Exploring the Impact of Data-Driven Optimization of Tutoring Systems on Student Learning

Karen Xiao[1,2], Conrad Borchers[2], Vincent Aleven[2]

Wellesley College[1], Carnegie Mellon University[2]

Human-Computer Interaction Institute

Motivation

Data-optimized Intelligent Tutoring Systems (ITS) aim to create more efficient and equitable learning environments by leveraging student learning data, though the real-world impact of such systems is still largely under-examined (Huang et al., 2021). This work aims to provide new insights into what constitutes a productive level of difficulty in data-optimized practice.

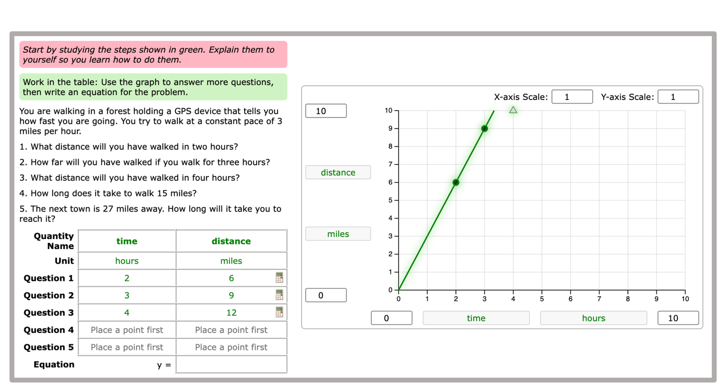
Research Questions

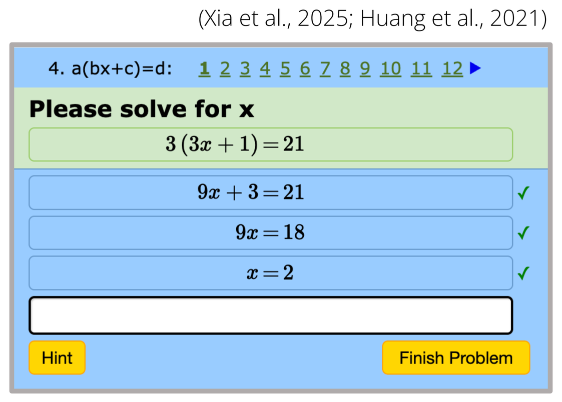
1. How do learning outcomes differ between practice with standard and data-optimized tutoring systems?
2. How do process level differences between specific optimizations impact learning outcomes between conditions?

Tutor Redesign

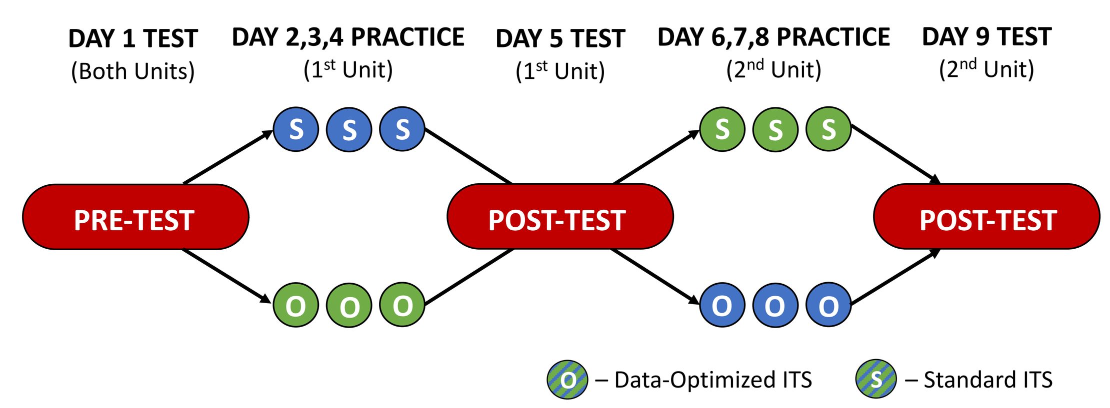
	STANDARD TUTOR	OPTIMIZED TUTOR
PROBLEM SELECTION ALGORITHM	Mastery Hard	Focused practice
SKILL MODEL	Broad skill model (e.g. division)	Split skill model (e.g. positive division & negative division)
PROBLEM POOL	Big problems practicing multiple skills	Focused practice tasks targeting a smaller set of skills
VISUAL INTERFACE (GRAPH UNIT ONLY)	Problems contain instructions, a graph, and a table for answer	Higher resolution graphs, better scaffolded instructions

