

## ***Design Principles for Virginia's Assessment System***

***Revised: August 19, 2015***

**Purpose:** *Utilizing research and best practices related to assessment, the Assessment 2.0 Subcommittee will work collaboratively to develop recommendations for the future of assessment in Virginia. The subcommittee will work to develop an assessment philosophy and model that supports effective teaching and learning in Virginia's classrooms.*

What do we want to measure?

- Literacy and numeracy at the elementary and middle school level
- Reading for different purposes
- Oral and written communication skills
- Civic responsibility
- Scientific literacy and an understanding of the scientific process
- Collaborative/social skills
- Self-awareness/agency
- Critical thinking/problem solving
- Financial literacy
- Ability to solve real-world problems
- De-emphasis on "Google-able" content with a focus on application
- Ability to create new ideas/solutions rather than simply locating information
- Life-ready skills (college, career, and citizenship skills)
- Ability to identify accuracy and bias
- Ability to "curate" available information

What are our beliefs related to effective assessment and an assessment system?

- Assessment should "drive" instruction.
- Assessment and instruction should be closely "linked together."
- The assessment system should measure students on a "lean" curriculum that only focuses on the most important standards. This allows for more depth rather than surface learning or memorization.
- Reading and mathematics assessments should be vertically aligned in order to measure student growth over time.
- Assessment should "engage" students in active, meaningful learning.
- Assessments should not be tests that students can "cram" for and still be successful.
- Assessment should be straight-forward for students and should include clear learning targets.
- Assessments should provide meaningful data in a timely manner.
- Diagnostic and pre-assessment should be part of the system.
- The assessment system should provide student "choice" where appropriate.

- Only the most valuable content and skills should be assessed.
- Flexibility should be a part of the overall assessment system in Virginia.

What are “design principles” that must be included in Virginia’s New Assessment System?

<b>Design Component</b>	<b>Notes</b>
Tested Population	<ul style="list-style-type: none"> <li>• There should be a balance of “census” testing vs. sampling.</li> </ul>
Amount of Testing	<ul style="list-style-type: none"> <li>• The amount and frequency of testing must be considered. Are there additional areas where state testing can be reduced or eliminated?</li> </ul>
Local Flexibility	<ul style="list-style-type: none"> <li>• There is a need to clearly define local testing requirements vs. state testing requirements.</li> <li>• There is a need to remove barriers to allow localities to create innovative assessment practices at the local level.</li> </ul>
Assessment Format	<ul style="list-style-type: none"> <li>• Project-based and performance assessments should be a part of the new system.</li> <li>• The assessments should not be primarily multiple choice assessments. In order to measure certain skills, performance tasks are necessary.</li> </ul>
Equity	<ul style="list-style-type: none"> <li>• All school divisions should have access to the same resources in order to increase equity in assessment practices across the state.</li> <li>• Assessment plans for ELL and SWD should be well defined.</li> </ul>
Teacher Involvement	<ul style="list-style-type: none"> <li>• Teachers must be actively involved in the state assessment system.</li> <li>• Training and resources should be readily available to support teachers and administrators with the new assessment system. This is a priority.</li> <li>• Teachers must be actively involved in the design, development, and scoring of statewide assessments.</li> </ul>
Assessed Content	<ul style="list-style-type: none"> <li>• The assessment system should focus on what graduates can do, not just what they know.</li> <li>• There should be a focus on assessing essential skills through content (application).</li> <li>• The assessments should measure important skills like critical thinking and problem-solving as well as higher levels of thinking (DOK and Bloom’s)</li> <li>• Assessments should measure a lean curriculum at a deeper level rather than too many standards at the surface level.</li> <li>• Students should be assessed on reading and communication skills at key grade levels.</li> </ul>

