18%)		Mathematical Decision Making (22%)			Patterns (39%)		
A- 7%	B - 6%	I - 9%	E - 11%	F - 7%	C-12%	D - 6%	J - 3%	G - 15%	H - 15%	K - 9%
39%	32%	29%	62%	38%	52%	23%	25%	38%	40%	22%
1 – 50%	1 – 70%	1 – 17%	1 – 29%	1 – 50%	1 – 21%	1 – 51%	1 – 100%	1 – 27%	1 – 27%	1 – 50%
2 – 50%	2 – 30%	2 – 30%	2 – 36%	2 – 50%	2 – 24%	2- 49%		2 – 29%	2- 19%	2 – 50%
		3 – 13%	3- 35%		3 – 20%			3 – 26%	3 – 27%	
		4 – 40%			4 – 21%			4 – 18%	4 – 27%	
					5 – 13%					

(→ BoR across standards)

To prioritize performance indicators for assessment in their LAS, the faculty team reviewed the data from the statewide Balance of Representation conducted by the MDOE, but decided to use data that they had previously generated for Balance of Representation. In June 2001, the East Coast School District mathematics faculty met to determine the "relative importance *(Measured Measures, p. 42)* of each of the performance indicators within each mathematics standard in MLRs. To accomplish this they used the process outlined in Measured Measures on pages 42 and 43.

The decisions found in Table 2 as to which performance indicators must be assessed and which are optional in East Coast School District is the result of this work. The criteria they used to apply the data follow.

Prioritization Criteria for Applying Balance of Representation

1) If a standard or cluster received a higher distribution (greater than 20 points) than the other standards or clusters, then all or most of the performance indicators within the cluster would be required for assessment in the LAS; (E.g., Shape and Size Cluster; Patterns)
2) If the performance indicators received equal weight within a standard and cluster that was not of high priority, then the performance indicators within the standard would be optional, but at least one performance indicator needed to be assessed. (E.g., A. Number of Number Sense)
3) If there were a difference greater than 30% between performance indicators within a standard, then the performance indicators with the greatest weight would be required. (E.g., B.1)

An additional decision made by the team was to require all performance indicators for standards J and K.

Table 2: East Coast School District Performance Indicator Sample for Certification

Standards and Related Performance Indicators	Sampling Rule	Prioritization Criteria Applied for Decision
A. Number and Number Sense 1. Describe the structure of the real number system and identify its appropriate applications and limitations. 2. Explain what complex numbers (real and imaginary) mean and describe some of their many uses.	At least one is required	1
B. Computation 1. Use various techniques to approximate solutions, determine the reasonableness of answers, and justify the results. 2. Explain operations with number systems other than base ten.	B.1 is required.	3
C. Data Analysis and Statistics 1. Determine and evaluate the effect of variables on the results of data collection. 2. Predict and draw conclusions from charts, tables, and graphs that summarize data from practical situations. 3. Demonstrate an understanding of concepts of standard deviation and correlation and how they relate to data analysis. 4. Demonstrate an understanding of the idea of random sampling and recognition of its role in statistical claims and designs for data collection. 5. Revise studies to improve their validity (e.g., in terms of better sampling, better controls, or better data analysis techniques).	C.1; C.2; and C.4 are required, and at least one other	3
D. Probability 1. Find the probability of compound events and make predictions by applying probability theory. 2. Create and interpret probability distributions.	D.1 required	3
E. Geometry 1. Draw coordinate representations of geometric figures and their transformations. 2. Use inductive and deductive reasoning to explore and determine the properties of and relationships among geometric figures. 3. Apply trigonometry to problem situations involving triangles and periodic phenomena.	E.2 required, and one optional	1
F. Measurement 1. Use measurement tools and units appropriately and recognize limitations in the precision of the measurement tools. 2. Derive and use formulas for area, surface area, and volume of many types of figures.	F.1 and F.2 required	1
G. Patterns, Relations, and Functions 1. Create a graph to represent a real-life situation and draw inferences from it. 2. Translate and solve a real-life problem using symbolic language. 3. Model phenomena using a variety of functions (linear, quadratic, exponential, trigonometric, etc.). 4. Identify a variety of situations explained by the same type of function	G.1; G.2; and G.3	1
H. Algebra Concepts 1. Use tables, graphs, and spreadsheets to interpret expressions, equations, and inequalities. 2. Investigate concepts of variation by using equations, graphs, and data collection. 3. Formulate and solve equations and inequalities. 4. Analyze and explain situations using symbolic representations.	All required	1
I. Discrete Mathematics 1. Use linear programming to find optimal solutions to a system. 2. Use networks to find solutions to problems. 3. Apply strategies from game theory to problem-solving situations. 4. Use matrices as tools to interpret and solve problems.	Choice of either I.2 or I.4 required	1 and 2
J. Mathematics Reasoning 1. Analyze situations where more than one logical conclusions can be drawn from data presented.	J.1 required	
K. Communication 1. Restate, create, and use definitions in mathematics to express understanding, classify figures, and determine the truth of a proposition or argument.. 2. Read mathematical presentations of topics within the Learning Results with understanding	K.1 and K.2 required	

Table 3: Relationship Between Cluster, Standards and Required Performance Indicators as Determined by Squirrel High School's Prioritization Process

Mathematics										
Number			Shape and Size		Mathematical Decision Making			Patterns		
A	B	I	E	F	C	D	J	G	H	K
1/2	1/2	1/4	2/3	2/2	3/5	1/2	1/1	3/4	4/4	2/2
3/8			4/5		6/8			9/10		
22/31 Performance Indicators Must be Assessed										

Note: The number of performance indicators reflected in this table does not necessarily reflect the balance of the final distribution of assessments as outlined in Table 5.

