Title: Uncovering the Black Box: Exploratory Mediation Analysis for a Science Teacher Professional Development Program

Authors and emails:

Thanh Nguyen, tnguyen@empiricaleducation.com
Andrew Jaciw, ajaciw@empiricaleducation.com
Jenna Zacamy, jzacamy@empiricaleducation.com

Conference Section:
Academic Learning in Education Settings
Teachers and School Leaders

Keywords:
Mediation
Educational evaluation
Classroom instruction

Background/Context

Randomized Control Trials (RCTs), often considered to be the gold standard in education research, if conducted properly, yield unbiased causal estimates for education programs and policies. However, causal impact findings alone tell us little about how the inputs led to the outcomes, whether our theorized causal pathways were proven or falsified, or how the program could be improved. A logic model (LM), which is a visual representation of a program's theoretical underpinnings, serves as the series of hypotheses of how inputs relate to outcomes. However, without empirically testing critical junctures of the LM, there is no empirical basis on which to identify areas of the LM that are supported or challenged, and to help program developers make informed decisions about program improvement.

Purpose/Objective/Research Question

Our objective is to uncover the black box for a science teacher PD program that aims to improve teacher content knowledge (TCK) and student science achievement. We ask the following research questions:

- What is the impact of the program on the intermediate outcomes, including teacher beliefs and attitudes, content knowledge, opportunity to learn (OTL), and school climate after two years of the program's implementation for participants in schools assigned to receive the intervention relative to those assigned to Business-As-Usual (BAU)?
- To what extent do these intermediate outcomes mediate the relationship between the program and student outcomes? To what extent do the results support and/or challenge the LM?

Setting

The study took place in 66 elementary schools serving low-income students (an average of 75% of students were eligible for the Free and Reduced Price Lunch program among schools in the sample), across seven school districts and two states.

Population/Participants/Subjects

The study includes over 2000 students, 300 upper elementary school teachers, and 60 administrators.

Intervention/Program/Practice

The program is a model for teacher PD aimed at raising students' science achievement through improving science instruction. The PD model focuses on the critical connections between science understanding, classroom practices, and literacy to support the implementation of the Next Generation Science Standards (NGSS). PD activities for teachers include summer PD and school-year professional learning communities for two

years. The model also builds capacity for school administrators, teacher leaders, district staff, and regional partners for supporting science instruction.

Research Design

Sixty-six schools were randomly assigned to either the intervention or to BAU. Impacts were assessed on intermediate and final outcomes after two years using standard methods for analyzing impacts from cluster randomized trials.

Data Collection and Analysis

For students, we administered a science achievement assessment and a survey of non-academic outcomes (e.g. aspirations for careers in science, enjoyment of science, self-efficacy around science) at the end of the two-year study. For teachers, we administered a CK assessment at baseline and again at the end of the study. We also administered surveys to teachers six times across the two years and to administrators in the spring of each year. We collected student demographic and achievement data from school districts.

After assessing the impact on the confirmatory outcomes, we analyze the impact on the intermediate outcomes. We use multi-level Hierarchical Linear Models (HLMs) with intermediate outcomes regressed against baseline covariates, a dummy variable indicating treatment assignment at the cluster level, and random effects at the appropriate levels, depending on whether the intermediate outcome is measured at the student, teacher, or school-level. Evaluating first stage impacts will show us where to apply further mediation analysis.

We also create indices from the teacher surveys for constructs of OTL, school climate, and teacher attitudes and beliefs related to science instruction. At the outset of the study, the program developers conjectured mediators of impact on student achievement. We are interested in assessing impacts on the core constructs and their role in mediation. However, we are also exploring impacts on sub-constructs identified through exploratory and confirmatory FA of the main scales, potentially yielding more fine-grained solutions. The resulting product will be an LM on which we overlay the impact findings for distal and intermediate outcomes (Figure 4). This will provide a visual aid for “back mapping” the LM to unpack confirmatory impact findings.

