NCEE 2012-4008

U. S. DEPARTMENT OF EDUCATION

Evaluation of the Effectiveness of the Alabama Math, Science, and Technology Initiative (AMSTI)

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

Full-length report located here:
http://ies.ed.gov/ncee/edlabs/projects/project.asp?ProjectID=69

Evaluation of the Effectiveness of the Alabama Math, Science, and Technology Initiative (AMSTI)

Final Report

February 2012

Authors:

Denis Newman, Empirical Education Inc., Co-Principal Investigator

Pamela B. Finney, Research Management Leader, Regional Educational Laboratory Southeast, Co-Principal Investigator

Steve Bell, Abt Associates Inc., Task 2 Methodology Leader

Herb Turner, ANALYTICA, Technical Consultant

Andrew P. Jaciw, Empirical Education Inc., Director of Experimental Design and Analysis

Jenna L. Zacamy, Empirical Education Inc., Senior Research Manager

Laura Feagans Gould, Academy for Educational Development, Senior Research Scientist

Project Officer:
Sandra Garcia, Institute of Education Sciences

NCEE 2012-4008

Richard Sawyer, Academy for Educational Development, was the original co-principal investigator.

Summary

Partly motivated by the 1996 National Assessment of Educational Progress scores, which were below the national average for Alabama's grade 4–8 students in mathematics and grade 8 students in science, the Alabama State Department of Education (ALSDE) developed a statewide initiative to improve mathematics and science teaching and student achievement in kindergarten through grade 12 (K–12). The Alabama Math, Science, and Technology Initiative (AMSTI) is a two-year intervention intended to better align classroom practices with national and statewide teaching standards—and ultimately to improve student achievement—by providing professional development, access to materials and technology, and in-school support for teachers.

AMSTI, a schoolwide intervention, was introduced in a set of 20 schools in 2002. Each year since then, the state has rolled out the program to additional schools within its 11 regions.[2] By 2009, about 40 percent of the state's 1,518 schools were designated as AMSTI schools. Funding for the program from the state legislature was $46 million in 2009.[3]

Given the policy relevance and level of investment in AMSTI, the Regional Educational Laboratory Southeast mounted a longitudinal, cluster randomized controlled trial to determine the effectiveness of AMSTI in grades 4–8, as implemented in five regions in the state. Previous evaluations of the program's effects on students in K–12 did not use randomized controlled trials. The most recent evaluation (Miron and Maxwell 2007) reported that students in grade 5 in AMSTI schools outperformed students in non-AMSTI schools in mathematics, science, and reading and students in grade 4 in AMSTI schools outperformed their counterparts in non-AMSTI schools in reading only. These evaluations used a study design that compared school-level test scores of AMSTI schools with non-AMSTI schools in the same district but did not establish preintervention comparability. This study's randomized controlled trial design improves on previous evaluations because it eliminates selection bias and establishes the preintervention comparability of the two groups.

The AMSTI theory of action posits that in order to improve student achievement, teacher instructional strategies should include higher levels of hands-on, inquiry-based instruction. The three components of the program that foster this type of instruction are comprehensive professional development delivered through a 10-day summer institute and follow-up training during the school year; access to program materials, manipulatives, and technology needed to deliver hands-on, inquiry-based instruction; and in-school support by AMSTI lead teachers and site specialists who offer mentoring and coaching for instruction. The full program is delivered over the course of two years. In each region, AMSTI site specialists partner with a local university or college. ALSDE oversees the professional development and implementation of the program.

---

[2] Alabama has 11 regional inservice centers (RICs), which were established by the Alabama Legislature in 1984 to provide "rigorous inservice training in critical needs areas for the state's public school personnel." The 11 AMSTI regions follow the same boundaries as the RICs.
[3] Accessed from the web on May 8, 2010 (http://www.alsde.edu/general/quick_facts.pdf).

