EDUCATION COMMISSION
OF THE STATES
Your education policy team.

Glossary of Assessment Terms

Constructed response questions entail a written response, single part or building and solving a single equation.

Criterion-referenced tests measures student performance against defined objectives, skill levels or knowledge and show a demonstrated mastery/knowledge of a specific subject.

Diagnostic tests are given at the beginning of the school year, or the beginning of a new unit of study, and attempts to quantify what students already know about a topic. It assess strengths, weakness, and prior knowledge and identifies a student's preconceptions of a topic and can help teachers link further instruction on what students have already mastered.

Extended constructed response questions require a written response, with multi-part questions, requiring students to answer the first part of the question before answering subsequent parts, which may require reflection on or further explanation of an answer given in a previous section.

Formative assessments are given throughout the learning process with an intention of determining how students are progressing through a certain learning goal. They are often low-stakes and ungraded. This type of assessment allows for student feedback. The timing of information is provided during the school year when adjustments and interventions can be utilized. Both diagnostic and summative assessments can be used as formative assessments.

Informal assessments can be a wide range of tasks ranging from checklists to observations and do not necessarily lead to a score. It helps teachers gather insights about students.

Interim assessments are a form of test is used to evaluate where students are in their learning progress and determine whether they are on track to performing well on future assessments, such as standardized tests or end-of-course exams.

Multiple-choice questions allow for possible answers to generally include one correct answer and three to four distracters, designed to mimic the common misconceptions students have about the concept being tested. This type of question can be efficient and practical. Carefully designed multiple-choice questions can provide valid information about students' knowledge and their ability to reason logically and apply complex thinking processes to solve problems. Often these tests are scored by computers and provide impartial, accurate results.

Norm-referenced tests measure students against a national "norm" or average in order to rank students against each other (SAT, ACT, Iowa Basic Skills Test), measure basic concepts and skills commonly taught in schools throughout the country. These tests are not designed as precise measures of any given curriculum or single instructional program.

Online tests require each student to have access to a device on which to take the assessment. Students also need strong digital literacy skills in order to successfully navigate the demands of online testing. Internet is also required for an online assessment.

Online adaptive tests adjust as the user progresses through the questions, meaning the difficulty level of questions are gauged to student responses. This type of test can provide a broader picture of students' ability levels.

Performance assessments directly assess pupil performance by requiring a written essay or short response, drawing of a conclusion, responding to a reading passage or performing a science experiment. They are generally scored through teacher observation of student performance who rate the outcomes. Performance assessments can also be standardized so that the test is given and scored the same way at each administration.

Reliability is the extent to which an assessment yields consistent information about the knowledge, skills, or abilities being assessed. An assessment is considered reliable if the same results are yielded each time the test is administered.

Standardized tests are an assessment that is always given in a consistent manner, with the same directions, the same questions and the same time limits. Scores can be compared with confidence in test validity and reliability.

Summative Assessment are designed to measure student learning at the end of a project, unit, course or school year. Summative assessment scores are usually recorded and factored into student's academic record in the form of letter grades and/or scores from tests like SAT or ACT. Summative assessments are used by states to evaluate a school or district's performance for federal and state accountability measures, may be used to evaluate a teacher's performance using her student's scores and may be used for graduation or advancement requirements.

Technology enhanced tests are computer delivered and require students to interact with the material through actions like dragging and dropping information, highlighting relevant text and completing sentences or equations from items in a drop-down menu.

Validity is defined as the extent to which an assessment accurately measures what it is intended to measure.