

The Appropriate Use of Student Assessments



A Union of Professionals

AFT Teachers



The Appropriate Use of Student Assessments

We know that skilled teachers who assess students using high-quality, relevant and timely assessments can use the results to modify, refine and individualize instruction. We also know that student assessments should be a reflection of what is taught in the classroom.

As more accountability provisions have been implemented at the national, state and district levels, the focus on test scores and their implications has dictated what students should learn. Although the federal government requires testing in most grades, these requirements are often supplemented at the state and district levels. This results in increased testing and test preparation without providing additional information to help teachers inform instruction and understand their students' academic standing and needs.

The AFT believes it is critical to define and describe appropriate testing practices, to advocate on behalf of teachers and their students for sound assessment policies, and to align standards and curricula at each grade level.

The Current Landscape

Polls conducted by the AFT and other organizations have shown that too many states and districts are using tests inappropriately by:

- Requiring teachers to use large-scale state assessments for diagnostic purposes, even though such tests do not provide diagnostic information.
- Requiring teachers to use large-scale test results to guide instruction, even though teachers often receive these test results too late to be used in this manner.

- Basing high-stakes decisions, such as whether a student should graduate from high school or be promoted to the next grade, on a single test score, even though assessment experts agree that this is not appropriate.
- Using tests, including large-scale and benchmark assessments, that are not aligned to the standards and curriculum that teachers are required to teach.
- Requiring teachers to use benchmark or interim tests—in some cases as often as every six weeks—to help predict performance on state-level assessments, even though such predictions can be unreliable, and can result in teachers narrowing their curriculum and focusing more on test preparation than instruction.
- Trying to measure the performance of individual teachers based on student test scores, even though the tests are not designed to provide such specific information, often are not aligned to the curriculum, and use methodology that education experts have questioned.
- Requiring teachers to spend valuable instructional time on test preparation, especially in reading and math, which leads to teaching testing strategies rather than academic content. This has resulted in narrowing the curriculum.
- Failing to provide adequate in-service opportunities to help teachers understand the appropriate use of tests they are required to administer, and to help them design the formative (also known as classroom-based) assessments they need to inform instruction.
- Implementing assessment policies for students with disabilities and English language learners that are unrealistic and serve only to discourage students, parents and teachers.

Recommendations for the Appropriate Use of Tests for Our Students

It is imperative that states, districts and teachers understand the purpose of various types of tests and test results. In some cases, new tests that yield more useful results may need to be developed. To help create a common language and understanding of what is appropriate, we've defined the most commonly used tests and provided recommendations about how these tests should be used. A list of commonly used tests is included in the Appendix.

1. **Norm-referenced tests (NRTs) are designed to compare students, schools, district and states with each other.** NRTs provide insight into how students in California, for example, compare with students in New York. These tests do not tell us how well any of these students performed in relation to a standard. Instead, students are scored based on how well they did compared with their peers. These results are typically reported as percentiles and as a bell-shaped curve, where half of students will fall below the 50th percentile and half will fall above.
2. **Criterion-referenced tests (CRTs) are designed to compare individual student performance against a specified standard.** CRTs provide information about whether students met certain academic standards. The results are typically reported as performance levels (basic, proficient, advanced). Student scores are based on how well they know the content and how they answer the questions—not on how well their peers performed on the same questions. Data from CRTs should be used to inform programmatic/instructional decisions, provided they reflect the same standards that are being taught in the schools.
3. **Formative assessments are used to guide instruction.** Formative (classroom-based) assessments occur during teaching and are embedded in instruction. Results are received instantly, which allows teachers to adjust their instruction immediately. These are typically teacher-developed and should be implemented based on teacher judgment.

4. **Summative assessments provide snapshots of how well students have mastered the standards at a particular point in time.** These assessments occur at the end of a unit of instruction and tell us whether students “got it.” Results are received anywhere from two weeks to two months later. As a result, these tests cannot guide instruction in the short term. Results can provide some information regarding programmatic/instructional decisions and can guide the future delivery of material covered during the unit, provided the test is aligned with the standards being taught.
5. **Benchmark/interim assessments are designed to reflect the content being taught in the classroom.** Benchmark assessments are standards-based assessments administered at regular intervals. The assessment results can be used to determine student growth and student performance relative to statewide grade-level achievement expectations. Unfortunately, many benchmark/interim tests are being used as a predictor of future success on assessments, and are not aligned to the curriculum currently being taught. As a result, these tests interrupt classroom instruction rather than complement it. Benchmark/interim tests should provide information to teachers about their instruction and where each student is in relation to the content he or she is learning.
6. **Diagnostic assessments ascertain each student’s strengths, weaknesses, knowledge and skills on a few concepts.** This type of assessment allows teachers to diagnose students’ content knowledge and remediate any weaknesses by adjusting curricula and instruction to meet student needs. For an assessment to provide educators with enough diagnostic information to make valid judgments about their students, it must include enough questions about a topic as well as both easy and difficult questions (called “outliers”). For this reason, many tests, including high-stakes and some benchmark/interim tests, are not appropriate for diagnostic use. These tests cover numerous topics, with often only a few questions

per topic. In addition, these tests are designed to eliminate “outliers” that could skew the data.

