

Strengthening and Streamlining Vermont Local Comprehensive Assessment Systems: Defining Essential Components

September 1, 2022



Introduction

A [local comprehensive assessment system](#) includes multiple measures across all content areas, for various purposes, and data that should be used to inform instructional and programmatic decisions. These measures include screening, benchmark, diagnostic, formative, progress monitoring, and summative assessments as well as state assessments.

Different types of assessments can be used to gather evidence of student learning, and, generally, no one piece of assessment information can fulfill all purposes. Assessment systems should be inclusive of all students, which means that some students may require accommodations to access the assessment.

Local Comprehensive Assessment Systems, one of the four levers related to Act 173, should be designed to provide parents, guardians, educators, and pediatricians with essential information that can follow a child from infancy through twelfth grade and beyond. The data that are collected enable caregivers and educators to determine how to provide learning opportunities and support that help each child to thrive.

Purpose

This document was created from a larger piece of work, Strengthening and Streamlining Local Comprehensive Assessment Systems, in order to develop a shared understanding of assessment terminology that is commonly used.

Types of Assessments

The type and purpose of the assessment dictate the frequency of its administration. “Done well and thoughtfully, assessments are tools for learning and promoting equity... Done poorly, in excess, or without a clear purpose, they take valuable time away from teaching and learning...” ([USDOE, 2015](#)).

Performance Assessments and Tasks

Performance assessments are any teacher- or student-designed learning activity or investigation in which students demonstrate their knowledge, understanding, and/or skills through a performance task. Performance assessments engage students in meaningful learning in authentic contexts, show genuine applications of knowledge, and yield a tangible product and/or performance that serves as evidence of learning.

A performance task is what students actually do within a larger performance assessment framework. Tasks built around student interests engage students and help them make connections to their personal lives.

Project-Based Learning is one method of teaching that incorporates a series of performance assessments and learning opportunities that lead to a culminating product or performance for an audience. The Agency of Education document, [Project-Based Learning: A Student-Centered Approach](#), explains:

According to [PBL Works](#), Project-Based Learning “unleashes a contagious, creative energy among students and teachers,” strengthening relationships and creating a life-long love of learning. Students gain academic skills and content knowledge while

honing critical thinking, collaboration, creativity, communication, and other transferable skills. When done correctly, Project-Based Learning is an effective methodology to support Proficiency-Based and Personalized Learning as it allows students to engage in authentic learning opportunities that are meaningful and relevant to their lives and provide opportunities to present evidence of their learning in creative ways. (p. 1)

Universal Assessments

Universal assessments are used with **all students**. They can inform teachers about where students are performing relative to grade-level standards, performance indicators, or developmental milestones; which students need intervention; and how to adjust instruction to affect students' success. (See Tables 1, 2, and 3 in the Appendix for more information.)

Screener

Universal screening is conducted to identify or predict students who may be at risk for poor learning outcomes. Universal screening assessments are typically brief, conducted with all students at a grade level, and followed by additional testing or short-term progress monitoring to corroborate students' risk status.

Interim Assessment

An interim assessment is a form of assessment that educators use to evaluate where students are in their progress toward overall proficiency as determined by state summative assessments. Concerns related to interim assessments are that some do not give teachers the level of detailed information needed to inform instruction and they may not be given frequently enough to impact instructional decisions. Before making a decision on purchasing or using a particular interim assessment tool, educators should determine the purpose of the assessment and how the data will be used.

Benchmark

A benchmark is a point of reference against which a learner's level of proficiency can be measured. Benchmark assessments can be used at a variety of local levels including, but not limited to, classroom, grade, or school level. While benchmark assessments can be used as both screeners and interims, they additionally should be used to collect data for future planning of instruction and curriculum, as well as communicating to learners their progress toward proficiency.

Formative Assessment

Formative assessments provide information to both educators and students about what has been learned, which objectives have been addressed, and what techniques have been successful. They are assessments *for* learning and the data should be used to adapt or adjust instruction accordingly. (See Table 4 in the Appendix for more information.)

Summative Assessments

Summative assessments confirm what students know and can do, typically at the end of a year, semester, course, or instructional unit. They are assessments *of learning*. (See Table 5 in the Appendix for more information.)

Diagnostic and Progress Monitoring: Assessment Data for and from Targeted and Intensive Instruction

Diagnostic and progress monitoring assessments are conducted with only some students but are often necessary to plan instruction and/or intervention to meet the needs of students who require additional supports or encounter new challenges. This may mean more formal and frequent assessments for some students in order to improve outcomes. (See Tables 6 and 7 in the Appendix for more information)

Diagnostic

The purpose of a diagnostic assessment is to investigate and analyze a specific student's strengths and challenges, identify an appropriate focus for intervention, and explore the foundational skills/concepts of a content domain more comprehensively.

