What Is the TIMSS 1999 Benchmarking Study?

The TIMSS 1999 Benchmarking Study enabled states and school districts in the United States to participate in TIMSS 1999, also known as TIMSS-Repeat or TIMSS-R. The Benchmarking Study is a voluntary component of TIMSS 1999 in which U.S. states and school districts were given an unprecedented opportunity to "benchmark" the mathematics and science achievement of their students against the world class performance of students in the top-scoring TIMSS 1999 countries. TIMSS 1999, a successor to the 1995 Third International Mathematics and Science Study (TIMSS), focused on the mathematics and science achievement of eighth-grade students. Thirty-eight countries including the United States participated in TIMSS 1999, and 27 jurisdictions from all across the U.S., including 13 states and 14 districts or consortia, participated in the Benchmarking Study.

In addition to measuring achievement, TIMSS 1999 and the Benchmarking Study investigated the contexts for learning mathematics and science in the participating entities through background questionnaires completed by students, teachers, school principals, and project coordinators from the participating entities. Information was collected about educational systems, curriculum, instructional practices, and characteristics of students, teachers, and schools, providing an extremely rich source of valuable insights into the teaching and learning of mathematics and science. The TIMSS results have stirred debate, spurred reform efforts, and provided important information to decision makers, researchers, and practitioners the world over.

Who Participated in the TIMSS 1999 Benchmarking Study?

The Participants Map (PDF), which lists the 13 states and 14 districts or consortia
that participated in TIMSS Benchmarking and the 38 countries that participated in TIMSS 1999.

Geographically the Benchmarking jurisdictions were from all across the U.S., although there was a concentration of east coast participants with six of the states and several of the districts and consortia from the eastern seaboard. There was substantial diversity across the Benchmarking jurisdictions in the size and socioeconomic composition of their student populations, as well as in their per pupil expenditure on education. Although taken collectively the Benchmarking participants are not representative of the United States, the effort was substantial in scope involving approximately 1,000 schools, 4,000 teachers, and 50,000 students.

States participating in the Benchmarking Study were required to sample at least 50 schools and approximately 2,000 eighth-grade students. School districts and consortia were required to sample at least 25 schools and at least 1,000 students. Where there were fewer than 25 schools in a district or consortium, all schools were to be included, and the within-school sample increased to yield the total of 1,000 students. All but five Benchmarking jurisdictions (four states and one consortium) included public schools only. For the most part, the United States TIMSS 1999 national sample was separate from the students assessed in each of the Benchmarking jurisdictions. Each Benchmarking participant had its own sample to provide comparisons with each of the TIMSS 1999 countries including the United States.

**What Are the Consortia?**

The consortia consist of groups of entire school districts or individual schools from several districts that organized together either to participate in the Benchmarking Study or to collaborate across a range of educational issues. Descriptions of the consortia that participated in the project follow.

**Delaware Science Coalition.** The Delaware Science Coalition (DSC) is comprised of 15 school districts working in partnership with the Delaware Department of Education and the business-based Delaware Foundation for Science and Mathematics Education to improve the teaching and learning of science for all students in grades K-8.

**First in the World Consortium.** The First in the World Consortium consists of 18
school districts from the North Shore of Chicago that have joined forces to bring a world-class education to the region’s students and to improve mathematics and science achievement in their schools.

Fremont/Lincoln/Westside Public Schools. The Fremont/Lincoln/Westside consortium is comprised of three school districts in Nebraska. These districts joined together specifically to participate in the Benchmarking Study.

Michigan Invitational Group. The Michigan Invitational Group is a socioeconomically diverse consortium composed of urban, suburban, and rural schools across Michigan. Schools invited to participate as part of this consortium were those that were using National Science Foundation (NSF) materials, had well-developed curricula, and provided staff development to teachers.

Project SMART Consortium. SMART (Science & Mathematics Achievement Required For Tomorrow) is a consortium of 30 diverse school districts in northeast Ohio committed to long term systemic change and improved student learning in science and mathematics in grades K-12. It is jointly funded by the Ohio Department of Education and the Martha Holden Jennings Foundation.

Southwest Pennsylvania Math and Science Collaborative. The Southwest Pennsylvania Math and Science Collaborative coordinates efforts and focuses resources on strengthening mathematics and science education in the entire southwest Pennsylvania workforce region that has Pittsburgh as its center. The Collaborative is composed of all 118 “local control” public school districts, as well as the parochial and private schools in the nine-county region.