Findings/Results

We obtained preliminary results for the confirmatory impact on student achievement and for intermediate outcomes to date. For student achievement, we observe an impact estimate of 0.076 standard deviations (p-value: .248). For the intermediate outcome of TCK, we observe an effect of 0.355 standard deviations (p = .092) (Figure 4).

For the Instruction construct of OTL, results of an exploratory factor analysis suggest a three-factor solution: 1) Hands-on work, 2) Literacy and discourse, and 3) Science ideas. The distribution of the factor-based scores for the three factors are shown in Table 1.The difference in means between treatment and control are 0.195 (Factor 1) ($p<.10$), 0.372 (Factor 2) ($p<.05$), and 0.176 (Factor 3) ($p<.10$) based on a Tobit analysis. We are in the process of obtaining cluster-adjusted Tobit results.

By the conference, we will have additional results for the impact on the main distal outcomes on students: science achievement on a constructed-response component of the student assessment and non-academic outcomes. Additionally, we will populate impact findings on the remaining intermediate outcomes. Given the no impact finding on student science achievement, we are especially interested in identifying any "inconsistent mediation models" that have at least one mediated effect having the opposite sign to and therefore counteracting the other mediated or direct effects in the model (MacKinnon, 2012).

Conclusions

Through this approach of systematically testing the LM and overlaying results, we hope to bring into focus more plausible causal pathways that program developers could leverage in strengthening the program, while also drawing attention to theorized pathways that have weak or no empirical evidence of a clear relationship and that challenge program developers to consider areas for improvements or further research.

References

MacKinnon, D. (2012). *Introduction to statistical mediation analysis*. New York and London: Routledge.