Progress Monitoring

The purpose of progress monitoring assessments is to measure a student's progress on a skill/concept in a predefined period of time, determine if the student has made sufficient progress to discontinue intervention or if they need a different or more intensified intervention, and decide if a child with an IEP is making progress on goals and objectives.

Reliability, Validity, and Fidelity

Equitable methods of assessing require appropriate design, administration, and interpretation. Whether designing or selecting assessments, reliability, validity, and fidelity are essential factors for making sound decisions. Additional considerations when selecting and using assessments can be found in the [VTmtss Field Guide](#) and the Framework for [Vermont's Early Childhood Comprehensive Assessment System](#).

Reliability

Reliability refers to the consistency of measurements when assessments are repeated – the degree to which results are free from measurement errors ([AERA, APA, and NCME, 1999](#)). Results or scores reported for individuals or schools must be accurate to support each intended interpretation ([AERA, 2000](#)). Reliability, in conjunction with validity and fidelity, increases confidence that the assessment will provide scores that consistently and accurately determine students' academic abilities. The reliability of local assessment data is increased when teachers have opportunities to calibrate scoring on student work samples. Commercially available assessments will detail their reliability in their technical reports. Large scale and commercial assessments will also provide data regarding reliability.

Validity

Validity refers to the accuracy of the interpretation of assessment results ([Reynolds, Livingston, and Wilson, 2009](#)). In other words, it is the degree to which evidence supports the interpretation of results, entailed by proposed uses of tests (AERA et al., 1999). Therefore, the validation process involves collecting evidence to support the proposed interpretation of results or test scores (Lane, 1999). For example, if assessment results are interpreted as determining reading comprehension, the assessment must actually measure degrees or levels of reading comprehension. Tests valid for one use may be invalid for another (AERA, 2000). Again, commercially available assessments will provide evidence of validity studies in their technical reports. Additional information about reliability and validity can be accessed on page 13 of the document [Criteria for High-Quality Assessments](#).

Fidelity

Fidelity refers to the degree to which a practitioner follows prescription or protocol when delivering an assessment, program, or intervention (e.g., Sanetti & Kratochwill, 2009; Mowbray, Holter, Teague, and Bybee, 2003). Fidelity of administration ensures that assessments produce valid and reliable results. For further reading about the importance of fidelity, please access the National Center on Response to Intervention

Appendix

Table 1: Universal Assessments: Screeners

Category	Summary
Purpose	<ul style="list-style-type: none"> ● Screeners are characterized by the administration of quick, low-cost, repeatable testing of age-appropriate skills to all students.
Frequency	<ul style="list-style-type: none"> ● 3 times a year → Fall, Winter, Spring* ● When a new student enrolls in school outside of a screening window.
What types of questions can this data answer?	<ul style="list-style-type: none"> ● Who is performing above, at, or below grade level? ● Is the child meeting developmental milestones? ● Are at least 80% (or other local target) of students demonstrating proficiency? ● What content should be emphasized in core instruction? ● What are areas of strengths for this cohort of students? ● How functional are the core curriculum, environment, and instruction in the school? ● Who may not be making expected progress and who may need additional diagnostic assessment and/or intervention, either in small groups or on an individual basis?
Tips and Considerations	<ul style="list-style-type: none"> ● Ensure that the screener is accessible to all students, enabling them to show what they know, and does not advantage some students over others. ● Clarify the critical skills and concepts to be assessed. ● Screening will identify students with advanced or emerging skills so schools must be ready to support students at either end of the spectrum. ● It is essential that educators differentiate between universal instruction and a responsive learning environment; additionally, educators must provide targeted or intensive supports when needed. ● Universal screening is an activity that can be part of the Local Education Agency's (LEA's) comprehensive Child Find system to locate and identify children who may be in need of special education and related services.

Resource: [Elementary Assessments: Universal Screening, Diagnostic, and Progress Monitoring](#)

**When a robust, comprehensive, and balanced system is in place, and when all assessment administrators have the necessary assessment and data literacy competencies, scheduling may vary, as determined by students' needs and progress.*

Table 2: Universal Assessments: Benchmarks

Category	Summary
Purpose	<ul style="list-style-type: none"> • Benchmark assessments are assessments administered periodically throughout the school year, at specified times during a curriculum sequence, to evaluate students’ knowledge and skills relative to an explicit set of longer-term learning goals.
Frequency	<ul style="list-style-type: none"> • Once or twice during a unit of study; no more than four times over a year for end-of-year goals (e.g., English language arts).
What types of questions can this data answer?	<ul style="list-style-type: none"> • What knowledge and skills are important to learn? • How do curriculum and instruction need to be adjusted to meet student learning goals? • How well are programs, curricula, or other resources helping students to achieve learning goals? • Are students, classes, schools, and districts on course to demonstrate proficiency?
Tips and Considerations	<ul style="list-style-type: none"> • Develop a clear description of the purpose(s) the benchmark assessment is intended to serve. • Identify what is most important for students to know and be able to do in a specific content area. • Ensure that the benchmark assessment is accessible, enabling all students to show what they know, and does not advantage some students over others. • Determine how useful the assessment will be in helping to accomplish intended purposes.

