

# RESEARCH STATEMENT

## Introduction

My research interests are in applied education program evaluation and student behavior through the use of experimental and quasi-experimental designs. Education research is interdisciplinary by nature, ranging all the way from cognitive scientists working on individual brain mapping as students learn to policymakers evaluating large scale school choice programs. I joined the Program for Interdisciplinary Research (PIER) last year to take advantage of the pre-doctoral training offered outside of my economics department in cognitive and developmental psychology, human-computer interaction, and instructional technology. PIER is funded by the U.S. Department of Education through the Institute of Education Sciences (IES) in an effort to effectively integrate similar research across disciplines. I have taken the program's ethos to heart throughout the pursuit of my PhD, working on a variety of projects ranging from district-wide school choice program evaluation to curriculum assessment to student motivation and cognition.

## Contributions of Previous and Ongoing Work

My first research project resulted in a joint paper on magnet programs that developed a new estimation technique for constructing bounds on treatment effects when researchers are faced with differential attrition rates between experimental and control groups. A common method for school districts to fairly distribute a limited number of seats in preferred academic programs is the administration of a blind lottery system. All applicants are then either "lotteried-in" or "lotteried-out" of the program, with lotteried-out students making a further education choice either elsewhere in the district or outside of the district. Differential attrition arises in these models because students that lose the lottery (control group) are more likely to pursue educational options outside the school district, and important outcome variables are then often not observed. The estimator exploits known quantiles of the outcome distribution for students remaining in the district to construct bounds on relevant treatment effects. The magnet programs in the mid-sized urban district that we investigated improved student behavioral outcomes such as offenses, timeliness, and attendance. I also plan to revisit this magnet data set soon, and incorporate outcomes (including marriage rates, voting participation records, and criminal records, available at the county level) observed later in life to make inferences about the civic treatment effects of the same magnet programs.

Turning toward curriculum choices, I studied the effects of a math cognitive tutor (MCT) system widely used throughout the U.S. after adaptation for use in Chilean and Mexican public middle schools. The MCT, titled Bridge to Algebra, covered material commonly referred to as pre-algebra in the U.S. The curriculum required large changes in pedagogy, including the use of computers for individual students to progress through the software. The study was conducted over a 6-month time period. We showed that students enrolled in schools which were randomly assigned to adopt the MCT significantly improved their math test scores on state standardized exams as compared to control group peers. However, the implementation of the changes in the schools and classrooms was not perfect. Those schools which were better prepared to make changes, especially those with sufficient computers and technical support services, saw their students master more of the software part of the curriculum. The Bridge to Algebra MCT was effectively translated and integrated into Chilean and Mexican public schools, and treatment students saw a subsequent increase in their math test scores. My recent work involving the MCT has focused on the other data produced by the software. There are large differences across schools in the hint-seeking behavior of students, plus smaller differences by gender (females ask for more hints, on average, than male counterparts). My plan is to exploit

these differences to demonstrate the effectiveness of the MCT's hint generator at improving student performance.

## Future Research Directions

In addition to completing my promising ongoing work regarding student hint-seeking behavior, I have three other major thrusts that I plan to pursue over the next few years. A larger scale Bridge to Algebra study in Chile is currently underway involving more schools and a more extensive curriculum designed to last a full school year. Students are currently using the software now, and we expect the first round of data, including pretest scores and demographics on teachers and students, to be available by the end of the 2013 calendar year. I will perform similar work as before that investigates the effectiveness of the curriculum for treated students, with the inclusion of additional background characteristics, a larger sample size, and a more extensive curriculum.

Exam wrappers are surveys given to students after they have taken and received feedback on exams. They are meant to encourage students to investigate their preparation and performance on the exam. Instead of worrying about a single score or grade as the only outcome of value, exam wrappers force students to “think about their thinking,” evaluate study techniques, and determine areas of strengths and weaknesses on the assessment. The questions on an exam wrapper usually fall into a few discrete categories: preparation and performance in the class, preparation for the exam itself, study environment, and relative strengths and weaknesses on the exam. Along with another collaborator, we have developed a promising exam wrapper for the introductory economics course at Carnegie Mellon University (CMU). We deployed it in the spring of 2013, and, using a matching estimator that compared those who completed the exam wrapper to those who did not, we investigated whether the completion of an exam wrapper in this course led to a subsequent increase in future exam performance. The initial results are encouraging, though the selection bias inherent in a matching design is less than desirable. In order to increase the internal validity of our study, in later iterations we will use a regression discontinuity design and only make the exam wrapper available to students scoring below average on the first exam. We can then compare the future exam performance of a subset of students just below the first exam average to a subset just above it (who did receive the exam wrapper) to measure the exam wrapper's effectiveness at teaching and facilitating the transfer of study skills in the course.

The last item on my research agenda has to do with student problem solving approaches when faced with “non-standard” multiple choice options on exams. Obviously these types of exams are quite prevalent in school-based assessments and state standardized exams, though there has been little work on student solving strategies for questions that include an answer choice such as “all of the above,” “none of the above,” or “not enough information given.” Questions with these answers offer teachers an opportunity to test students' abilities to consider more challenging options when answering multiple choice questions. This opportunity may have important metacognitive consequences for students. In an experimental setup with treatment groups receiving questions with non-standard answer options and a control group receiving questions with standard answer options, the first stage of the research plan would verify that the non-standard exams provide an appropriate level of difficulty, discriminate similarly across student abilities, and provide reliable test questions. Conditional on an appropriate exam, I am interested in the time differences between the two groups. My hypothesis going in is that non-standard answer options demand greater cognitive load from students, and hence they should either require more time or students should underperform their control group peers on the exams. There is also the possibility that these outcome measures differ when the non-standard option is the correct choice as opposed to when it is an incorrect one. Ideally, a final piece in this project would check whether students employed different strategies to answer questions with non-standard answer options. For example, if treatment students who understand that eliminating one answer choice when “all of the above” is another option effectively eliminates two answer choices outperform their treatment

peers who do not make this connection, then simple test-taking strategies can be devised and taught to improve student performance on standardized exams without detracting from the learning goals of classes.