Assessment Types and Distribution

The East Coast School District decided to rely heavily on smaller assessments administered frequently within the instructional program over larger assessments. However, the faculty also felt it was important that every East Coast School District student complete one major mathematics project completed during their third year of school. It is the responsibility of the faculty members to assure that this opportunity is provided to every Squirrel High School mathematics student.

In distributing the assessment types, the district has decided to sample every content cluster 10 times instead of five times to be able to provide enough information to report student level information on a cluster.

Table 4: Distribution of Assessment Types

Assessment Type	Number of Assessments	Number Common
Bundles	14	14
Structured Response	8	8
Projects	1	1
	23	23

Table 5: Assessments Distribution Blueprint

Mathematics											
	Number			Shape and Size		Mathematical Decision Making			Patterns		
Assessments	A	B	I	E	F	C	D	J	G	H	K
Bundle 1	4										
Bundle 2				4							
Bindle 3					4						
Bundle 4				8							
Bundle 5					4						
Bundle 6						8					
Bindle 7							4				
Bundle 8									8		
Bundle 9										4	
Bundle 10	4										
Bundle 11			8								
Bundle 12				8							
Bundle 13						8					
Bundle 14	4		4				4				
Buried Treasure				4					4		
The Number Line	4										
Buying a Jet Ski									8	4	
Bagels and Donuts									4	4	4
Surprise	4		4								
Community Growth									4	4	
Probability Booth	4						4	4			
Probability Again							8	4			
Exploratory Project I				12							
Total Points By Clusters	40			44		44			48		
				176							

The bolded assessments (MAP and LAD) have been or are being developed to meet the Standards for Assessment. The faculty members agreed on a process for developing Bundles, and the required project to meet Standards for Assessments as required in Chapter 127.

Maine Educational Assessment (MEA)

The district team members explored ways in which MEA could be included as a part of the Certification decision. At first they wanted to include MEA as an assessment type. However, they read the Guidelines provided by MDOE about standardized assessments.

> Standardized norm-referenced tests and other commercial assessments (norm-referenced or criterion-referenced) can play an important role in a local assessment system, although they are not necessary for certification. The loosely defined class of assessments referred to as standardized tests (norm-referenced achievement tests like the TerraNova, ITBS, Stanford-10 or other commercial assessments) are not classified as separate assessment types in and of themselves. The variation in the content and complexity among those instruments does not permit identification of the standardized test as an assessment type. Each standardized assessment *item* considered for inclusion in the system must be reviewed to ensure alignment with Maine's *Learning Results*.

As the discussion continued they realized that there were other reasons why it would be difficult to include it as an assessment type.

1) There would not be opportunities for replacement;
2) The MEA does not provide the level of timeliness of feedback to inform individual instruction that they wished to have as a feature of their system.

The faculty decided that they could use the MEA as a standard against which to compare their common assessments for consistency.

Comparability, Replacement, Performance Standards, and Public Reporting

A district wide panel will decide how East Coast School District will implement the criteria for Comparability, Replacement, Performance Standards, and Public Reporting. Members of the mathematics faculty will serve on that working group to assure that the unique features of mathematics are represented in the district wide decision-making.

After considerable discussion the faculty identified the following values as they relate to Comparability, Replacement, Performance Standards, and Public Reporting.

Comparability

The mathematics faculty made the following recommendations relating to comparability.

6) All students are included in local assessments through any combination of standard administration, administration with accommodation, or by alternate assessments. The distribution of assessment types is the same for all[8] ECSD students.
7) Since all students take the same 3-year sequence *all assessments are common.*
8) Item Specifications for each assessment type have been developed for each assessment type.
9) All assessments used for certification are on a 4-point scale. Each 4 – point scale is linked to a specific performance indicator.
10) All assessments used for certification have at least 20% of the student papers for double scored. Prior to scoring student work for certification, all teachers will have been trained to score to a rubric, and then qualify to score using prescored papers.
11) All assessments for certification are selected, adopted or developed to meet the Standards for Assessment.

---

[8] Students with an Individualized Education Plan (IEP) or 504 Plan, and students with limited English proficiency are provided with appropriate accommodations identified in Individual Educational Plans.

Replacement:

ECSD Faculty recommends the following guidelines regarding replacement for certifying a student for graduation. These gu8delines will be reviewed at the end of each year of implementation. Adjustments will be made, as needed, that respond to balancing concerns about fairness and resources.

