Impact Study of a Personalized Learning Model

Elizabeth McCarthy – WestEd

Ying Liu - University of Southern California

Background/Context

Personalized learning (PL) has expanded rapidly (EdWeek, 2017), with policies and funding sources supporting PL growing significantly (Banister, Reinhart, & Ross, 2015; U.S. Department of Education, 2017). PL is described as instructional practices in which students' needs and goals are accounted for during design and implementation of instruction (see Pane, et al., 2017). Digital technologies show great promise in supporting PL (Bingham, Pane, Steiner, and Hamilton, 2016). Though several studies of PL have reported positive results (U.S. Department of Education, 2017), there is a lack of research about the effectiveness of PL models.

Purpose

The purpose of the research was to test the effectiveness of an innovative PL model in improving achievement in mathematics, reading and language usage in a diverse population of K-5 school students. Research questions included:

1. Is there any impact on students' academic achievement over the course of the building phase and after one year of full implementation of the PL model? If so, what is the magnitude?
2. What is the student growth trajectory during the years of implementation? What are the trajectories for two disadvantaged groups, socioeconomically disadvantaged and English Learners?

The research also highlights a relatively new methodology, the use of a virtual comparison group.

Setting

The study was implemented in a high-poverty mid-size school district in the Western U.S. as part of an initiative funded by the U.S. Department of Education.

Participants

The treatment group included 2,304 K-5 students who were enrolled in the district in fall 2013 and participated in the pre-test. The analytic sample was balanced by gender (51% females) and majority Hispanic (60%). Most were socioeconomically disadvantaged (64%) and a large percentage were English-language learners (ELL, 25%) at baseline.

Intervention

The intervention was implemented over a four-year period and included: a growth model, individual PL plans, computer-adaptive assessments, a learning management system, blended and extended learning, student goal-setting and reflection, strengths-based learning, cross-curriculum ELD, and educator PL. Students experienced PL in classrooms, extended learning environments and, starting in year 3, at home via free internet connectivity to technology and the district learning management system.

Research Design

Because the intervention was carried throughout the entire school district, it was impossible to conduct random assignment of conditions. Alternatively, we used a pre-post quasi-experimental design with a matched "business-as-usual" comparison group. Data included students' performance on the NWEA's MAP (NWEA, 2017) assessments. For the comparison group, we used a national database of students' performance on MAP to construct the virtual comparison group (VCG, Ma & Cronin, 2009) using the baseline data. This VCG was followed longitudinally during the same time period in which the treatment students had MAP data.

Data Collection and Analysis

The MAP assessment suite is a widely-used interim assessment system designed to measure continuous learning and growth for K-12 students. MAP was administered to the treatment group three times per year from fall 2013 to spring 2017.[1] We considered fall 2013 administration as the baseline or pre-test, and the final administration (i.e., spring 2017 for mathematics and reading, and spring 2016 for language usage) as the post-test.

Our choice of analytic techniques was based on three considerations. First, observed intra-class correlations below 0.03 suggested that the inter-school variability on MAP scores was almost ignorable compared to student differences within the same school. Therefore, we ignored the student-in-school structure in the analysis. Second, because the matching was only based on a handful of student- and school- level variables, and there may likely be other risk factors that were not matched such as student-level SES and ELL statuses, we considered the two samples as being independent. Lastly, for other individual-level background variables, such as student's ethnicity and ELL or SES status, the reference sample consisted of aggregated matches, therefore it was not feasible to construct meaningful aggregation on such covariates. Instead, we interacted them with the missing data indicator, which coincided with the treatment status in our models.

To address Research Question 1, we regressed the treatment status on post-intervention scores, adjusting for covariates including baseline MAP scores. For Research Question 2, we extended the primary model to a mixed effect model, using all waves of spring MAP data as the dependent variable. The model was estimated using the maximum likelihood estimation method.

Findings/Results

Table 1, below, shows baseline equivalence of MAP pre-intervention scores. The impact was positively significant on all three subjects (mathematics, reading, and language usage), all with an effect size greater than 0.10. This suggests that the PL model intervention was effective. Results

[1] The district did not collect MAP language usage data in the 2016-17 school year because other assessments provided information related to student achievement in language usage.

of the student growth trajectory analysis suggests that after an initial dip in scores in the early building period of the initiative, particularly in mathematics, the treatment group scores in mathematics, reading, and language usage grew continuously and significantly in the remaining years. Table 2 indicates that compared to the comparison group, the performance of the treatment group improved steadily over time.