	<ul style="list-style-type: none"> <li>• Clear targets should be set for the elementary, middle, and high school level.</li> <li>• It will be important to define and communicate “what” is being assessed and “why” these areas are being measured.</li> <li>• Where appropriate, state tests should integrate more than one subject or content area. For example, writing can be assessed on a content area assessment.</li> </ul>
Timing of the Test	<ul style="list-style-type: none"> <li>• High school students should have the option to meet graduation requirements earlier, if appropriate.</li> <li>• Personalization and individualization should be included in the system, as appropriate.</li> <li>• “On-Demand” testing should be a component to the assessment system.</li> </ul>
Student Growth	<ul style="list-style-type: none"> <li>• Student growth in literacy and mathematics should be measured K-8 (for diagnostic <i>and</i> accountability purposes).</li> <li>• A vertical scale for reading and mathematics K-8 must be developed in order to effectively measure student growth in these areas.</li> </ul>
High School Flexibility	<ul style="list-style-type: none"> <li>• Credentialing should be a priority at the high school level.</li> <li>• Several options for high school should be explored to provide opportunities for students to meet graduation requirements (examples: credentials, portfolio assessment, college credit, and apprenticeships).</li> </ul>
Developmentally Appropriateness	<ul style="list-style-type: none"> <li>• Reading assessments should be developmentally appropriate (learning to read vs. reading to learn).</li> <li>• Elementary assessments should be shortened to be more appropriate for young students.</li> <li>• Reducing the test length for students at <u>all</u> levels must be considered.</li> <li>• Language and vocabulary used on the assessment must represent the appropriate reading level for the grade level being assessed.</li> </ul>
Scoring	<ul style="list-style-type: none"> <li>• If TEI items continue to be utilized, students should receive partial credit for correct responses, even if they do not get all items for a particular question correct.</li> </ul>

	<ul style="list-style-type: none"> <li>• All TEI items need to be developmentally appropriate and closely aligned to the content area being assessed.</li> </ul>
Student Choice	<ul style="list-style-type: none"> <li>• Student choice should be a part of the assessment system as appropriate. For example, students can have an option to choose among more than one writing prompt or select a task to complete from 2-3 performance task options.</li> </ul>
Reporting of Assessment Results	<ul style="list-style-type: none"> <li>• A clear, understandable interface needs to be developed to report assessment results in a timely manner.</li> <li>• Teachers need immediate access to these results in order to provide remediation in a timely manner.</li> </ul>
Test Development	<ul style="list-style-type: none"> <li>• Teachers need to be actively involved in test development with limited involvement from outside testing companies.</li> <li>• An evaluation process must be put in place to determine if the revised assessment system is meeting the intended goals.</li> <li>• All assessments should assess what the curriculum framework requires, rather than assessing standards that go beyond the state framework (alignment).</li> </ul>

Updated May 19, 2015

## **A Glossary of Assessment Terms**

Revised 5-19-2015

**Accountability Systems**—The mechanisms used (generally by states) to evaluate the performance of their education systems.

**Alternative (or alternate) Assessments**—Alternative assessments are used primarily to determine what students can and cannot *do*, in contrast to what they do or do not *know*. In other words, an alternative assessment measures applied proficiency more than it measures knowledge. There are multiple types of alternative assessments, of which performance assessment is one.

**Authentic Assessments**—An alternative assessment that asks students to perform real-world tasks. The student will typically have to employ critical thinking and problem-solving skills to successfully address the challenge presented. The more authentic an assessment task is, the more closely it approximates the way a similar task would be done in a setting outside the classroom (a workplace or community for example). Student performance on a task is typically scored on the basis of a list of desired outcomes (known as a rubric).

**Balanced Assessment System**—An assessment system that employs multiple types of assessments so that:

- (1) achievement and growth are taken into account;
- (2) assessments are matched to learning goals (both core content mastery and skills for success in the modern world); and
- (3) the need for accountability measures is met, but not at the expense of meaningful information that informs classroom instruction.

**Computer Adaptive Testing (CAT)**—A test taken using a computer program that adjusts the difficulty of test questions based on student responses.

**Common Diagnostic Items**—A collection of assessment items that can be utilized to pre-assess students on specific content and skills.

**Criterion-Referenced Assessment**—A test or assessment that is designed to measure students against a fixed set of predetermined criteria or learning standards. In K-12 education, these are typically aligned to the curriculum taught in a particular course, academic program, or specific content area.

**Developmental Appropriate Assessment**—An assessment that respects the age and individual needs of the student who is being assessed. This may include areas such as the intellectual, social, emotional, and physical needs of the particular age group being assessed.

