

**Position Paper:
ASSESSMENT, TESTING AND LEARNING**

NYS PTA POSITIONS

All Educators must:

- establish balanced use of assessments to inform instruction without over testing or overemphasizing test performance
- create and use high quality assessments to measure what is important in a meaningful manner that meets the standards with validity, reliability, fairness and developmental appropriateness - especially for students with disabilities or new English learners
- maintain a positive environment for students during assessment/test administration
- minimize the use of testing and test preparation so to maximize the learning environment and support whole child learning
- avoid the use of a single test or single assessment measure for the placement, promotion, or retention of a student, or within high stakes decisions affecting students
- establish developmentally appropriate seat time when assessing students
- use assessment as part of a holistic improvement of learning and strategic planning for the individual student, groups of students and school systems, rather than using assessment results for punishment or competition
- never use standardized multiple-choice tests and school readiness tests with preschool and early elementary children for any purpose

Boards of Education must:

- examine assessment data to inform and support improved instruction, such as curriculum design, improved instructional materials and professional development, and assistance to students when gaps are identified
- annually review all assessment data with parents and the general public to engage in a dialogue when establishing priorities for programs, staffing, budgeting and strategic planning

Principals and Teachers must:

- score assessments and present student data in a timely, easy to understand manner to parents and students, where possible
- provide mutually convenient opportunities for parents and students to review assessment data for the purposes of establishing individual learning goals, describing measures of expected learning and providing support for learning outside the school day

BACKGROUND

Assessment is an essential aspect of the learning process since it provides feedback intended to show the level at which instruction should begin, the effectiveness of ongoing learning strategies, and the level of proficiency attained as a result of the instructional process. It is intended to benefit the student, to inform teachers who guide each student's progress and to assist parents in understanding how to support their child's learning. There are many types of assessment, with what we know as tests representing a relatively small subset of the numerous assessment tools available.

Types of Assessment

Benchmark – is the initial measurement of knowledge or skill. Subsequent measures show student growth and skill acquisition over a period of time.

Formative – also called diagnostic, are relatively short assessments intended to provide teachers with very quick feedback that can shape day to day instruction. Formative assessment may be as simple as a answering a question within or at the end of class, or be a short quiz or demonstration.

Summative – often found in the form of final exams or achievement tests, measure what a student knows at the completion of some period of instruction. Regents exams and annual portfolios, and certification exams provide a summation of learning for the year or program period.

Performance – beyond measuring what a student knows, these assessments demonstrate what a student can do with that knowledge. Not typically multiple choice or “bubble tests,” performance assessments require multiple steps to

complete a task, often based on higher level thinking questions. Students are asked to show their work and are typically given credit for the process they use, along with arriving at the “correct” answer.

Portfolio – draws conclusions of student performance based on samples of student work over a period of time, for example reading /writing logs. While this form of assessment may provide the best evidence of what a student knows and can do, it requires clearly defined rubrics or benchmarks of performance.

Adaptive – is a relatively new form of assessment usually completed on a computerized platform. Each test item/question is selected based on the response to previous items. If a student responds correctly to an item, the program selects a more difficult item as the next question. Likewise, if a response is incorrect, a less difficult item is presented. The purpose is to pinpoint a student’s actual achievement level in the least amount of time.

Standardized – while individual teachers construct their own benchmark, formative and summative assessments; standardized tests constructed by multiple people, have undergone a broad review to determine that they fairly and consistently measure what they are intended to measure. Such tests are typically developed by research/testing/publishing companies and reviewed by testing experts (psychometricians) to be administered to large groups. The NYS Regents exams, SAT, ACT, NAEP and PISA are examples of standardized assessments.

Norm -Referenced – uses the score of one test taker compared with or against others who have taken the test. For example, a score of 70 could mean that the student scored *better than 70% of the other people who took the same test*.

Criterion-referenced – measures whether students have mastered specific content without comparison to others who take the test. For example, a score of 70 would typically mean that the student got *70% of the items* correct.

Assessment Standards

Regardless of the assessment intent, tests are most useful in shaping instruction if they are fair, achieve consistent results, measure what they are intended to measure and are free of bias. The following terms are often heard in discussion of tests or assessments.

Validity. Does the assessment measure what it is supposed to measure?

Reliability. If the same assessment were administered multiple times to the same student, would the results be similar or consistent?

Fair or Bias free. Are questions or items meaningful for students of different cultural backgrounds such that items measure achievement rather than the impact of unintended factors, such as race, residence, or family wealth?

Developmental Appropriateness. Do assessment items use language and examples that measure what is intended in language and context familiar to students’ age and educational development? Are the time on task and organization of the test developmentally appropriate for students?

Assessment Data

Assessment results or data may be used to provide entrance to specific programs or assign learning supports for a student, revise instruction or curriculum, highlight professional development for teachers, evaluate teachers and/or compare school effectiveness.

Scaled Score – is determined by the number of points earned on a test so that the test results mean the same thing year after year even though different students are taking the test with different questions.

Cut Score/Proficiency Level – is a performance judgment, typically established for state tests once raw (number correct) scores are known. For example, a “cut” score of 70% correct may be judged as the minimum definition of proficiency in one year where a “cut” score to define proficiency in a subsequent year may be set at 80% correct answers. Cut scores are aligned with levels of proficiency and percentile rank, which is a comparison to all students who have taken this specific test.

Growth Scores – compare a prior assessment score (benchmark) against the current assessment score and other students’ performance. Growth scores may consider additional impacting factors such as poverty or designation as a student with disability or new English learner.

Adopted July 2015