Table 3, below, shows that compared to the comparison group, the performance of the treatment group improved steadily over time. Figures 1 through 3 show the adjusted means for treatment and comparison groups over the 4 years of the building period and full intervention (years 1 to 4 of the intervention) for mathematics, reading, and language usage. Figure 4 through 6 show the estimated treatment group difference in mathematics, reading, and language usage for the high-poverty and ELL subgroups, cumulated along years.

Conclusions

This study contributes to the growing literature in the field of PL by contributing evidence related to a successful PL model. The study also builds on the growing practice of using VCGs to study educational interventions (see Pane et al., 2017). VCG designs can be relatively low-cost and allow for rigorous studies of educational interventions when randomization is not practical or possible. Overall, this study's findings will be valuable to educators, researchers, and policymakers.

References

Banister, S., Reinhart, R., & Ross, C. (2015, March). Designing an effective massive open online course for educators: Evidence from student experience. In Society for Information Technology & Teacher Education International Conference, 2015(1), 1086–1091.

Ma, L., & Cronin, J. (2009). *Evaluating the effect of random selection on virtual comparison group creation*. Retrieved from Northwest Evaluation Association website: https://www.nwea.org/content/uploads/2014/08/EVALUATING-THE-EFFECT-OF-RANDOM-SELECTION-ON-VIRTUAL-COMPARISON-GROUP-CREATION.pdf

Northwest Evaluation Association. (2017). *Measures of Academic Progress (MAP): Interim assessments for K–12*. Retrieved from https://www.nwea.org/content/uploads/2015/09/MAP-Comprehensive-Brochure-Jan17.pdf

Pane, J. F., Steiner, E. D., Baird, M. D., Hamilton, L. S., & Pane, J. D. (2017, July 11). *Informing progress: Insights on personalized learning implementation and effects*. Retrieved from RAND Corporation website: https://www.rand.org/pubs/research_reports/RR2042.html

Personalized learning: Vision vs reality. (2017, November). *Education Week*. Retrieved from http://www.edweek.org/ew/collections/personalized-learning-Full-report-2017.pdf

U.S. Department of Education, Office of Educational Technology. (2017, January). *Reimagining the role of technology in education: 2017 National Education Technology Plan update*. Retrieved from https://tech.ed.gov/files/2017/01/NETP17.pdf

Table 1

Baseline Equivalence of MAP Pre-Intervention Scores

	Treatment M (SD)	Control M (SD)	Standardized M_{Diff}
Mathematics	180.2 (26.2)	180.3 (26.0)	0.00
Reading	179.1 (25.6)	179.2 (25.4)	0.00
Language usage	192.0 (19.3)	191.9 (18.9)	0.01

Table 2

Treatment Impact Estimates for 1 Year After Full Implementation

	Treatment		Control		Estimated Difference					
	N	Adjusted *M*	*n*	Adjusted *M*	(*SE*)		95% CI	Effect Size[a]	*p* value	R^2
Mathematics	1,899	219.39	1,893	217.42	1.96	(0.42)	[1.15, 2.78]	.12	< .001	.80
Reading	1,878	212.95	1,870	211.26	1.69	(0.40)	[0.91, 2.46]	.12	< .001	.75
Language usage	1,345	213.45	1,344	212.22	1.22	(0.37)	[0.49, 1.96]	.10	.001	.79

[a]Effect sizes were calculated using Hedge's *g*, consistent with the guidance in the *What Works Clearinghouse Procedures and Standards Handbook* (Version 4.0). The mean difference is standardized by the unadjusted student-level pooled standard deviation of posttest scores. The unadjusted student-level standard deviations were 18.38 for the treatment group and 15.06 for the control group in mathematics; 15.86 for the treatment group and 12.81 for the control group in reading; and 13.52 for the treatment group and 11.17 for the control group in language usage.