**Why a Benchmarking Study?**

To meet the challenge of preparing children around the world for a technologically oriented 21st century, policy makers and educators need information about students’ understanding of mathematics and science to improve learning and instruction. Over the last decade, many states and school districts have created content and performance standards targeted at improving students’ achievement in mathematics and science. There has been an enormous amount of energy expended in states and school districts not only on developing mathematics and science content standards but also on improving teacher quality and school environments as well as on developing assessments and accountability measures.
Participation in the TIMSS 1999 Benchmarking Study was intended to help U.S. states and school districts assess the comparative international standing of their students’ achievement, evaluate the rigor and effectiveness of their mathematics and science programs in a global context, and improve the teaching and learning of mathematics and science. Regardless of its performance, each state, district, and consortium will have a better idea of the challenges ahead and access to a rich array of information about various facets of its educational system. The TIMSS 1999 Benchmarking results provide an excellent basis for examining how best to move from developing curriculum frameworks or content standards in mathematics and science to meeting the extraordinary challenge of actually implementing the standards in schools and classrooms often characterized by considerable cultural, socioeconomic, and experiential diversity.

**What Was the Nature of the Test?**

The mathematics and science tests were based on the TIMSS curriculum frameworks, which were developed by groups of educators with input from the TIMSS National Research Coordinators (NRCs). The tests were developed through a consensus process by international experts in mathematics, science, and educational measurement, and were endorsed by all participating countries. Working within the frameworks, test specifications were developed that included items representing a wide range of mathematics and science topics and eliciting a range of skills from the students. The mathematics test covered five content areas: fractions and number sense; measurement; data representation, analysis, and probability; geometry; and algebra. The science test covered six content areas: earth science; life science; physics; chemistry; environmental and resource issues; and scientific inquiry and the nature of science.

The tests included multiple-choice questions, comprising about three-fourths of the items, and open-response items requiring students to solve problems and explain their answers. To achieve broad content coverage, a matrix sampling technique was used in which the 308 test items (162 mathematics and 146 science) were systematically distributed across eight test booklets, and the booklets were randomly distributed to students. Each student in the sampled classrooms responded to one test booklet that included about 80 mathematics and science questions, requiring 90 minutes to complete. About half the items used in 1999 have been released for public use (available at the TIMSS 1999 International).
What Is the Comparability of the Results?

To conduct the Benchmarking Study, the TIMSS 1999 assessments were administered to representative samples of eighth-grade students in each of the participating jurisdictions in the spring of 1999, at the same time and following the same guidelines as those established for all 38 countries. Procedures used throughout both TIMSS 1999 and the Benchmarking Study ensure that the results are comparable across participating entities. To ensure comparability in testing, rigorous procedures were designed to translate the tests where necessary, and numerous training sessions were held in data collection and scoring activities. Quality control monitors observed testing sessions in all jurisdictions and reported back to the International Study Center at Boston College, which manages the TIMSS studies. The samples of students selected for testing were scrutinized according to rigorous standards designed to prevent bias and ensure comparability. In general, the Benchmarking samples were drawn in accordance with the TIMSS standards, and achievement results can be compared with confidence. Prior to analysis, the data from each participating entity were subjected to exhaustive checks for accuracy and consistency. In short, TIMSS 1999 and the Benchmarking Study used the same achievement tests and background questionnaires, the same sampling definitions and procedures, the same test administration procedures, and the same data analysis and scaling methods, all of which ensure the comparability of the results.

How Are the Results Reported?


The reports contain rankings of all participants and jurisdiction-by-jurisdiction comparisons of mathematics and science achievement overall and for each content area; comparisons of performance against international benchmarks (see below); and gender differences in performance. The achievement data are accompanied by extensive questionnaire data about the home, classroom, school, and jurisdictional contexts within which mathematics and science learning take place. In some cases results for the Benchmarking participants are reported in comparison to the results for all TIMSS 1999 countries, and in other cases in comparison to the results for selected reference countries, comprised of the United States as well as a dozen European and Asian countries of interest. These include several high-performing
European countries (Belgium (Flemish), the Czech Republic, the Netherlands, and the Russian Federation), countries that are major economic trading partners of the United States (Canada, England, and Italy), and the top-scoring Asian countries of Chinese Taipei, Hong Kong SAR, Japan, the Republic of Korea, and Singapore.