The AMSTI theory of action provided the theoretical basis for selecting the research questions addressed in this report. The primary confirmatory analyses address the effect of AMSTI on student achievement in mathematics problem solving and science after one year. These outcomes, though related and expected to be positively correlated, are from different content domains. The primary research question looks at whether the intervention had an effect on mathematics problem solving or science knowledge.[4]

The secondary research question addresses the effect of AMSTI on classroom practices, which are the mediating link between the intervention components and student achievement. The effect of AMSTI on classroom practices is measured by a composite variable of teacher self-reported time (in minutes) using hands-on instruction, inquiry-based instruction, and instruction promoting student use of higher-order thinking skills. This composite "active learning" score was computed separately for mathematics and science instruction. As the initiative may be successful at increasing active learning instruction for one subject area but not the other, the study examines whether the intervention had an effect on either domain—active learning instruction in mathematics or active learning instruction in science.

The study addresses the following confirmatory research questions:

Primary confirmatory research question: effects on student achievement after one year

- What is the effect of AMSTI on:
  a. student achievement in mathematics problem solving *after one year*?
  b. student achievement in science *after one year*?

Secondary confirmatory research question: effects on classroom practice after one year

- What is the effect of AMSTI on:
  a. the use of active learning instructional strategies by mathematics teachers *after one year*?
  b. the use of active learning instructional strategies by science teachers *after one year*?

The study also addresses the following exploratory research questions:

Exploratory research question: effects on student achievement after two years

- What is the effect of AMSTI on:
  a. student achievement in mathematics problem solving *after two years*?
  b. student achievement in science *after two years*?

[4] The decision to examine these as separate outcomes was further warranted by the program design elements (see table 1.1 in chapter 1). During the professional development, trainers use content- and grade-specific instructional methods; there are separate mathematics and science specialists; and separate curriculum modules.

Exploratory research question: effects on student achievement in reading after one year

- What is the effect of AMSTI on student achievement in reading *after one year*?

Exploratory research questions: effects on teacher content knowledge and student engagement after one year

- What is the effect of AMSTI on:
  a. mathematics teachers' reported level of content knowledge *after one year*?
  b. science teachers' reported level of content knowledge *after one year*?
- What is the effect of AMSTI on:
  a. mathematics teachers' reported level of student engagement *after one year*?
  b. science teachers' reported level of student engagement *after one year*?

Exploratory research questions: variations in effects on student achievement for specific subgroups of students after one year

- Do the one-year effects of AMSTI on student achievement in (a) mathematics problem solving, (b) science, and (c) reading vary by student pretest scores? What is the effect of AMSTI on these outcomes after one year for students with pretest scores that fall in the low, middle, and high ranges?
- Do the one-year effects of AMSTI on student achievement in (a) mathematics problem solving, (b) science, and (c) reading vary by low-income status, proxied by enrollment in the free or reduced-price lunch program (as part of the National School Lunch Program)? What is the effect of AMSTI on these outcomes after one year for students enrolled in the free or reduced-price lunch program and students who are not enrolled?
- Do the one-year effects of AMSTI on student achievement in (a) mathematics problem solving, (b) science, and (c) reading vary by racial/ethnic minority status? What is the effect of AMSTI on these outcomes after one year for racial/ethnic minorities and for White students?
- Do the one-year effects of AMSTI on student achievement in (a) mathematics problem solving, (b) science, and (c) reading vary by gender? What is the effect of AMSTI on these outcomes after one year for boys and for girls?

Although AMSTI is a two-year program, the confirmatory analyses address the effect of the program after the first year. The effect of AMSTI after the full intervention was implemented (that is, after two years) cannot be estimated without additional assumptions because, as detailed in chapter 2, the control group entered the program after its first year and was no longer a pure control group in the second year. Researchers selected an appropriate method to estimate the two-year effects; however, the necessity of additional assumptions makes the analyses exploratory rather than confirmatory. This limitation on the study's design means that only the one-year effect on mathematics problem solving and science can be considered confirmatory.

Beyond the two-year impacts, the exploratory questions pertain only to the first year of AMSTI. Unlike questions concerning two-year effects, they can be answered without additional assumptions necessitated by the entry of the control group into the AMSTI program in the second year of the study. These analyses address the effect of AMSTI on student achievement in reading, teacher content knowledge, student engagement, and variations in effects on student achievement for particular subgroups of students. These questions are important to understanding the full effects of AMSTI and in potentially identifying ways to improve the program. The rationale for selecting these questions arises from several sources: the AMSTI theory of action; interest from program developers; prior research on AMSTI and within the fields of science, technology, engineering, and mathematics; and measured state achievement gaps.