**Formative Assessment**—The overall goal for formative assessment is to collect detailed information in an informal manner that can be used to improve instruction and student

Updated May 19, 2015

learning during the learning process. It is used to “inform” the learning process so that the teacher can make in-process adjustments and learning modifications based on the data collected from students. Formative assessments are generally referred to as assessments “for” learning.

**Higher-order thinking**—A category of thinking skills that increase the cognitive load, requiring students to go beyond understanding content and replicating skills. Students employing higher-order thinking may make connections, solve problems different from those given in classroom examples, and use content to reach and justify conclusions.

**Integrated or Interdisciplinary Assessments**—An assessment that measures student performance on content and/or skills across content areas.

**Inter-rater Reliability**—The degree of agreement among raters who are tasked with scoring a performance task or product.

**Norm-Referenced Assessment**—A standardized test that is designed to compare and rank students (test takers) in relation to other students who participated in the assessment. Norm-referenced tests report how a particular test taker performed in comparison to the hypothetical average student, which is determined by comparing scores against the performance results of a statistically selected group of test takers, typically from the same age group and grade level, who have already taken the assessment.

**“On Demand” Testing**—This type of testing allows flexibility to assess students when they are ready to be tested on required content rather than testing students according to an established testing window that does not account for student readiness.

**Performance Assessment** – A type of alternative assessment in which students demonstrate the use of their acquired knowledge and skill. Performance assessments typically include, but are not limited to, exhibitions, investigations, demonstrations, written or oral responses, journals, and portfolios. Performance assessments are typically scored using rubrics (see Rubrics), which explicitly describe levels of performance and designate which levels meet standards.

**Personalized Learning**—An approach in which students’ individualized learning needs are the primary consideration in making instructional decisions.

**Portfolio Assessments**—A type of assessment that is a systematic collection of student work and artifacts that demonstrate mastery of course and/or subject knowledge and skills.

**Problem-Based Learning**—Problem-based learning or PrBL is a form of inquiry-based instruction that is frequently utilized in mathematics to engage students in the investigation and resolution of real-world problems.

**Project-Based Learning/Assessments**—Project-based learning or PBL is a teaching

Updated May 19, 2015

approach that engages students in sustained, collaborative real-world investigations. Projects are organized around a driving question, and students participate in a variety of hands-on tasks that seek to meaningfully address this question (Buck Institute).

**Reliability** – The consistency or stability of test performance. Tests must be constructed and administered so that measurement error (for example, from ambiguous scoring, unclear questions/directions, or environmental factors) is minimized.

**Rubric**—A description of the criteria for success and levels of achievement for a task. Rubrics are used during instruction to help students maximize and improve the quality of their work, and as scoring tools for multiple types of alternative assessments (see Performance Assessment).

**Student Achievement**—Student demonstrated mastery of certain knowledge and/or skills as measured by a particular assessment.

**Student Growth**—The change in student achievement for an individual student between two or more points in time.

**Student Portfolios**—A type of assessment that is a collection of student work and artifacts that demonstrate mastery of course and/or subject knowledge and skills. The collection should include evidence of student reflection and self-evaluation, guidelines for selecting the portfolio contents, and criteria for judging the quality of the work included in the portfolio (Venn, 2000, pp. 530-531).