Table 3

Contrast of Treatment Group Differences Over Time

	Mathematics			Reading			Language Use		
	Est.	SE	p.	Est.	SE	p	Est.	SE	p
S15 vs. S14	-1.23	0.23	0.00	0.66	0.24	0.01	1.10	0.24	0.00
S16 vs. S15	0.59	0.23	0.01	0.26	0.24	0.29	0.50	0.24	0.04
S17 vs. S16	1.28	0.23	0.00	1.11	0.24	0.00			

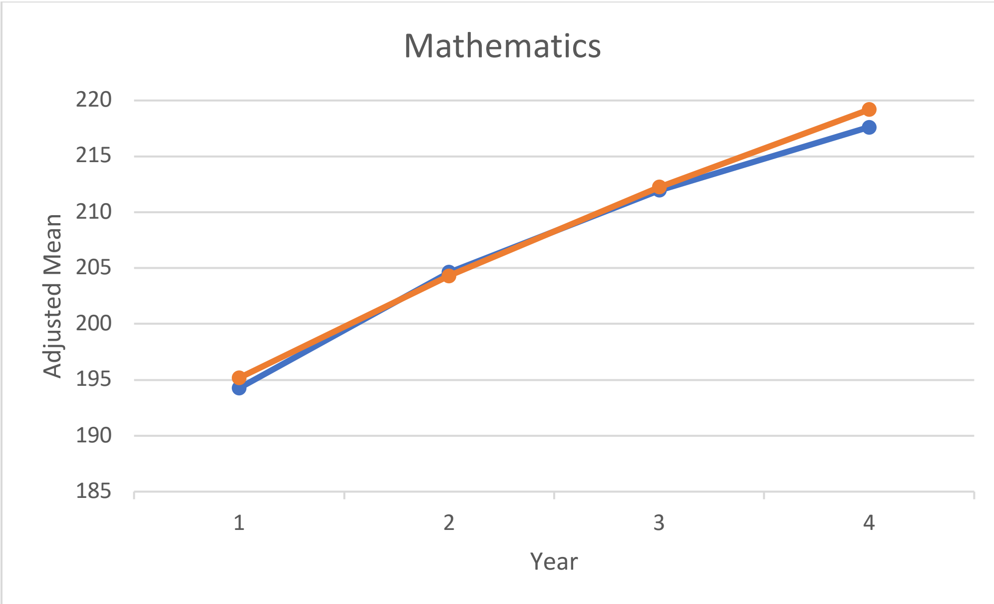

Figure 1. Adjusted means of treatment (in red) and comparison (in blue) groups over Years 1 through 4 of the intervention for mathematics.

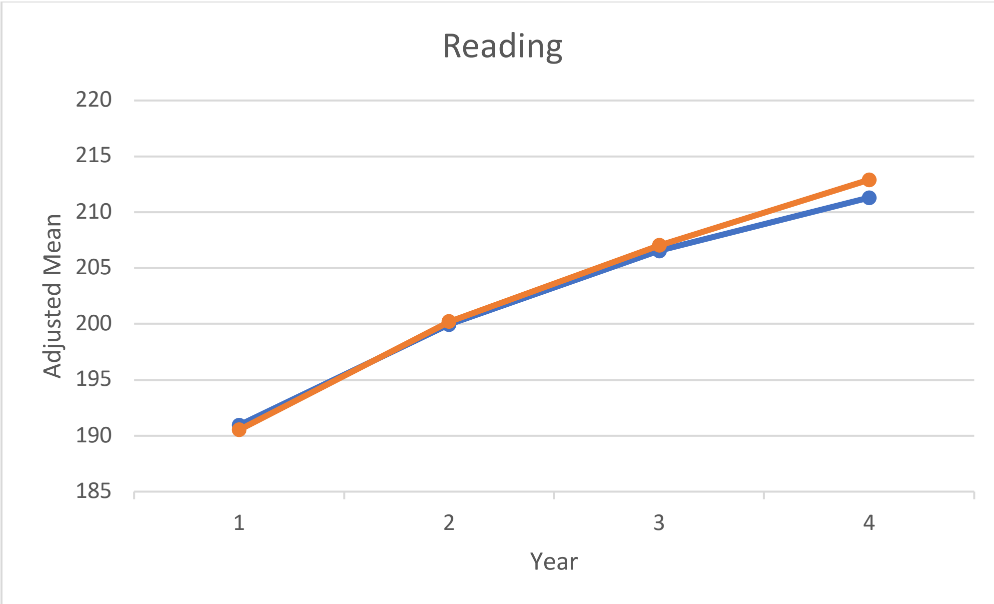

Figure 2. Adjusted means of treatment (in red) and comparison (in blue) groups over Years 1 through 4 of the intervention for reading.