Resource: [Benchmark Assessment for Improved Learning: AN AACC POLICY BRIEF](#)

Table 3: Universal Assessments: Interims

Category	Summary
Purpose	<ul style="list-style-type: none"> An interim assessment is a form of assessment that educators use to determine whether students are on track for performing at the proficient level on state assessments.
Frequency	<ul style="list-style-type: none"> Interim assessments are often administered three times during a school year. However, the frequency in which interims are administered may vary depending upon the specific assessment tool that is administered.
What types of questions can this data answer?	<ul style="list-style-type: none"> How will students perform on the state assessment? What proportion of students are at risk of scoring below Proficient on the end-of-year state assessments? On which content standards are the students performing relatively well (or poorly) (for a student, classroom, school, district, or state)? How does this student’s performance compare to the performance of other students in the class?
Tips and Considerations	<ul style="list-style-type: none"> Clarify what you hope to learn from this assessment. Determine how the information gathered from this assessment will be used and identify action steps that will be taken as a result. Decide what professional learning or support structures need to be in place to ensure the action steps are taken and are successful. Discuss how student learning improves as a result of using this interim assessment and whether it will improve more than if the assessment was not used.

Resource: [The Role of Interim Assessments in a Comprehensive Assessment System](#)

Table 4: Formative (Assessment *for* Learning)

Category	Summary
Purpose	<ul style="list-style-type: none"> ● To inform the educator about possible facilitators and barriers to instruction for students. ● To provide actionable feedback to students regarding their learning. ● To identify appropriate focus for instruction. ● To inform future lessons and units. ● To determine the efficacy of Universal Instruction.
Frequency	<ul style="list-style-type: none"> ● Several times throughout a learning opportunity.
What types of questions can this data answer?	<ul style="list-style-type: none"> ● Which students or groups of students need re-teaching and reinforcing of concepts? ● Which students or groups of students need acceleration/enrichment? ● In which areas are students proficient? ● What preconceptions/misconceptions do students have about the content/concept? ● How did instruction impact student learning? What could be improved? ● Where are students on a continuum of proficiency?
Tips and Considerations	<ul style="list-style-type: none"> ● Teachers and students need to share a common, clear, and accurate understanding of learning targets. ● Formative assessments are targeted at specific skills and content and can be personalized for individual students. ● Formative assessments occur as part of the learning process so adjustments can be made immediately or in the near future. ● Formative assessments can take many forms such as verbal, written, virtual, or technology-based. ● Fidelity of teacher instruction must also be measured. ● An assessment is not formative if the results are not used to inform instructional decisions.

Resource: [The Future of Assessment Practices: Comprehensive and Balanced Assessment Systems](#)

Table 5: Summative (Assessment of Learning)

Category	Summary
Purpose	<ul style="list-style-type: none"> • To verify learning and confirm what students know, understand, and can do relative to state standards or performance indicators. • To determine the efficacy of Universal Instruction. • To develop and monitor student progress on a personalized learning plan to meet proficiencies.
Frequency	<ul style="list-style-type: none"> • End of unit, course, semester, year
What types of questions can this data answer?	<ul style="list-style-type: none"> • In which standards do students demonstrate proficiency? • Are at least 80% (or other local benchmark) of students demonstrating proficiency? • Which students or groups of students need re-teaching and reinforcing of concepts? • Which students or groups of students need acceleration/enrichment? • Are there students who were not identified through Universal Screening who need intervention or enrichment?
Tips and Considerations	<ul style="list-style-type: none"> • Strategies should be developed to extract the relevant information from the assessment. • Data from summative assessments should be used to inform programmatic decisions. • Assessments need to adhere to principles of validity and reliability.