The study took advantage of ALSDE's rollout of AMSTI to specific regions during the study years. To participate in the study, schools must have housed at least one grade between grades 4 and 8, and at least 80 percent of a school's mathematics and science teachers must have agreed to participate. From the eligible schools that applied to the program, researchers made a purposeful effort to select a sample that was representative of the population of schools in the regions involved. Pairs of similar schools were selected from the pool of applicants based on similarity in mathematics achievement, the percentage of minority students, and the percentage of students from low-income households. Within each pair, schools were randomly assigned either to the AMSTI condition, in which teachers received AMSTI training and program materials, or to the control condition, in which teachers used their existing mathematics and science programs.

Because Alabama did not plan to introduce the program in the number of schools required by the experiment in one year, the experiment combined two "subexperiments", one starting in 2006 and the other starting in 2007. The full sample combined the two samples from the two "subexperiments" and included 82 schools, with about 780 teachers and 30,000 students in grades 4–8 across the two subexperiments.[5] In Subexperiment 1, the first set of 40 schools (within three regional AMSTI sites) was randomized to conditions in the winter of 2006. In Subexperiment 2, the second set of 42 schools (within two regional AMSTI sites) was randomized to conditions in the winter of 2007. To estimate the effects of AMSTI after one year (confirmatory analysis), data from both subexperiments were pooled and analyzed together after their respective first year. The integrity of the samples used in the confirmatory analysis was maintained, because the difference in attrition between the intervention and control groups was less than 5 percentage points and overall attrition was 2.5 percent or less for all outcomes. To estimate the effects of AMSTI after two years, data from both subexperiments were pooled and analyzed together after the respective second year.

Data were collected at multiple levels. Sources included classroom rosters, student achievement and demographic data, professional development training logs and observations, professional development teacher surveys, interviews with teachers and principals, classroom observations, and web-based surveys of teachers and principals.[6] In both subexperiments,

---

[5] This number represents the approximate number of teachers and students in the 82 study schools during years 1 and 2 of Subexperiment 1 and Subexperiment 2. For precise numbers of teachers and students used in each analysis, see chapter 2.

[6] Training logs, in-person interviews with teachers and principals, and classroom observations were conducted only with Subexperiment 2, because researchers did not receive approval from the Office of

teachers in AMSTI schools were trained in the program the summer following randomization and before their first year of implementation (2006/07 for Subexperiment 1, 2007/08 for Subexperiment 2).

Inferential tests on web-based teacher survey data were conducted to examine the differences between AMSTI and control schools in the presence of the three main intervention components (summer professional development, access to materials and manipulatives, and in-school support). AMSTI teachers were more likely to have participated in summer professional development than were control teachers (87 percent versus 24 percent for mathematics teachers, 84 percent versus 24 percent for science teachers). AMSTI teachers also reported having greater access to materials than did control teachers (78 percent versus 41 percent for mathematics teachers, 61 percent versus 33 percent for science teachers). AMSTI teachers were more likely to receive in-school support than were their control counterparts (59 percent versus 40 percent for mathematics teachers, 65 percent versus 25 percent for science teachers). All these differences were statistically significant at $p < .05$ (for specifics see chapter 3).[7]

The effect of AMSTI on student achievement in mathematics after one year, as measured by end-of-the-year scores on the Stanford Achievement Test Tenth Edition (SAT 10) mathematics problem solving assessment of students in grades 4–8, was 2.06 scale score units (figure 1). The difference of 0.05 standard deviation in favor of AMSTI schools is equivalent to a gain of 2 percentile points on the SAT 10 mathematics problem solving assessment for the average control group student had the student received AMSTI. The 0.05 standard deviation is statistically significant but smaller than the effect the research team believed would be detectable by the experiment as designed. Whether this size effect is educationally important is an open question. It may be useful to convert this effect into a more policy-relevant metric—additional student progress measured in days of instruction. In these terms, the average estimated effect of AMSTI was equivalent to 28 days of additional student progress over students receiving conventional mathematics instruction.[8] The effect of AMSTI on student achievement in science,

---

Management and Budget in time to collect these implementation data during the 2006/07 school year for Subexperiment 1. Student-level data and web-based survey data from teachers and principals were collected for Year 1 of Subexperiment 1 through a research grant (IES: #R305E040031) from the Institute of Education Sciences to Empirical Education Inc., with permission from the IES program officer.

