Stanford
GRADUATE SCHOOL OF
EDUCATION

Digital Initiatives
Office of Innovation and Technology
Spring 2014

LONGITUDINAL STUDY TRANSITIONS FROM PAPER-AND-PENCIL TO MOBILE

After two years of conducting interviews, collecting, and hand-coding paper-and-pencil survey data, GSE's Claude Goldenberg, and UC Santa Cruz Faculty Associate Research Scientist Peggy Estrada, have embraced a digital and paperless approach to field research. Their longitudinal study examines the process of reclassifying English language learners (ELs) to fluent English proficient status in California, and the academic and course-taking outcomes of either reclassifying to fluent English proficient or remaining in EL status. In the 2012-13 school year, there were approximately 1.5 million English learners in California public schools. Some of these students, however, have met some of the minimum requirements for being reclassified to fluent English proficient yet retain their EL status. This study, for which GSE alumna Estrada is Principal Investigator, poses the questions: why, and what are the consequences of students remaining English learners when they might actually be ready to transition to mainstream English instruction? Schools ostensibly provide ELs with instructional services such as English language development and sheltered instruction. What happens when students who qualify to be reclassified continue to receive EL services and curricular placement they don't need? Could it restrict their learning opportunities by denying them access to core academic content, the full curriculum, and non-EL peers?

The mixed methods research design seeks to answer these questions by following 7 cohorts of students for 4 years and examining quantitative measures including students' scores on the California English Language Development Test (CELDT) and academic content standards achievement test, classification and reclassification status, course taking, course credits, and other academic outcomes. The quantitative data reveal broad patterns. The qualitative data inform those patterns with interviews of district staff regarding EL policies and interviews of teachers, principals, EL coordinators, and other school staff regarding their awareness and understanding of, and participation in the reclassification process, school procedures, instructional practices, and rationales for student curricular placement. "It's probably not a surprise that secondary English learners tend to take the least core and advanced courses and the most intervention courses," says Estrada. Funded by the Institute of Education Sciences, the project is a partnership between Stanford, UC Santa Cruz, and SRI International.

The research team, which includes Goldenberg, Estrada, Director of SRI Education Patrick Shields (also a GSE alum), SRI Senior Researcher Haiwen Wang, UCSC Junior Research Specialist Aleshia Barajas, and Stanford Ph.D. candidates SoYoung Park, Claudia Rivas, and Claudia Rodriguez-Mojica, used an interview-only format in Year 1 and an interview plus paper-and-pencil survey format in Year 2, to collect field data. "The project is designed so that every year the instruments are revised on the basis of what we discovered previously. The idea is that each year we are going more in depth, so that each set of questions is informed by what we learned from the previous set of questions as well as changes in policy in the district. The process allows us to develop new hypotheses based on what we've learned, and through the interview and survey process, test them," says Estrada.

"When we came back from our field work in Year 2, we were daunted by the task of having to code the paper survey. We had to input everything manually, come up with a coding scheme from scratch, and then check everything. This process took us about 3 months," says Barajas. This was when the team began considering using an electronic survey to collect participant responses. "There'd be a lot of advantages, for example, it's all coded immediately." However, the team had concerns about maintaining their high response rate with an electronic survey, as well as the new technical and logistical aspects implicit in this approach. Barajas recounts, "How would we give this to the teachers? Would we have to send them the link electronically, and would we need their email addresses in advance? How would we keep track of their surveys? A lot of these practical questions started coming up."

The team consulted with GSE-IT's Pamela Levine and Shawn Kim on an approach to electronic data collection in the field--a growing practice amongst GSE faculty and staff members for their research. Levine and Kim considered the project's particular logistical and IRB constraints in distributing a survey via email, as well as concerns about response rate. Estrada told them, "We know that our presence and our connection with [participants] motivates them to turn in the survey--we had a 98% response rate in the district during Year 2." Levine and Kim ultimately recommended creating the survey in Qualtrics and using the GSE iPads to collect response data on the spot.

Continued on page 2.

Continued from page 1.

Kim then provided in-depth guidance about the survey flow and design, in order to help the team program the survey to adapt to each respondent (based on input demographic criteria and responses). She also helped them think beyond setting up the questions, to the output each item would produce, and whether or not this would provide the type and level of information needed during analysis. "At first we were still thinking about the paper survey, so we didn't fully explore all of the capacities Qualtrics had to offer," says Barajas.

> **"It was really helpful to discover what we could really do with Qualtrics."**

"That's when Shawn came in, saw our question formats, and said 'You can try this'. We learned so much and it was really helpful to discover what we could really do with Qualtrics.