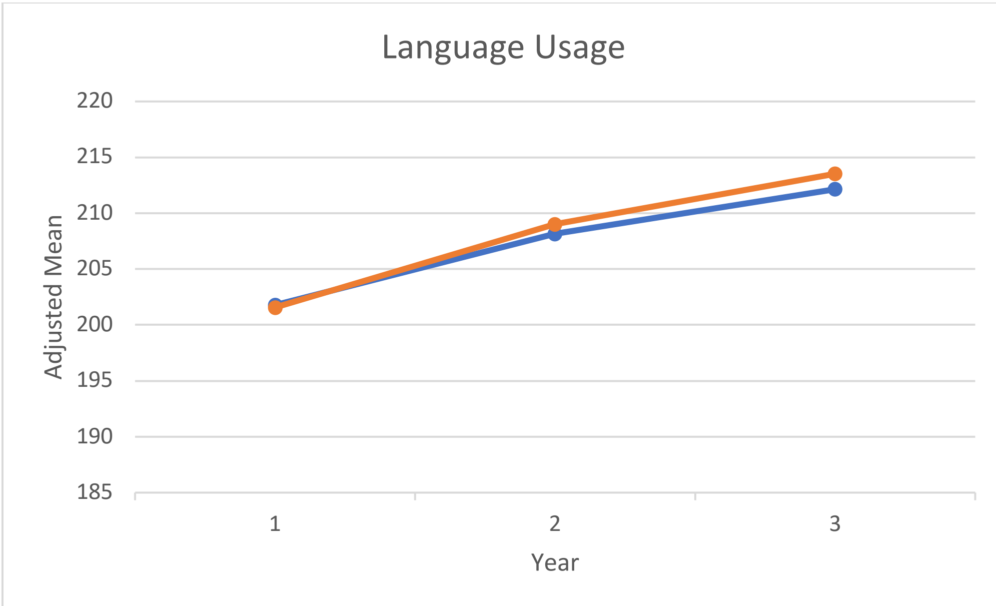

Figure 3. Adjusted means of treatment (in red) and comparison (in blue) groups over Years 1 through 4 of the intervention for language usage.

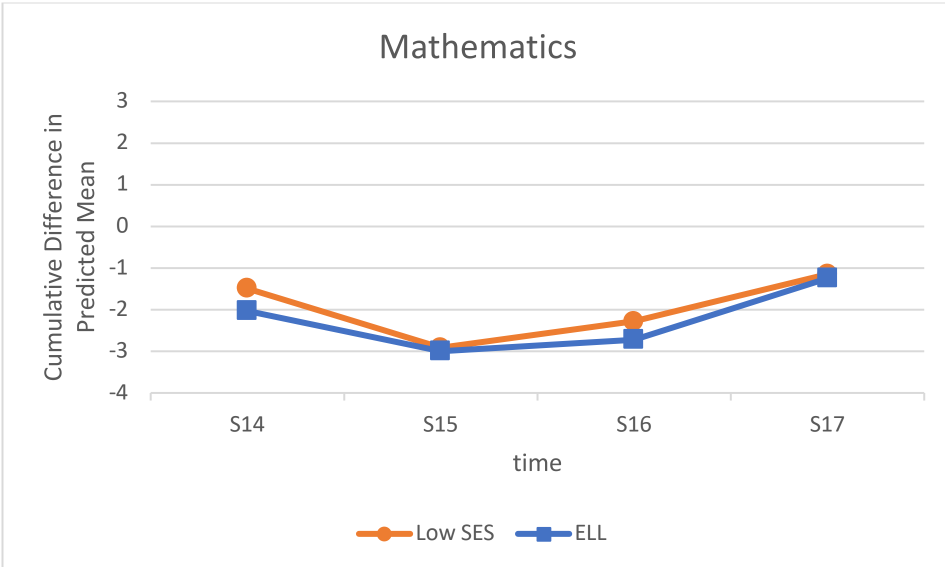

Figure 4. Trajectories of predicted means for treatment high-poverty and ELL subgroups for mathematics.