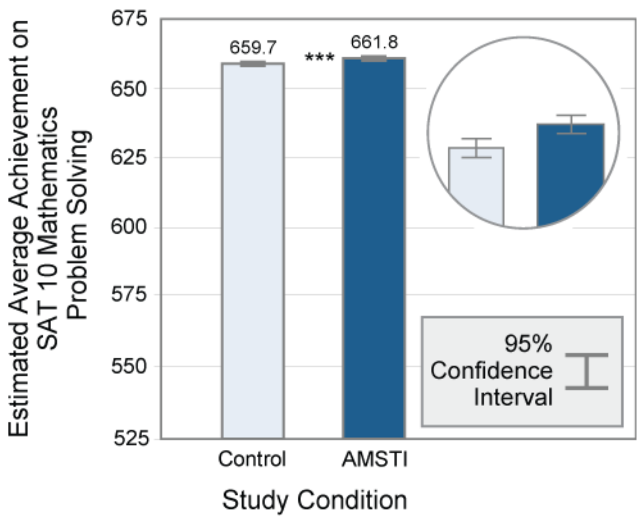
[7] The implementation analyses presented in this report aim simply to describe program implementation for each program component. The study design did not include assessment and analysis of the AMSTI implementation quality since objective benchmarks for AMSTI implementation do not exist.

[8] To obtain this value, we express the estimated average score gain in the treatment group as a proportion of the score gain in the control group (T=treatment, C=control):

$$\frac{\overline{\hat{Y}}_{T(post)} - \overline{\hat{Y}}_{T(pre)}}{\overline{\hat{Y}}_{C(post)} - \overline{\hat{Y}}_{C(pre)}} = \frac{\overline{\hat{Y}}_{C(post)} - \overline{\hat{Y}}_{C(pre)} + \hat{T}}{\overline{\hat{Y}}_{C(post)} - \overline{\hat{Y}}_{C(pre)}} = 1 + \frac{\hat{T}}{\overline{\hat{Y}}_{C(post)} - \overline{\hat{Y}}_{C(pre)}}$$

. We then multiply this value by 180 (assuming a 180-day school year in Alabama) which yields the estimated projected number of days of schooling by the control group, had they been in the treatment condition. Subtracting 180 from this as measured by end-of-the-year scores on the SAT 10 science assessment, required only in grades 5 and 7, was not statistically significant after one year (figure 2).

Figure 1 Effect of the Alabama Math, Science, and Technology Initiative (AMSTI) on Stanford Achievement Test Tenth Edition (SAT 10) mathematics problem solving achievement after one year



** Significant at $p < .05$; *** Significant at $p < .01$
Note: n = 82 schools; 18,713 students
Source: Student achievement data from tests administered as part of the state's accountability system.

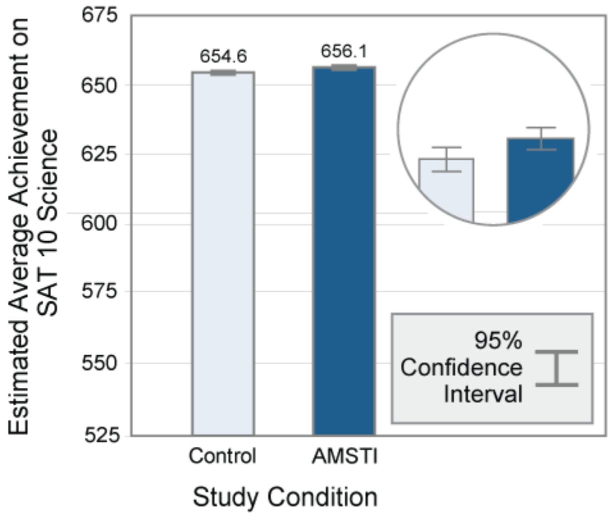
---

quantity yields an estimate of the treatment effect in terms of additional learning growth as translated into additional days of

schooling: $$IMPACT = \left(1 + \frac{\hat{T}}{\hat{\overline{Y}}_{C(post)} - \hat{\overline{Y}}_{C(pre)}}\right) \times 180 - 180 = \left(\frac{\hat{T}}{\hat{\overline{Y}}_{C(post)} - \hat{\overline{Y}}_{C(pre)}}\right) \times 180 = 28 days.$$

This calculation assumes that the treatment effect accumulates linearly with time over the course of a grade. A formal test of the linearity of the accrual of the treatment effect was not conducted. If the treatment effect does not accrue linearly then this extrapolation of the number of days may not be accurate.

