Title: Year 1 of an Efficacy Trial of the Promoting Science among English Language Learners (P-SELL) Intervention in Grade 5 classrooms: Intervention, Results, and Limitations

Author(s): Randall D. Penfield & Okhee Lee
**Background / Context:**
Current classroom practices have largely been shaped by changing student demographics, including English language learners (ELLs), and evolving accountability policies. The educational intervention in this project takes place against this backdrop. The intervention is grounded in three areas of literature: (a) reform-oriented practices in science, (b) science instruction with ELL students, and (c) high-stakes testing and accountability policy in science with ELL students.

**Purpose / Objective / Research Question / Focus of Study:**
The goal of the project is to conduct a cluster randomized field trial in order to evaluate the efficacy of the *Promoting Science among English Language Learners* (P-SELL) curriculum in grade 5 classrooms. The project examines the following research questions:

1. What are the effects of the intervention on teachers’ knowledge and practices?
2. What are the effects of the intervention on high-stakes science test (FCAT) of all students?
3. To what extent does ESOL level moderate the effect of the intervention on FCAT Science?
4. Within the treatment group, what are the effects of fidelity of implementation on FCAT Science?

**Setting:**
The project involves fifth grade students who take the state science test (FCAT) that counts toward school accountability in Florida. The project takes place in Miami-Dade County Public Schools (M-DCPS), the fourth largest school district in the nation, with high percentages of ELL and ethnic minority students.

**Population / Participants / Subjects:**
Of the 238 elementary schools, the district provided a pool of 206 elementary schools, excluding 23 elementary schools that are under close monitoring due to critically low performance and 9 schools that had participated in our previous project. From the pool of 206 schools, 64 schools were randomly selected and then randomly assigned to one of two groups: (1) a treatment group of 32 schools participating in the program for two years and sustaining the program for the final year and (2) a control group of 32 schools using the district-adopted science textbook. An invitation was sent, and all 64 schools agreed to participate.

During the first-year implementation, the project has close to 130 fifth grade teachers in the 32 treatment schools and close to 100 fifth grade teachers in the control schools. The teachers in the sample come from diverse linguistic and ethnic backgrounds, which is reflective of the diversity of the student sample.

**Intervention / Program / Practice:**
The curriculum includes the student book, teachers’ guide, and science supplies. The project addresses science standards and assessment for fifth grade in Florida, as our stand-alone fifth grade science curriculum (encompassing the nature of science, Earth and space science, physical science, and life science) promotes students’ scientific inquiry and understanding and
ultimately prepares them for the state science test as part of school accountability. To promote science achievement of ELLs, the curriculum highlights activities or strategies to foster literacy development (i.e., reading and writing) and language support strategies for ELLs (i.e., typically identified as ESOL strategies).

During the first-year implementation of the project, five full-day workshops are being offered, including a three-day summer workshop (or a two-day make-up workshop for teachers who did not participate in the summary workshop), one full-day workshop prior to the state science test, and one full-day workshop at the end of the school year. The focus of the workshops during the first-year implementation is on familiarizing the teachers with the science content and hands-on activities, and also the state science content standards and assessment.

Research Design:

The study employs a cluster randomized field trial design, involving two groups (treatment group and control group), with each group consisting of 32 randomly selected and randomly assigned schools, for a total of 64 schools. The school district (M-DCPS) is organized according to five regions, and the random selection and assignment procedure is stratified on region such that the 64 schools included in the study are evenly distributed across the five regions (i.e., 6 or 7 treatment and 6 or 7 control schools were selected from each of the five regions).

The student-level outcome measure for all students across all 64 schools is fifth grade FCAT Science score obtained in the spring of the school year. Although the project had aspired to administer a project-developed science test at the beginning and end of the school year, this component of data collection was not achieved due to restrictions placed on data collection by the school district, and represented a substantial boundary encountered in the actual implementation of the study.

Three different teacher-level outcome measures are employed to measure teacher knowledge and instructional practices. First, a teacher science knowledge test is administered at the beginning and end of the school year to every participating science teacher in the treatment and control schools. Second, using a questionnaire, teacher self-reported knowledge and instructional practices are obtained at the beginning and end of the school year for every participating science teacher in the treatment and control schools. Third, classroom observations of teacher knowledge and instructional practices are obtained at three time points throughout the school year for one randomly selected science teacher in each of the treatment and control schools. The observation of only one teacher was necessitated due to a large amount of resources required to conduct such observations. These teacher outcome variables will be used in answering research questions pertaining to the impact of the intervention on teacher knowledge and practices, as well as providing evidence of implementation fidelity.

Data Collection and Analysis:

Student science achievement outcomes (FCAT Science scores) are provided to the research team by the school district at the conclusion of the school year. Teacher outcome variables (teacher science test, teacher questionnaire, and classroom observations) are obtained by the research team.

The teacher science test and questionnaire are standardized instruments (i.e., paper and pencil instruments). These instruments have been shown to generate measures of teacher
knowledge and instructional practices that have a reliability on the order of .8 (Cronbach’s alpha ≥ .8).

The classroom observations involve observational ratings using rating scales and thus require training and calibration of rater observations. To accomplish this, initial training of the three team members involved in conducting classroom observations was facilitated by a project member with extensive previous experience using the rating instrument. This training ensured a common understanding of how to assign observational ratings. To ensure that the ratings were being assigned in an equivalent manner across the three team members, pairs of observers periodically observed the same instructional session and compared ratings to ensure consistency of rating assignment. Any discrepancies in ratings were vetted to ensure a common calibration of rating assignment.

To answer the research questions a series of hierarchical linear models (HLMs) are employed. One set of HLMs employs a three-level structure (student nested in teacher nested in school) to model (a) student FCAT Science scores as a function of treatment group status included as a level-3 school fixed factor (Research Question 2), (b) student-level effect of ESOL status (i.e., the level-1 effect of ESOL status on FCAT Science scores) as a function of treatment group status included as a level-3 school fixed factor (Research Question 3), and (c) student FCAT Science scores a function of fidelity of implementation measures for students in the treatment group (Research Question 4). A second set of HLMs employs a two-level structure (teachers nested in school) to model teacher outcome variables (measures of teacher knowledge and instructional practices) as a function of treatment group status (Research Question 1).

Findings / Results:
At the time of writing this abstract, results are not yet available. During the symposium, results will be described that address each of the research questions stated above.

Conclusions:
The implementation of the intervention has been widely successful across the treatment schools, an outcome that relied heavily on the close collaboration (and relationship with) key district administrators. However, several obstacles were encountered in conducting the efficacy trial, including (a) the inability of the project to administer a project-developed science test at pretest and posttest to participating students; (b) difficulties in tracking science teachers in the control schools; (c) low motivation for teachers in completing the teacher science test, particularly for teachers in the control schools; and (d) the occasional control teacher who gained access to the P-SELL curriculum and used it during instruction despite being in a control school. Despite these limitations, sufficient data are being collected to enable the research team to answer the specified research questions. Results concerning each of the research questions will be presented, and conclusions stemming from the obtained results will be provided.