Abstract Title Page
Not included in page count.

Individual *paper* proposal

Title: Locating Differential Effectiveness of a STEM Initiative Through Exploration of Moderators

Authors and Affiliations:

Denis Newman, Empirical Education Inc. dn@empiricaleducation.com
Boya Ma, Empirical Education Inc. bma@empiricaleducation.com
Andrew Jaciw, Empirical Education Inc. ajaciw@empiricaleducation.com

Sections:
1. Classroom Instruction and Context, or
2. Research Methods.

Abstract Body

Limit 4 pages single-spaced.

Background / Context:

Description of prior research and its intellectual context.

In randomized control trials, such as the one we conducted in Alabama, moderators are potentially an important source of variation in treatment effects. Whether moderators used in the RCT's analysis are pre-existing characteristics of students, their teachers, or the school or other organizational context, they can point us toward differences in the way the program (treatment) is implemented, the way teachers interact with different kinds of students when implementing the program, or how effects are operating at levels in the organization above the level where the differences were measured. We appreciate Cronbach's (1975) metaphor of the "hall of mirrors" to point to the endless possibilities for higher-order interactions complicating the explanation of what was found. But the goal of this paper is to illustrate a disciplined approach to exploring important interactions guided by our interest in narrowing the working hypotheses as to how the program is having its effects and thus guiding practitioners, program developers, and policy makers in making the most effective use of STEM programs.

The RCT from which this paper draws is complete and the reporting finalized except for the final approvals by IES but as of the deadline for this abstract, the report has not been made public. Therefore, this abstract cannot provide the actual results or details on the specific moderators on which it will focus. We feel we can disclose that the results will provide sufficient substance for the analyses that we will focus on in the paper. In addition, explorations that are not reported in the IES report will be included in this paper.

The Alabama Math Science and Technology Initiative (AMSTI) is a major program developed and deployed in Alabama schools. 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 K-12 schools. AMSTI, a school-wide 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. 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. 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.

The RCT measured the impact of AMSTI. Key outcomes included student achievement in mathematics problem solving, science, and reading; the use of active learning instructional strategies in mathematics and science classrooms; and teacher content knowledge and student engagement in mathematics and science.

Purpose / Objective / Research Question / Focus of Study:

To reflect the theme of SREE Spring 2012 we will focus on moderator analyses. In the first part

of the proposed paper we will report estimates of differential impact as well as impacts for subgroups. In the second part of the paper we will discuss these results and consider how extensions of the basic moderator analyses may be used to understand better the differential impacts and how such exploratory work may inform both educational practice and refinement of the program to make its benefits more widespread. This approach – to first present results of moderator analyses, and then discuss next stages of analyses for helping with interpretation of treatment heterogeneity for informing program adaptation – we believe fits into the conference theme *Understanding Variation in Treatment Effects.*

Setting:
The RCT was conducted in public schools in Alabama beginning in the 2006/07 school year.

Population / Participants / Subjects:
In all, 82 schools, approximately 780 teachers of math and science, and 30,000 upper-elementary and middle school students in five regions in Alabama participated in the study.

Intervention / Program / Practice:
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 developers posited that teacher quality and effectiveness were the keys to improving student test scores in mathematics and science. The AMSTI model is based on the hypothesis that intensive, comprehensive professional development; in-school support (e.g., teacher coaching provided by technical assistance staff); and associated resources and materials (e.g., curricular materials, manipulatives, and microscopes) will lead to teachers' use of effective instructional strategies that are aligned with statewide and national standards. These changes in instructional strategies were hypothesized to lead to improved student achievement in mathematics and science. Specifically, the program includes comprehensive professional development delivered through a 10-day summer institute and follow-up training during the school year; in-school support by AMSTI lead teachers and site specialists who offer mentoring and coaching for instruction; and access to program materials, manipulatives, and technology needed to deliver hands-on, inquiry-based 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.

Research Design:
This was a cluster randomized trial. 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 teachers must have agreed to participate in AMSTI. 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. Data from both subexperiments were pooled and analyzed together. This paper examines data only from the first year of each subexperiment. The statistically unbiased estimates of the effect of AMSTI were generated under authentic conditions in volunteer schools. The study did not alter implementation specifically for the experiment but followed schools as they participated in the standard initiative.