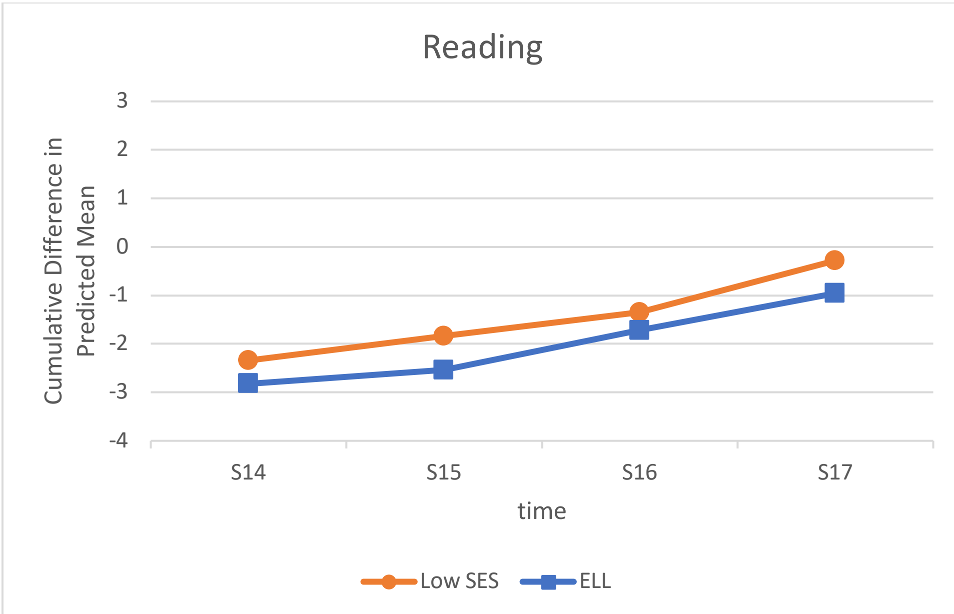

Figure 5. Trajectories of predicted means for treatment high-poverty and ELL subgroups for reading.

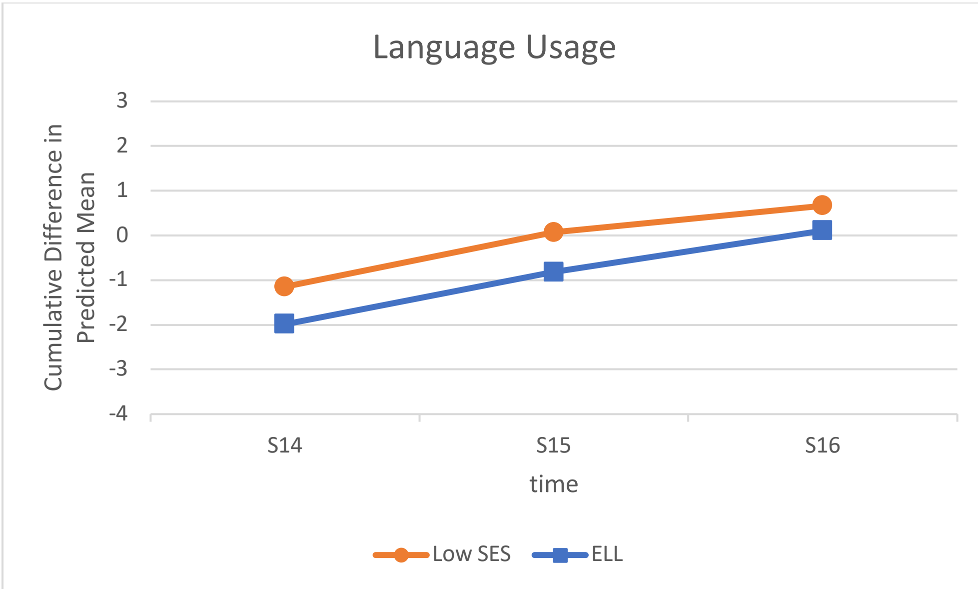

Figure 6. Trajectories of predicted means for treatment high-poverty and ELL subgroups for language usage.