

Large-Scale Mathematics Assessments and High-Stakes Decisions

A Position of the National Council of Teachers of Mathematics

Question: How should large-scale mathematics assessments be used in making significant, high-stakes decisions about schools, teachers, and students?

NCTM Position

The results of large-scale mathematics assessments should not be used as the sole source of information to make high-stakes decisions about schools, teachers, and students. High-stakes decisions should also take into account relevant and valid data on classroom-based performance, such as formative and summative assessments of high quality that offer students a range of opportunities to demonstrate their mathematical knowledge. Moreover, educational systems—states, districts, and schools—should be held accountable for providing essential support for high-quality mathematics teaching and learning before teachers and students are held accountable for assessment results.

Decisions about school performance necessitate the use of multiple measures designed to supply relevant and valid information about student learning outcomes. Large-scale mathematics assessments, such as PARCC and SBAC, provide one set of data, which, when used as part of a comprehensive assessment system, can help to determine how well schools are serving students in their learning of mathematics. A vital role for large-scale mathematics assessments in schools and districts is to afford opportunities to identify systematic gaps in student mathematical content knowledge and examine the alignment of the district mathematics curriculum with the standards in place. The results from large-scale mathematics assessments, when used along with data from more finely grained assessments of student learning, provide an opportunity for schools and districts to self-assess and make appropriate revisions to their mathematics curriculum, instruction, and assessment.

High-stakes decisions about students must be informed by and based on evidence from multiple data sources. The *Code of Fair Testing Practices in Education* (2004) recommends that educators “avoid using a single test score as the sole determinant of decisions about test takers” (p. 9). Standardized test results can be overanalyzed and misinterpreted (NCTM, 2014) and can lead to invalid inferences. Results from large-scale mathematics assessments are one source of information and provide a snapshot of student mathematical understanding on a particular set of problems on a particular day. By contrast, ongoing formative and summative assessments provide a moving picture of what students know and are able to do and offer evidence of how well students are progressing toward established learning goals. Structures must be created to ensure that the results of all assessments are used to strengthen teaching, curriculum and support for all students (*Principles to Actions*). Major decisions about student placement, promotion, or graduation should be made on the basis of valid inferences drawn from a comprehensive set of reliable evidence. Furthermore, the basis for such decisions should be transparent and accessible to parents and students to reduce inequities in the implementation of such policies.

References

- American Psychological Association. (2004). *Code of fair testing practices in education*. Joint Committee on Testing Practices. Washington, DC: Author. Retrieved from <http://www.apa.org/science/programs/testing/fair-testing.pdf>
- Romagnano, L. (2006). *Mathematics assessment literacy: Concepts and terms in large-scale assessment*. Reston, VA: NCTM.
- National Council of Teachers of Mathematics (2014). *Principles to actions: Ensuring mathematical success for all*. Reston, VA: Author.