Efficacy Study of Perceptual Learning Technology in Mathematics Education

Christine M. Massey[1], Andrew C. Porter[2], Laura M. Desimone[2], & Philip J. Kellman[1]
With Janie Scull[2] and Jennifer Kregor[2]
University of California, Los Angeles[1] & University of Pennsylvania[2]

Abstract

This project tested the efficacy of a set of web-based Perceptual Learning Modules (PLMs) focused on fractions and measurement with two successive cohorts of sixth-grade students. The project conducted a randomized controlled trial of the intervention consisting of four PLMs integrating (1) principles of perceptual learning that accelerate learners' abilities to recognize and discriminate key structures and relations in mathematics, and (2) adaptive learning algorithms that use a constant stream of performance data, combined with principles of learning and memory, to improve the effectiveness and efficiency of learning for each individual.

Data sources include content-aligned constructed assessments drawn from publicly available items; student-level covariate data; and teacher implementation surveys. Multilevel modeling indicates that students in the treatment arm performed significantly better ($p < .05$) than students in the control arm for three of the four modules in Cohort 1 and two in Cohort 2, with the largest effects for the module focused on Area Measurement (ES=.48). Treatment students also scored significantly higher on a composite one-year delayed posttest, indicating particularly long-lasting learning effects.

Research Questions

1) Are there differences in math learning gains for students in classrooms randomly assigned to use the Perceptual Learning Modules (PLMs), compared to students in business-as-usual control classrooms?
2) Are learning gains durable over time, as tested by a one-year delayed posttest?
3) What factors affected implementation and retention in the study?

Method

Perceptual learning is defined by *discovery effects*, or learning what features or relations are relevant to a particular concept or task, and *fluency effects*, or improving the speed and automaticity of extracting task-relevant information. The technology studied here was designed to bolster students' perceptual learning by exploiting natural human abilities to extract invariant structure from unique interactive learning episodes.

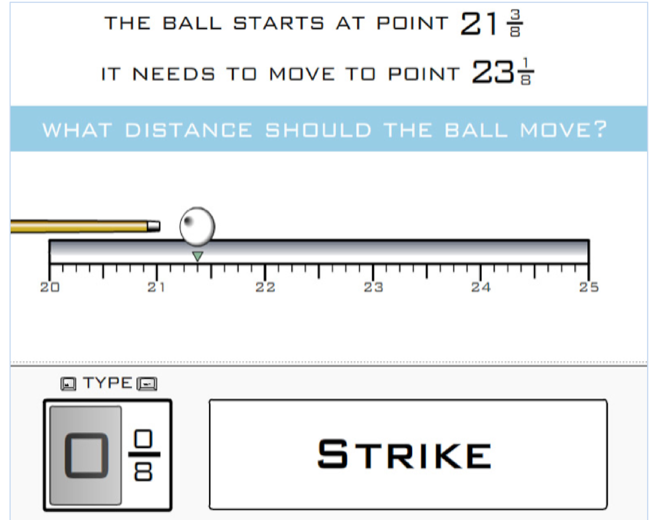
Using a within-teacher design for teachers with two 6th grade math classes, we randomly assigned one class to serve as the control and one to receive the treatment. Control classrooms received business-as-usual instruction, while students in intervention classrooms completed four PLMs over the course of the school year.

We defined two learning units and administered end-of-unit assessments for each to measure student learning. The units were composed as follows: Unit 1, which typically was taught earlier in the year, covered the modules *Fractions: Part 1*, *Fractions: Part 2*, and *Linear Measurement*. Unit 2 covered the module *Area Measurement*.

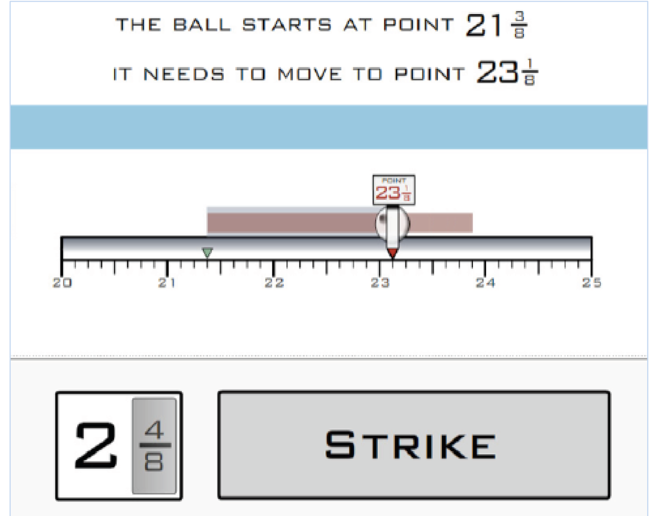
Attrition & Retention

- 7 out of 41 schools and 11 out of 52 teachers withdrew or became ineligible across the two cohorts. Every Cohort 1 teacher who remained eligible continued to participate in Cohort 2. The within-teacher design prevents differential attrition between conditions.
- Instability factors such as re-rostering of classes, changes to teachers' positions, medical leaves, and lack of technology infrastructure were the most common reasons for attrition or compromised implementation.
- Additional resources, such as laptops loaned from the project and on-site technology support, helped to retain teachers in the study, thus maintaining intended power.
- On average, Cohort 2 students were lower performing than Cohort 1 students, possibly because more high-challenge schools were retained in the Cohort 2 sample.

