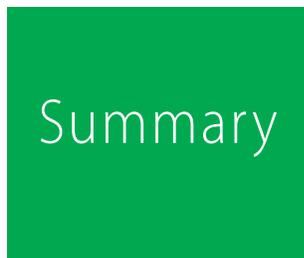


Measuring how benchmark assessments affect student achievement



Institute of Education Sciences
U.S. Department of Education



Measuring how benchmark assessments affect student achievement

Summary

December 2007

Prepared by

**Susan Henderson
WestEd**

**Anthony Petrosino
WestEd**

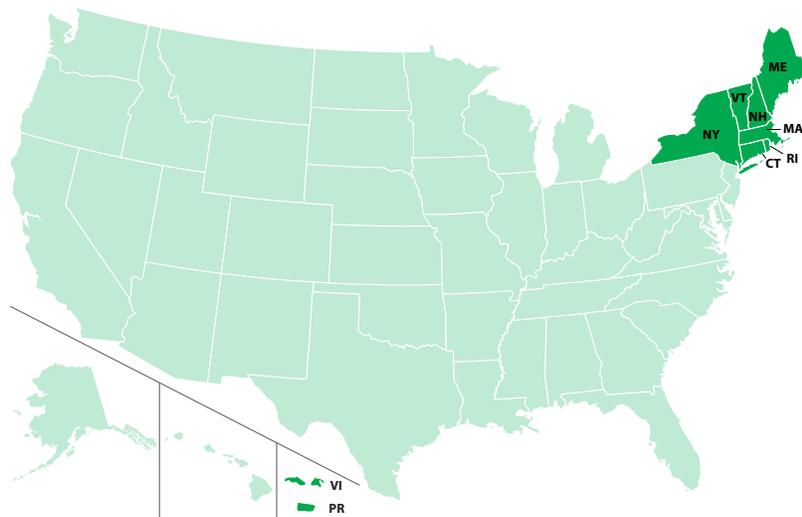
**Sarah Guckenburg
WestEd**

**Stephen Hamilton
WestEd**



Institute of Education Sciences

U.S. Department of Education



Issues & Answers is an ongoing series of reports from short-term Fast Response Projects conducted by the regional educational laboratories on current education issues of importance at local, state, and regional levels. Fast Response Project topics change to reflect new issues, as identified through lab outreach and requests for assistance from policymakers and educators at state and local levels and from communities, businesses, parents, families, and youth. All Issues & Answers reports meet Institute of Education Sciences standards for scientifically valid research.

December 2007

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-06-CO-0025 by Regional Educational Laboratory Northeast and Islands administered by Education Development Center, Inc. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

This report is in the public domain. While permission to reprint this publication is not necessary, it should be cited as:

Henderson, S., Petrosino, A., Guckenbug, S., & Hamilton, S. (2007). *Measuring how benchmark assessments affect student achievement* (Issues & Answers Report, REL 2007–No. 039). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast and Islands. Retrieved from <http://ies.ed.gov/ncee/edlabs>

This report is available on the regional educational laboratory web site at <http://ies.ed.gov/ncee/edlabs>.

Summary

Measuring how benchmark assessments affect student achievement

This report examines a Massachusetts pilot program for quarterly benchmark exams in middle-school mathematics, finding that program schools do not show greater gains in student achievement after a year. But that finding might reflect limited data rather than ineffective benchmark assessments.

Benchmark assessments are used in many districts throughout the nation to raise student, school, and district achievement and to meet the requirements of the No Child Left Behind Act of 2001. This report details a study using a quasi-experimental design to examine whether schools using quarterly benchmark exams in middle-school mathematics under a Massachusetts pilot program show greater gains in student achievement than schools not in the program.

To measure the effects of benchmark assessments, the study matched 44 comparison schools to the 22 schools in the Massachusetts pilot program on pre-implementation test scores and other variables. It examined descriptive statistics on the data and performed interrupted time series analysis to test causal inferences.

The study found no immediate statistically significant or substantively important difference

between the program and comparison schools. That finding might, however, reflect limitations in the data rather than the ineffectiveness of benchmark assessments.

First, data are lacking on what benchmark assessment practices comparison schools may be using, because the study examined the impact of a particular structured benchmarking program. More than 70 percent of districts are doing some type of formative assessment, so it is possible that at least some of the comparison schools implemented their own version of benchmarking. Second, the study was “underpowered.” That means that a small but important treatment effect for benchmarking could have gone undetected because there were only 22 program schools and 44 comparison schools. Third, with only one year of post-implementation data, it may be too early to observe any impact from the intervention in the program schools.

Although the study did not find any immediate difference between schools employing benchmark assessments and those not doing so, it provides initial empirical data to inform state and local education agencies.

The report urges that researchers and policymakers continue to track achievement data in the program and comparison schools, to

reassess the initial findings in future years, and to provide additional data to local and state decisionmakers about the impact of this benchmark assessment practice.

Using student-level data rather than school-level data might help researchers examine the impact of benchmark assessments on important No Child Left Behind subgroups (such as minority students or students with disabilities). Some nontrivial effects for subgroups might be masked by comparing school mean scores. (At the onset of the study, only school-level data were available to researchers.)

Another useful follow-up would be disaggregating the school achievement data by mathematics content strand to see if there are any effects in particular standards. Because the quarterly assessments are broken out by mathematics content strand, doing so would connect logically with the benchmark assessment strategy. This refined data analysis might

be more sensitive to the intervention and might also be linked to information provided to the Massachusetts Department of Education about which content strands schools focused on in their benchmark assessments.

Conversations with education decisionmakers support what seems to be common sense. Higher mathematics scores will come not because benchmarks exist but because of how a school's teachers and leaders use the assessment data. This kind of follow-up research, though difficult, is imperative to better understand the impact of benchmark assessments. A possible approach is to examine initial district progress reports for insight into school buy-in to the initiative, quality of leadership, challenges to implementation, particular standards that participating districts focus on, and how schools use the benchmark assessment data.

December 2007