7. **Adaptive tests are designed to identify the level at which students are performing for a particular subject or concept.** Adaptive testing is usually done by computer and asks students more difficult or less difficult questions based on their answers to previous questions. This approach allows teachers to better focus instruction on each child’s strengths and weaknesses by helping to identify the specific concept or process where learning has broken down.
8. **Value-added assessments are designed to measure the growth of a student’s knowledge over time.** Value-added assessments can help schools and classroom teachers make data-informed decisions regarding the effectiveness of instructional strategies and programs for individuals and groups of students. This type of test is an estimation tool and, therefore, should not be used to make high-stakes decisions about students, teachers or other school staff. Effective value-added assessments must be of high quality and be closely aligned with classroom instruction. In addition, states or districts must create data systems that include unique student identifiers to track individual students from year to year. Officials must also be able to compare test results from year to year on a single scale.

Providing a common language around assessments is a small, albeit important, step in creating a valid and useful assessment system. States and districts also must ensure that the following issues are addressed:

- Standards must be provided for each grade at the K-8 level and for each course at the high school level, and must be clear, sequenced and focused on specific content.
- The tests must be aligned to the standards and the curriculum that teachers are expected to deliver in their classrooms. A 2006 study from the AFT found that only 20

percent of the states administered tests that reflected the content and skills required in their standards. This means that most states have significant work to do to ensure that what teachers are expected to teach in the classroom is aligned with the knowledge and skills students are expected to demonstrate on state tests.

- Clear curriculum and classroom resources must be provided to enable teachers to provide quality instruction. This will allow teachers to expose students to subject-area content while also ensuring that students are prepared for the state tests.
- Professional development opportunities on assessments must be provided to help staff understand the use of tests and test results. Topics that need to be covered include: appropriate uses of testing, developing quality classroom assessments, assessing special populations, incorporating formative assessment techniques in instruction, and analyzing student performance data.
- High-quality professional development for teachers must be provided on an ongoing basis. This professional development should help teachers provide individualized instruction, particularly for students with special needs.
- Teachers must be involved in the development of tests. This involvement will help align assessments to curriculum and instruction, as well as facilitate teacher buy-in. Teachers involved in developing these assessments may become a resource for other teachers. They can provide insight into how and why the tests were created, and can build confidence among teachers that the tests are aligned to curriculum and support instruction.

Conclusion

States and districts must create reporting systems that provide teachers and parents with clear, timely information. These systems also should give teachers the resources to improve instruction on skills and content where student performance is not adequate.

Now, more than ever, there is a need for assessments that are based on common, content-rich standards and curricula. Students must have opportunities to delve deeper into a wide variety of concepts. This is not possible in a testing environment that uses tests and test results for unintended purposes, and that requires teachers to spend endless hours in test preparation. America's students and teachers deserve much better.

**Appendix:
Examples of Commonly Used Assessments**

Assessment Type	Examples
Norm-Referenced Test	Stanford 10 Iowa Test of Basic Skills (ITBS) SAT
Criterion-Referenced Test	Standards-based tests Advanced Placement (AP) exams National Assessment of Educational Progress (NAEP)
Formative Assessment	One-on-one and group questioning Observation checklists Short quizzes
Summative Assessment	State tests End-of-course exams Advanced Placement exams
Benchmark / Interim Assessment	District benchmark tests 4Sight, Inc. Data Driven Software Corporation (D2SC)
Diagnostic Assessment	Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Phonological Awareness Literacy Screening (PALS) Direct Reading Assessment (DRA)
Adaptive Assessment	Measures of Academic Progress (MAP) Children's Progress
Value-Added Assessment	Tennessee Value-Added Assessment System (TVAAS) Project SOAR (Sharing Opportunities for Reading) Pennsylvania Value-Added Assessment System (PVAAS)





A Union of Professionals

American Federation of Teachers, AFL-CIO
555 New Jersey Ave. N.W.
Washington, DC 20001
202/879-4400
www.aft.org
June 2008