We also started thinking about how to connect our questions to what we wanted in the end in the output."

To others exploring developing and distributing an electronic survey on iPads, Estrada recommends thorough testing of the instrument and devices within the research team even before a pilot phase. In relation to using mobile devices for data collection, she says, "We would try to reduce the amount of typing required by replacing text fields with touch buttons," because using the iPad's built-in keyboard to enter text was cumbersome. She also recommends still giving respondents the option to use a paper survey.

The team was receptive to using this approach in future projects. "I've got this other study going on in Rwanda which also has teacher surveys and several hundred teachers. We would love to get more efficient data collection," says Goldenberg.

To get started with Qualtrics, contact Shawn Kim at **shawnkim@stanford.edu.**

PROMOTE YOUR EVENT

Submit your event to promote it on the GSE online event calendar, visit:

ed.stanford.edu/events

Digital Initiatives is written by Pamela Levine and is published 3 times a year.

For more information, contact **pblevine@stanford.edu**.

STAY CONNECTED

Follow GSE on Twitter and Facebook for education and technology-related news and events.

@StanfordEd

/StanfordEducation

EXPERIMENT-FOCUSED INSTRUCTIONAL DESIGN

A/B testing is a randomized controlled experiment method used to gain insight into online user behavior, where members of a website's audience are randomly assigned to see a version ("A" or "B") of a web page or feature, and their actions (such as click-throughs or time spent) are recorded and compared. Specific variables that can be manipulated include the content or presentation of text, images, video messages, and interactive elements. Generally faster and simpler than other types of testing, these experiments can expedite decisions about small design tweaks and enable iteration based on data rather than intuition.

In online or blended learning environments, comparisons of student responses to various instructional features can reveal relationships about learning, attention, and collaboration, and can help continuously improve course content and delivery. Joseph Jay Williams, a Postdoc in the Graduate School of Education and Office of Online Learning, spends a lot of his time thinking, "how do I get A/B testing into this?" Whereas face-to-face classrooms typically have not offered the scale necessary to attempt this strategy and produce statistically significant results, with the microscopic level of data collected at MOOC scale, "we can all of a sudden do randomized experiments in a real world setting," says Williams.

Despite the prevalence of experimental comparisons through A/B testing in web and user experience design, Williams found that "when MOOCs first came out it really wasn't possible to perform experiments on most of the platforms." Functions for randomassignment and manipulation of variables that enable these experiments were not built into the early MOOC systems. Nevertheless, "there are always ways to adapt something," says Williams, who began embedding outside services like Qualtrics within MOOC courses to pull information (such as anonymous student identifiers) from the platform, randomly assign students to conditions, and pass that information back into the system where students experience and respond to the experimental or control variables. "This functionality isn't built into the platforms yet, but that's not a reason to not incorporate experimentation." Even as Williams' work is informing and supporting building capacity for randomized controlled experiments in platforms like edX, he finds that "being able to link to Qualtrics' robust authoring tools is still really a benefit" within a MOOC. Doctoral Candidate and Lytics Lab Student Lead, René Kizilcec has created a MOOC Researcher Wiki which includes scripts for randomized controlled experiments using Qualtrics within MOOC platforms.

Once an instructor or researcher has satisfied the technical constraints of implementing an experiment in an online course, there is the challenge of packing something theoretically important or interesting into a small design change. Typical experimental comparisons examine finite changes and very narrow hypotheses. However, these individual experiments can compound to evolve courses and their impact on students.

Williams urges instructional designers and researchers to consider "MOOClets," or modular content and resources, which are a more natural fit for experimentation and easier to change based on findings. Because of the heavy investment of time and resources required to develop and produce semester-long MOOCs, many are delivered without any changes or attention to student data and feedback between offerings. "Doing experiments with modular resources is a lot easier because instead of changing a whole course, you can just

videos" says Williams. "They are then easier to iteratively improve. This is the ideal practice, but it isn't happening in a lot of MOOCs."

Williams also emphasizes the importance of "matchmaking" between online instructors and researchers, with experiment-focused design as a strategy for bringing the pedagogical expertise of instructors to bear through empirical evaluation of instructional design decisions. "With experimentation I think there's actually a lot of work we need to do there because the vocabulary is very unfriendly. We're always experimenting, and [in a conventional class] 100% of students receive the same intervention." Randomized controlled experiments in online and blended courses enables instructors to shift this continuum, and to improve the educational process by allowing data rather than intuition to drive course delivery.

Interested in learning more about experiment-focused instructional design and A/B testing for your teaching or research? Contact Joseph Jay Williams at **josephjaywilliams@stanford.edu**.