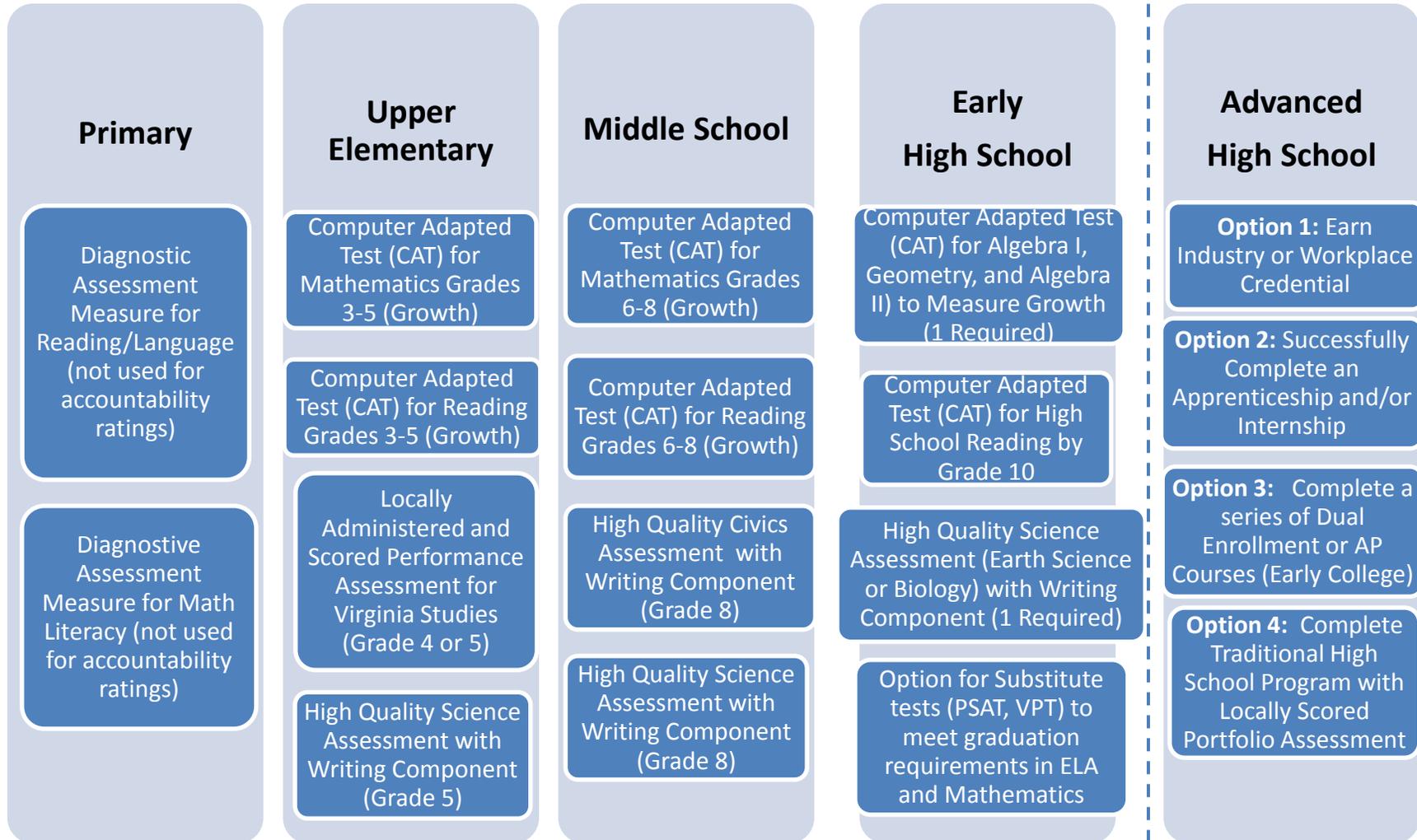
**Summative Assessment**—Assessments that are used to evaluate student learning, skill acquisition, and academic achievement at the conclusion of a defined instructional period. Summative assessments typically are administered at the end of a unit, project, course, semester, program, or school year. These are frequently described as assessments “of” learning.

**Test Equating**—The statistical process used to determine comparable scores for different forms of a particular test or assessment. It is intended to ensure that scores from different forms of the test are interchangeable. The equating process adjusts for different levels of difficulty across test forms.

**Validity** – The degree to which an assessment actually measures the learning it is intended to measure. Assessment designers use tools—both design and statistical tools—to maximize and collect evidence of assessment validity.

**Proposed Assessment Model pre-K through Graduation**

*Revised: August 19, 2015*



\*It will be necessary to refer to the work of the Accountability 2.0 Subcommittee to determine how these assessments factor into accreditation ratings for schools.

\*\*In order to fully implement these recommendations, significant changes must be made to the state assessment and accountability system.

\*\*\*The high school options will require a strong partnership with community colleges and post-secondary institutions across the state.