Example Perceptual Learning Module (PLM) Item



In this *Linear Measurement* trial, the learner is given a starting point and an ending point and is asked to enter the distance traveled. After the user keys in a response and clicks on STRIKE, an animation carries out the action.



The learner has responded incorrectly. The software illustrates a correct response in comparison to the student-provided incorrect response. Actual feedback is fully animated.

Sample Characteristics

Student Characteristics by Cohort and Treatment Condition

	Cohort 1 (Unit 1)		Cohort 2 (Unit 1)	
Student Characteristics	Control	Treat-ment	Control	Treat-ment
% Minority	0.66	0.63	0.68	0.66
% Female	0.52	0.53	0.53	0.52
% Free or Reduced-price Lunch	0.69	0.68	0.64	0.69
% Disability	0.09	0.09	0.06	0.09
% Limited English Proficient	0.04	0.08*	0.07	0.05
Math (centered scaled score)	0.23	0.12	0.11	0.00
Reading (centered scaled score)	0.21	0.14	0.11	0.01
Student *n*	515	498	562	539
Teacher *n*	24	24	29	29

Students in analysis sample are those enrolled in participating classrooms with complete end-of-unit assessments and available covariate data. *Note:* Significant contrasts across control and treatment groups within unit indicated by * $p < .05$.

The treatment and control groups are largely similar within each unit. Where they differ, the treatment group is the lower-performing or more disadvantaged. At best, then, these groups are equivalent, providing for unbiased treatment effects; and at worst, the estimates will be slightly biased against the treatment, erring on the side of more conservative treatment effect estimates. Cohort 2 students are lower-performing than Cohort 1 students, likely due to differential attrition/retention patterns.

Conclusions

- Learning software based on principles of Perceptual Learning was consistently effective in improving student learning outcomes for modules focused on Linear and Area measurement. Outcomes for Fraction modules were mixed.
- Learning gains were remarkably long-lasting, with significant effects of the PLM treatment demonstrated on a composite delayed posttest given one year after the end of the intervention.
- Results were obtained in chronically under-resourced urban schools serving high-proportions of low income and minority students, indicating that this type of computer-assisted instruction is a cost-effective resource for improving opportunity gaps in math learning for middle school students, especially as school districts continue to invest in technology.
- Ongoing analyses are examining more fine-grained patterns in the large datasets generated by this study. These include detailed item analyses and analytics conducted on the internal PLM performance and completion data collected by the software.

Results

Estimated Treatment Effects, By Cohort and Module

Module	Cohort 1		Cohort 2	
	Treatment Estimate (SE)	*Effect Size*	*Treatement Estimate (SE)*	*Effect Size*
Linear Measurement PLM	**3.96*** ** (1.054)	**.24**	**2.93** ** (0.975)	**.18**
Area Measurement PLM	**6.75*** ** (1.003)	**.48**	**6.27*** ** (1.116)	**.38**
Fractions 1 PLM	**2.05*** (0.925)	**.14**	1.89 (1.442)	
Fractions 2 PLM	1.01 (9.85)		-1.31 (1.209)	
Unit 1 Control Content	0.94 (1.194)		**-2.91** ** (1.024)	**-.19**
Unit 2 Control Content	0.45 (1.394)		-1.01 (1.035)	
One-year Delayed Posttest	**4.29*** (1.637)	**.31**	**3.01*** (1.197)	**.21**

Summary of Treatment Estimates and Effect sizes from Mixed Models for End-of-Unit PLM and Control Subscales and Delayed Posttest (Scored from 0 to 100)
Note: Significant fixed effects indicated by * $p < .05$; ** $p < .01$; *** $p < .001$.

- Significant treatment effects were found for three of the four PLM modules in Cohort 1 and two of the four modules in Cohort 2, with effect sizes ranging from .14 to .48.
- Area and Linear Measurement modules showed larger and more consistent treatment effects than Fraction modules.
- Learning gains related to the PLM treatment condition were still evident on a posttest administered after a one-year delay, with effect sizes of .31 for Cohort 1 and .21 for Cohort 2.
- A separate assessment of non-PLM related control material for Units 1 and 2 tested whether gains on PLM-related content might come at the expense of other material in curriculum. We found no statistically significant evidence of differential performance on the control-curriculum subtests for either Unit 1or 2 in Cohort 1 or for Unit 2 in Cohort 2, but a difference was detected for Unit 1 in Cohort 2, with an effect size of -.19. However, results from teachers' Surveys of Enacted Curriculum indicate that curricular alignment did not differ for treatment and control classrooms and also did not mediate positive treatment gains.

The research reported here was supported in part by the Institute of Education Sciences, U.S. Department of Education, through Grants #R305B090015 and #R305A120288 to the University of Pennsylvania and UCLA. The opinions expressed are those of the authors and do not represent views of the U.S. Dept. of Education.