In addition to reporting achievement as scale scores, student performance is also described in terms of international benchmarks of performance. In order to provide meaningful descriptions of what performance on the achievement scale could mean in terms of the mathematics and science that students know and can do, TIMSS identified four points on the scale – Top 10%, Upper Quarter, Median, Lower Quarter (90th, 75th, 50th, and 25th percentiles, respectively) – for use as international benchmarks, and conducted an ambitious scale-anchoring exercise to describe performance at these benchmarks. The percentage of students in each jurisdiction that reached each benchmark is reported. The benchmark descriptions are accompanied by example test items illustrating student performance at each benchmark.

What Publications and Resources Are Available?

In addition to the two Benchmarking Study reports, the results for the 38 countries participating in TIMSS 1999, including those for the United States, were reported in December 2000 in two companion reports, the TIMSS 1999 International Mathematics Report and the TIMSS 1999 International Science Report. Performance in the United States relative to that of other nations was reported by the National Center for Education Statistics of the U.S. Department of Education in Pursuing Excellence. The TIMSS 1999 and Benchmarking Study publications and resources are listed below.

Released April 4, 2001:

- Science Benchmarking Report: TIMSS 1999 – Eighth Grade
- TIMSS 1999 Mathematics Released Item Set
- TIMSS 1999 Science Released Item Set

Released December 5, 2000:

- Pursuing Excellence: Comparisons of International Eighth-Grade Mathematics and Science Achievement from a U.S. Perspective, 1995 and 1999
TIMSS 1999 International Mathematics Report
TIMSS 1999 International Science Report
TIMSS 1999 Technical Report

For future release:
TIMSS 1999 Benchmarking Technical Report
TIMSS 1999 Benchmarking Database and User Guide
TIMSS 1999 International Database and User Guide

All the above are published by the International Study Center at Boston College, with the exception of the U.S. national report Pursuing Excellence, published by the National Center for Education Statistics.

When Is the Next TIMSS?

TIMSS 1999 was the second phase of a long-term study designed to measure trends in mathematics and science achievement, much like the regular cycle of national assessments in the U.S. conducted by the National Assessment of Educational Progress (NAEP). Work has begun on TIMSS 2003, which will assess students in grades 4 and 8. A trend study such as this has the potential for affecting policy and practice by investigating the effects on achievement of efforts in educational improvement.

Who Conducted the TIMSS 1999 Benchmarking Study?

The TIMSS 1999 Benchmarking Study was a shared venture. In conjunction with the Office of Educational Research and Improvement (OERI) of the U.S. Department of Education and the National Science Foundation (NSF), the National Center for Education Statistics (NCES) worked with the International Study Center (ISC) at Boston College to develop the study.

The TIMSS studies are conducted under the auspices of the International Association for the Evaluation of Educational Achievement (IEA), an independent cooperative of national and governmental research agencies with a permanent secretariat based in Amsterdam, the Netherlands. Its primary purpose is to conduct large-scale comparative studies of educational achievement to gain a deeper understanding of the effects of policies and practices within and across systems of education.
The IEA delegated responsibility for the overall direction and management of TIMSS 1999 to the International Study Center in the Lynch School of Education at Boston College. In carrying out the project, the International Study Center worked closely with the IEA Secretariat, Statistics Canada in Ottawa, the IEA Data Processing Center in Hamburg, Germany, and Educational Testing Service in Princeton, New Jersey. Westat in Rockville, Maryland, was responsible for sampling and data collection both for the Benchmarking Study and the U.S. component of TIMSS 1999.

Funding for the overall design, administration, data management, and quality assurance activities of the TIMSS 1999 Benchmarking Study was provided by NCES, NSF, and OERI. Each Benchmarking participant contracted directly with Boston College to fund data collection activities in its own jurisdiction. Funding for the international coordination of TIMSS 1999 was provided by NCES, NSF, the World Bank, and participating countries. Each participating country was responsible for funding local project costs and implementing TIMSS 1999 in accordance with the international procedures.

TIMSS 1999 Benchmarking is a project of the International Study Center
Boston College, Lynch School of Education