Data Collection and Analysis:

Data were collected at multiple levels. Sources included:

- Classroom rosters, student achievement and demographic data were collected from the district and state-level. These sources were used to estimate the effect of AMSTI on student achievement
- Web-based surveys of teachers and principals. Teacher surveys were administered four times.

The key outcome measures included:

- Student achievement in mathematics problem solving, science and reading as measured by the SAT 10 assessment;
- The use of active learning instructional practices in mathematics and science classrooms, as measured by a composite variable of teacher self-reported time.

For this paper, analysis will focus on student outcomes. The key moderators of interest are at the student level. These include

- Pretest: For the math outcome we used both the math and the reading pretest. For the science outcome we used reading since there was not a science pretest available.
- SES: based on free/reduced lunch status
- Minority status: This was coded as White or Minority. Minority consisted predominantly of Black students with a small number of other ethnicities. (English learners did not have a significant presence)
- Gender

Findings / Results:

Part 1: Findings for Moderator Effects:

As noted above, because the report has not been released by IES, this abstract will not be able to provide the actual results or indicate the outcomes for which differential effects were found. The report is the final stages of review and is expected to be released in the next two months. For SREE we will report estimates of the interactions between the indicator of treatment assignment and each of the moderators for each outcome listed in the previous section. In addition, for context, the average impact findings will be reported. This paper will address only the results after one year of AMSTI implementation.

Part 2: Extensions

For SREE we will take the analyses several steps further and investigate both the robustness of the interactions and whether they are attributable to individual or compositional effects. This effort is consistent with the priority of moving educational programs through cycles of continuous improvement: it is important investigate *AMSTI* in terms of its processes and effects

so that it can be better understood and further developed to ultimately produce stronger positive effects for a broader range of students.

(1) *Robustness of interactions with random slopes for moderators.* For group randomized trials it is common to model a random effect at the level of randomization. As a result, the variance component at that level figures into the standard error of the impact estimate. For moderator analyses that examine differences in the impact for different subgroups of students, the main effect of the moderator is often represented as a fixed effect. We will examine whether there is sufficient variation across schools in the difference in performance between subgroups to warrant modeling the difference as randomly varying across schools. Ignoring this variation in moderator analyses is akin to ignoring the differences in average performance among schools and the corresponding design effect in estimates of average impact. Where variance in the subgroup performance differential is statistically significant, we will consider whether the effect is large enough to alter conclusions about differential impacts.

(2) *Decomposing the interactions into within- and between-school components.* Compositional effects (also known as contextual effects) "are said to occur when the aggregate of a person level characteristic…is related to the outcome…even after controlling for the effects of individual level characteristics…" (Raudenbush and Bryk, 2002). Similarly, we will consider whether a moderating effect of a student characteristic persists even after we include the moderating effect of the school aggregate of the characteristic in the same model. The purpose of these analyses is to determine whether moderating effects are due to organizational differences across schools or individual differences within schools. This informs program improvement efforts by identifying the potential sources of the differential effects. The implications of such analyses for the continuing improvement of AMSTI are large: if the effects of AMSTI are moderated by school-average characteristics this suggests targeting school-level malleable factors, such as organizational processes and practices, to enhance the impact of AMSTI. On the other hand, if the variation is accounted for by attributes of students, it suggests supplementing AMSTI so that it works better for students with certain attributes. We will report the results and discuss the implications.

(3) *Comparing multiple-interaction versus single-interaction models.* Modeling more than one moderator effect simultaneously allows us to address the question of whether treatment varies by levels of one posited moderator after adjusting for the effects of other posited moderators. For example, we can examine whether there is a differential effect of treatment depending on minority status after controlling for differences in impact associated with SES. Exploration that examines differences in impact across levels of one moderator within a stratum of another moderator allows us to develop more-precise theory of the locus of differential impact, and gives the developer information about where to focus program improvement efforts. We will report the results of several such analyses.

Conclusions:

This abstract does not outline our conclusions given our inability to report our findings before IES releases the full report. However, we will address the policy implications of our findings as well as the implications for developers and implementers of AMSTI and similar programs. We hope to illustrate the importance of examining moderator effects while carefully exploring the locus of the moderating process.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologis*t, 116-127.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage.

Appendix B. Tables and Figures

Not included in page count.