1) Students will be provided the opportunity to replace a weaker performance with a stronger performance at the end of every school year if the total number of aggregated points at the end of the third marking period of a given year indicates that the student is not on schedule for graduation. Not on schedule means that the "mean" performance based upon assessments taken to date is less than the mean expected for graduation.
2) ECDS will inform all students and their parents/guardian if the student is not on schedule for graduation at the end of the third marking period of each school year. The notification will indicate the content area, the cluster(s), and the standard(s) in which the performance is low.

Table 6: Yearly 3rd Quarter Review (Shaded area)

	1st Quarter	2nd Quarter	3rd Quarter	4th Quarter
Integrated 1				
Integrated 2				
Integrated 3				

3) The student must receive additional instruction before they take a replacement assessment. (Note: It is the student's responsibility to arrange for additional instructional time with the classroom teacher.)
4) A student may have only two replacement opportunities per assessment, and only if the student is in danger of not meeting certification requirements graduation.
5) Replacement of assessments must maintain the distribution of assessment types across the grade span and follow Form and Function.

District mathematics faculty agrees that students should be allowed to replace weaker performances with stronger performances only if they are in jeopardy of not graduating. They think there should be a limit on the number of times (3) that a student is allowed to replace an assessment for the following reasons.

1) There are limited available resources – both replacement assessments and instruction.
2) Instructional time – The need to move on to additional topics for students to progress;
3) There are a number of assessments in the system that provide other opportunities to assess the same content standards and related standards and performance indicators.

The district team believes that Replacement should happen at the standard level.

Reporting

The district mathematics faculty agrees the decision for Certification should be made at the content level. To meet the criterion for reporting school level information at the cluster level, the district elected to assess a cluster at least ten times, understanding that the minimum requirement is five times.

Standard Setting

ECSD has recommends adopting the MDOE performance standards, "Percent of Points Earned".

Recommended: "Percent of Points Earned"

Students must earn the specified percentage of all possible points for the content area and no fewer than the specified percentage of the available points for any content cluster to achieve each level of performance.

Performance Level	All Points	No Cluster Lower Than
1. Does Not Meet	0-37.4%	
2. Partially Meets	37.5-62.4%	
3. Meets	62.5-87.4%	37.5%
4. Exceeds	87.5%-100%	62.5%

Appendix A: DRAFT Item Specifications for Assessment Types

Assessment Type Structured Response	Recommended Cluster/Standards/Performance Indicators Assessed	Recommended Structure/Format/Setting	Example MAP or LAD Assessments
General Description: **A structured response assessment is defined by students being provided a set of guiding questions and/or formats in which to respond to a topic or problem. To respond to this assessment type, the student does not have to make decisions about the questions that need addressing, or the format in which to respond.** **Each structured response should assess content and process either by assessing standard J or K or through the implied cognitive demand of the performance indicator assessed.** Replacement: **A structured response should be replaced with another structured response that assesses the same standard.**	All clusters, standards and performance indicators	**Student interaction:** Same as general description. **Structure:** A prompt or set of prompts that describes a problem situation and asks for a student response with a clear expectation of what is expected of students. **Setting:** Structured response items are treated as a whole. The items take at least one class period to complete and may require extended time.	**Examples** Buying a Jet Ski **Points:** Depending on the number of dimensions tested a structured response ranges from 8 (2 dimensions) to 16 (4 dimensions) points

Assessment Type: Projects		Structure	Points Possible
General Description: **There are 2 types of projects at Moose High School.** ❑ **Concept Development Project** Replacement: **Projects should be replaced with another project within the same content cluster, or same standard. If the project is the only assessment of a standard then the replacement should be within the same standard in order to maintain the distribution.**	**Concept Development Projects E.2, K, and L.**	**Students Interaction:** Students may select the topic or the teacher may assign the topic or project to be completed. Students work independently. **Structure:** Common guidelines are being developed for the three types of projects. **Concept Development Project:** Students identify or are provided with a concept to study. There are two aspects to this project: 3) Fully explaining the concept using models and other representations. (E.g. written summary, posters, other) 4) Providing examples of how the concept is applied.	**Points:** **Concept Development: 20** points

Assessment Type: Bundle	Recommended Cluster/Standards/Performance Indicators Assessed	Recommended Structure/Format/Setting	Example MAP or LAD Assessments
General Description: **A bundle is a set of selected response, short answer, or short constructed response questions that assess a single performance indicator that has multiple components, or multiple related performance indicators.** Replacement: **A weak performance on a bundle should be replaced by Bundle in the same standard.**	**Mathematics**: All clusters but not J and K.	**Students Interaction:** Student responds to prompt. **Structure:** • Each component in a bundle should include a series of questions that relate to a single performance indicator or related performance indicators. • The format of the questions could include a combination of selected response, short answer, and constructed response but should not be limited to selected response. **Interaction of process and content:** Bundles assess only content performance indicators. **Setting:** Bundles should be administered within the class setting.	High School ❏ Down Below ❏ Ode to a Fraction, ❏ Which Base are bundles?