Resource: Baylor University: [Summative Assessment](#)

Table 6: Diagnostic

Category	Summary
Purpose	<ul style="list-style-type: none"> ● To investigate and analyze a student’s strengths and challenges. ● To identify appropriate focus for intervention. ● To explore the foundational skills/concepts of a content domain more comprehensively (math, literacy).
Frequency	<ul style="list-style-type: none"> ● One time per data cycle or as needed to determine a student’s academic need for intervention or instruction. ● Students with IEPs have a formal (triennial) reevaluation every three years as required by IDEA and no more than once a year if requested by the family.
What types of questions can this data answer?	<ul style="list-style-type: none"> ● What foundational skill/concept is preventing the student from being successful in grade level content/classroom? ● What prior knowledge does the student bring to the task/assessment? ● Have the student’s needs changed? ● Diagnostics in Special Education can help to answer: Is the child a "child with a disability" as defined by IDEA?
Tips and Considerations	<ul style="list-style-type: none"> ● Create a common understanding of foundational concepts and skills. ● Use valid, research-based, norm-referenced tools to diagnose student need. ● Calibrate the administration and scoring of diagnostic assessments. ● Only use diagnostics with students demonstrating a need for closer examination. ● Gather necessary information regarding teacher implementation of evidence-based practices.

Table 7: Progress Monitoring

Category	Summary
Purpose	<ul style="list-style-type: none"> ● To measure students’ progress on a skill/concept in a predefined period of time. ● To determine if students have made sufficient progress to discontinue intervention. ● To determine if a student needs a different or more intensified intervention. ● To determine if a child with an IEP is making progress on goals and objectives.
Frequency	<ul style="list-style-type: none"> ● As determined by the student’s IEP team, EST Team, or intervention tool (Targeted or Intensive).
What types of questions can this data answer?	<ul style="list-style-type: none"> ● What patterns do you notice in the data? ● What are the obstacles to student learning? ● Does this intervention meet the needs of this student? ● What other strategies can be used to support the learner? ● Is the student learning the skills/concepts needed to meet learning targets independently? ● Does the student need more, less, or a different intervention? ● Has the intervention made a difference for the child in gaining skills/concepts within the child’s classroom environment?
Tips and Considerations	<ul style="list-style-type: none"> ● If possible, use valid, research-based, norm-referenced tools to monitor progress; track student progress over time and across different contexts; and set a growth goal for a student based on their current level of performance. ● The frequency of progress monitoring is determined by the student’s needs and progress, which means that monitoring could occur multiple times throughout the day, daily, weekly, or monthly. ● Reference the student’s IEP and review goals and objectives as determined by the IEP team. ● Embedding observation opportunities within the predefined period of time can further inform the educator of the student’s progress.

References

Allensworth, E. (2005). Graduation and dropout trends in Chicago: A look at cohorts of students from 1991 to 2004. Chicago, IL: Consortium on Chicago School Research.

AERA, APA, & NCME (1999). Standards for Psychological Testing. Washington, DC: American Psychological Association.

CCSSO (2013). Toolkit for evaluating alignment of instructional and assessment materials to the common core state standards. Washington, DC: Author.

CCSSO (2015). Comprehensive statewide assessment systems: A framework for the role of the state education agency in improving quality and reducing burden. Washington, CD: Author.

Darling-Hammond, L., Herman, J., Pellegrino, J., et al. (2013). Criteria for high-quality assessment. Stanford, CA: Stanford Center for Opportunity Policy in Education.

Ikeda, M.J., & Neesen, E. (2007). Best practices in universal screening. *Best Practices in School Psychology V*.

Lane, S. (1999). *Validity Evidence for Assessment*, Reidy Interactive Lecture Series.

Linn, R. L. (1993). *Educational Assessment: Expanded Expectations and Challenges CSE Technical Report 351*, CRESST/University of Colorado at Boulder.

Mowbry, C.T., Holter, M.C., Teague, G.B., & Bybee, D. (2003). Fidelity Criteria: Development, Measurement, and Validation. *American Journal of Evaluation*. 24(3), 315-340. doi: 10.1177/109821400302400303.

Reynolds, CR., Livingston, R.G., & Willson, V. (2009). *Measurement and Evaluation in Education*. Upper Saddle River, NJ: Pearson.

Ryan, J. (2010). *Envisioning a state educational system: Improving learning through a comprehensive assessments system*. Office of Superintendent of Public Instruction. Olympia, Washington.

Sarlo, R. (n.d.) *Early Warning Systems: Moving from Reaction to Prevention*. RTI Action Network.

Sanetti, L. M., & Kratochwill, T. R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review*, 38(4), 445–459.

U.S. Department of Education (2015). *Fact Sheet: Testing Action Plan*.

Vermont State Board of Education Rule 2000 (2014). *Education quality standards*. Montpelier, VT: Vermont Agency of Education.

Vermont Statewide Steering Committee on Response to Instruction and Intervention (2014).

Vermont multi-tiered system of supports response to intervention and instruction field guide. Burlington, VT: Author.