

Galileo® K-12 Online Assessments

from Assessment Technology Incorporated / Identified by Massachusetts Department of
Elementary and Secondary Education in Partnership with WestEd as
Appropriate for Use as District-Determined Measures (DDMs)

Alignment Specifications	
Criterion	Galileo K-12 Online Customized DDMs and Assessment Technology Incorporated Ongoing Research Program
Growth: Measures change over time, not just achievement	Galileo K-12 Online pretests and posttests assess student knowledge on the same or vertically aligned standards before and after the learning period. Since Assessment Technology Incorporated (ATI) research staff places assessment scores on a common scale across assessments using procedures based on Item Response Theory, the change in scores is a direct measure of growth.
Growth: Identifies how much growth is sufficient for the period covered by the assessment	ATI conducts annual research using a large multi-state sample of students to establish the amount of expected growth that is sufficient for students to maintain their performance level throughout the year. This research is used to establish expected growth for the relevant time period between any two tests administered by a given district. In cases where the research to establish expected growth constants is still in progress, ATI uses equipercentile equating procedures to align the distribution of student scores on each test directly to the Massachusetts Comprehensive Assessment System (MCAS) distribution, yielding a measure of sufficient growth based on student performance in that district during that time period.
Growth: Measures change relative to an academic peer group	The Galileo K-12 <i>Aggregate Multi-Test Report</i> quantifies the change in a student's ability across time. Filters on the report enable the user to select student peer groups reflecting diverse academic, demographic, and other characteristics to evaluate the student's change relative to his or her peer group.
Consistency of administration	When a test is administered online, the date(s) when the test may be taken and the time allocated for the test are controlled electronically by test scheduling tools. When the test is administered offline, Galileo provides test scheduling tools indicating when the test will be available for administration as well as recommendations that assist districts in ensuring consistent and secure administration procedures. When offline administration scheduled end date has been reached, the test is closed preventing further administration.

Alignment Specifications

Criterion	Galileo [®] K-12 Online Customized DDMs and Assessment Technology Incorporated Ongoing Research Program
Alignment to standards	ATI item banks contain over 110,000 items written according to specifications that reflect the full depth and breadth of targeted grade-level standards. Specifications are designed to reflect all of the necessary elements of each standard.
Content Validity: Instructional sensitivity	ATI provides tests through Galileo that are customized to meet district needs. Items are written by content development specialists in ATI's Assessment and Instructional Design Department. Department staff members have many years of classroom experience. Items are written to specifications defining essential item characteristics including sensitivity to the needs of students from diverse backgrounds. Each item written undergoes an initial review and a final review to ensure its conformance to specifications. In addition, draft tests are provided to districts for their review to ensure that items reflect what is being taught in district classrooms.
Reliability: Items	Galileo item banks are very large. ATI maintains information on estimated item parameters affecting reliability including item difficulty, item discrimination, and (for selected-response items) guessing. ATI computes reliability coefficients for all district-wide assessments. Reliability is a function of test length. Reliability coefficients for Galileo district determined measures of adequate length are generally in the mid to high .80s or low .90s.
Reliability: Scoring of open-ended responses	Scoring rules (rubrics) are supplied for all open-ended response items. These rules indicate in concrete terms what the student must do to achieve the different levels of scores that may exist for a given open-response item.
Reliability: Rater training	ATI provides training guidelines for establishing reliability. Guidelines include recommendations related to assigning raters to items and students, to measuring agreement among raters and to training until a criterion is reached. ATI also provides procedures for checking the reliability of scoring.
Reliability of scores	ATI uses mathematical models to measure inter-rater reliability. The results provide evidence as to scoring reliability.
Fairness and freedom from bias	ATI assessments utilize universal design principles to ensure that assessments are free of elements that would prevent some subgroups of students from showing their capabilities.