Tables and figures

Figure 1. Program logic model and selected results

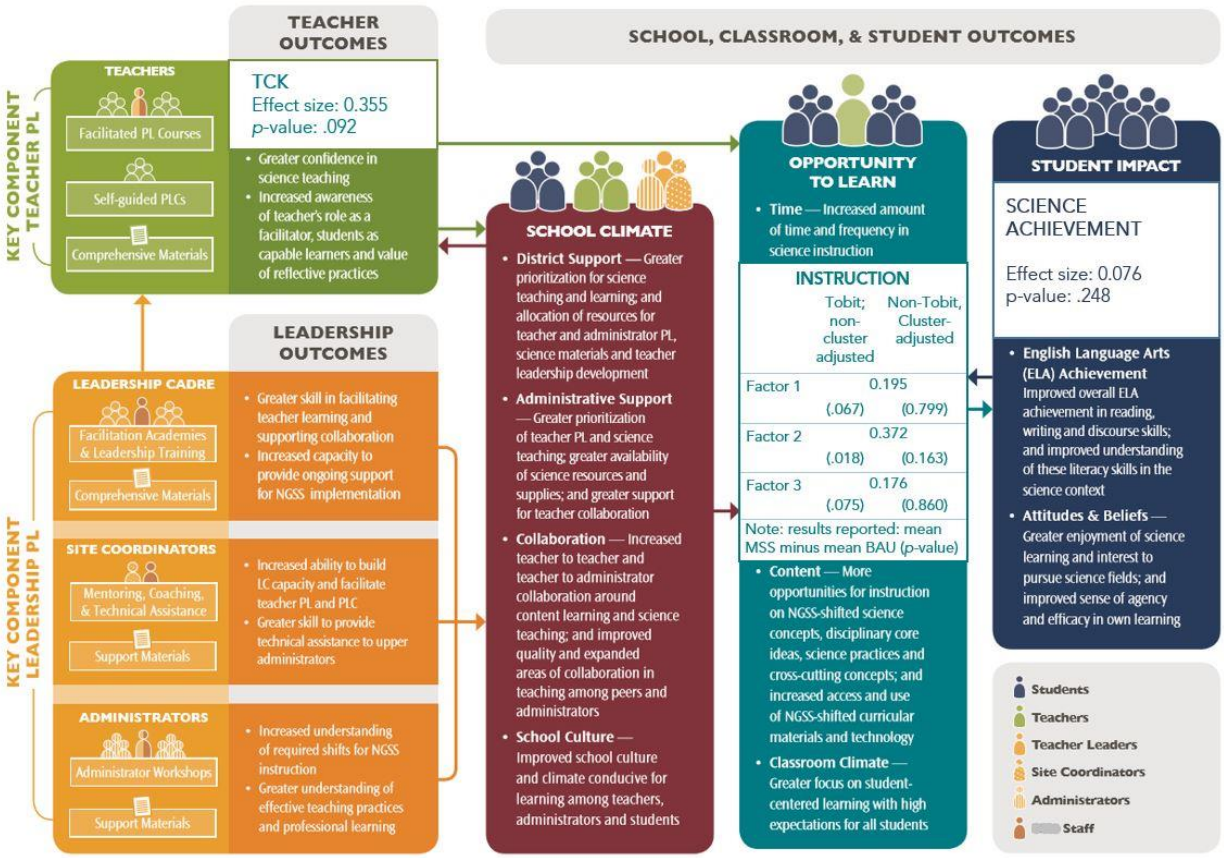

Table 1. Distribution of factor-based scores for the OTL Instruction construct

Factor1	Factor2	Factor3
Hands-on work	Literacy and Discourse	Science Ideas

	All	Treatment	BAU	All	Treatment	BAU	All	Treatment	BAU
N	137	76	61	137	76	61	137	76	61
missing	10	5	5	10	5	5	10	5	5
min	0	0	0	0	0	0	0	0	0
max	5	5	5	5	5	5	5	5	5
mean	3.00	3.06	2.93	2.69	2.85	2.50	3.08	3.13	3.02
median	3.80	4.00	3.60	3.25	3.50	3.00	4.00	4.00	4.00
sd	1.84	1.89	1.77	1.68	1.74	1.59	1.81	1.82	1.81

Figure 2. Distribution of Factor-Based Scores for Hands-on Work (Factor 1)

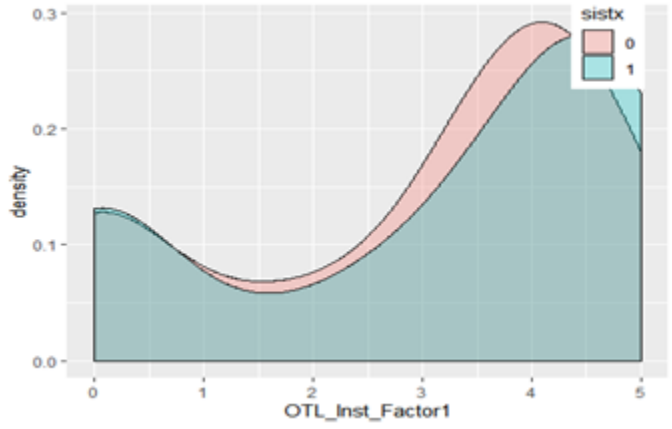

Figure 3. Distribution of Factor-Based Scores for Literacy and Discourse (Factor 2)

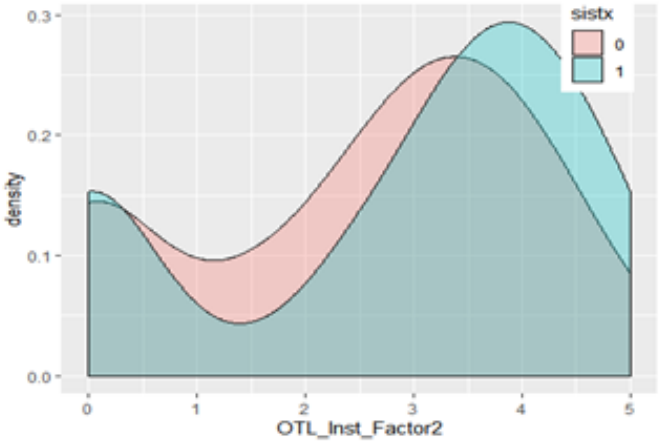

Figure 4. Distribution of Factor-Based Scores for Science Ideas (Factor 3)