Figure 2 Effect of the Alabama Math, Science, and Technology Initiative (AMSTI) on Stanford Achievement Test Tenth Edition (SAT 10) science achievement after one year



Note: *n* = 79 schools; 7,528 students
Source: Student achievement data from tests administered as part of the state's accountability system.

AMSTI also had a positive and statistically significant effect on classroom practices in mathematics and science after one year. Based on multiple surveys in which teachers reported the number of minutes of active learning strategies used during the previous 10-day period, AMSTI mathematics teachers averaged 49.83 more minutes, and AMSTI science teachers averaged 40.07 more minutes than control teachers. These estimated effects are equivalent to 0.47 standard deviation in mathematics and 0.32 standard deviation in science. Although teachers in both the AMSTI and control groups reported using active learning instructional strategies, teachers in AMSTI schools reported spending more time engaged in this type of instruction.

The exploratory investigation of the two-year effect of AMSTI on student achievement on the SAT 10 mathematics problem solving test found a positive and statistically significant result of 3.74 scale score units. This effect represents a difference of 0.10 standard deviation in favor of AMSTI schools, equivalent to a gain of 4 percentile points for the average control group student had the student received AMSTI for two years. This estimate of the average effect of AMSTI after two years can be translated into an estimated 50 days of additional student progress over students receiving conventional mathematics instruction.

The exploratory investigation of the two-year effect of AMSTI on student achievement in science also found a statistically significant result, with a magnitude of 4.00 scale score units. This effect represents a difference of 0.13 standard deviation in favor of AMSTI schools, equivalent to a gain of 5 percentile points for the average control group student had the student received AMSTI for two years.[9]

The effect of AMSTI on student achievement in reading after one year, as measured by end-of-the-year scores on the SAT 10 reading assessment of students in grades 4–8, was 2.34 scale score units. The statistically significant difference of 0.06 standard deviation in favor of AMSTI schools is equivalent to a gain of 2 percentile points on the SAT 10 reading assessment for the average control group student had the student received AMSTI. This difference can be translated into an estimated 40 days of additional student progress over students receiving conventional reading instruction.

The effect of AMSTI on teacher-reported content knowledge after one year was not statistically significant in either mathematics or science. AMSTI did have a positive and statistically significant effect on student engagement after one year, measured on a 5-point scale ranging from “not engaged” to “fully engaged.” AMSTI teachers were more likely than control teachers to rate their students as achieving higher levels of engagement.

An exploration of the differential effects of AMSTI on student achievement for subgroups of students found no statistically significant differential effects on student achievement in mathematics or science based on racial/ethnic minority status, eligibility for free or reduced-price lunch, gender, or pretest level. In reading, however, AMSTI had a statistically significant differential effect for minority and White students of 3.04 scale score points ($p < .001$). This difference can be translated into days of student progress, where progress is measured as the average gain in test scores over the course of the school year by the control group using conventional reading instruction. In this metric, White students in AMSTI made an estimated 52 more days of progress than minority students in AMSTI. The effect of AMSTI on reading achievement for minority students was not statistically significant ($p = .294$); for White students, there was a statistically significant positive effect of AMSTI on reading achievement ($p < .001$).

[9] The analysis of the two-year impact of AMSTI on student achievement is exploratory. Readers should exercise caution in interpreting the results. For instance, we remind the reader that with exploratory analyses we do not perform multiplicity adjustments. As a consequence, a less strict criterion is used with exploratory analyses for deciding whether a particular result achieves statistical significance, with the drawback that it increases the probability of finding a spurious impact. For the two-year impact on mathematics problem solving ($p = .030$) and science ($p =.038$) the results reach statistical significance under the less strict criterion (alpha = .05). Under the more strict criterion used with the primary confirmatory analyses (alpha = .025) these results would not have been considered statistically significant.