(Xia et al., 2025; Huang et al., 2021)





Experiment Details

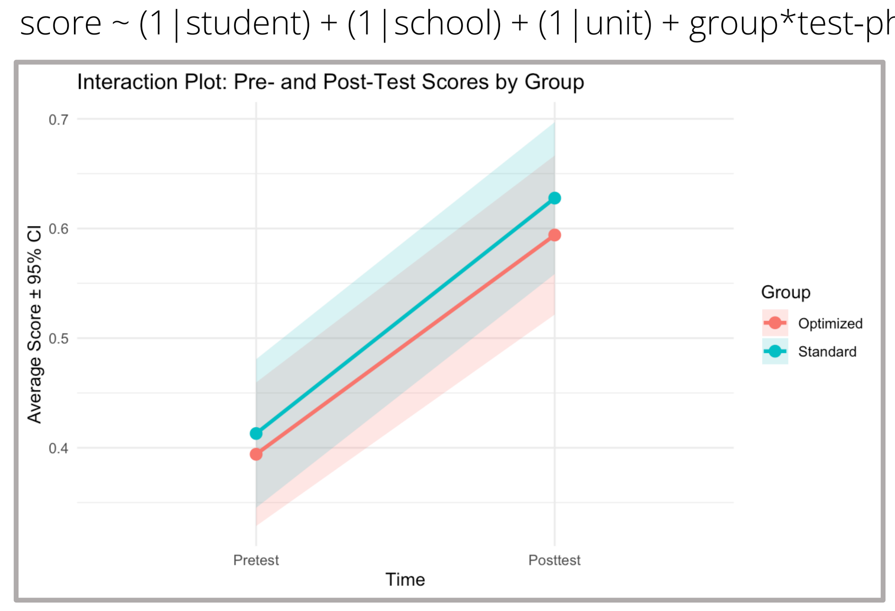


- 95 middle schoolers broken into groups
- Randomized two-condition within-subjects crossover design

Learning Outcomes

LMM FOR LEARNING GAIN ANALYSIS:

score ~ (1|student) + (1|school) + (1|unit) + group*test-phase



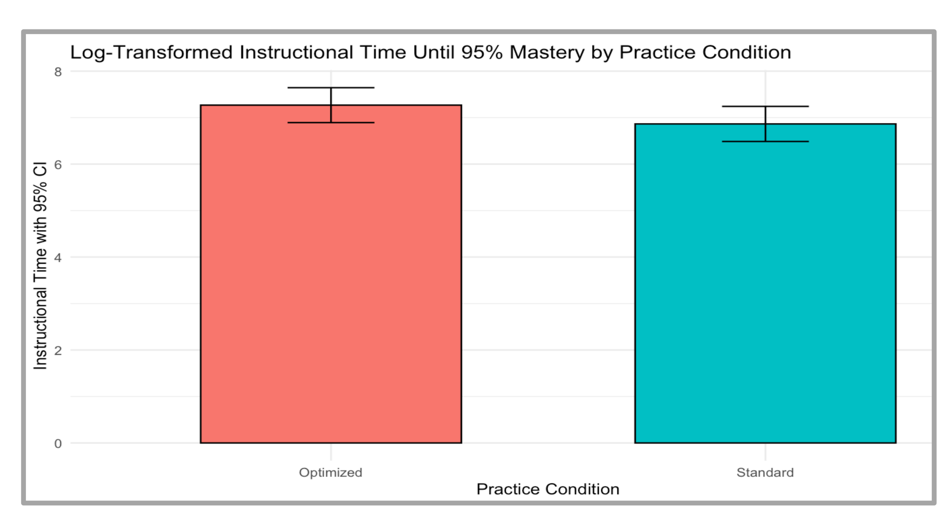
- **Significant evidence** of learning in both conditions ($p < .001$)
- **No statistically significant differences** in learning gain ($p = .19$)

iAFM FOR LEARNING RATE ANALYSIS:

Fixed Effects Model Summary for Learning Rate Analysis

Effect	Estimate	Std. Error	z-value	p
(Intercept)	0.297	0.215	1.385	.166
Opportunity	0.144	0.033	4.417	< .001
Condition	0.053	0.069	0.759	.448
Opportunity*Condition	0.011	0.008	1.420	.156

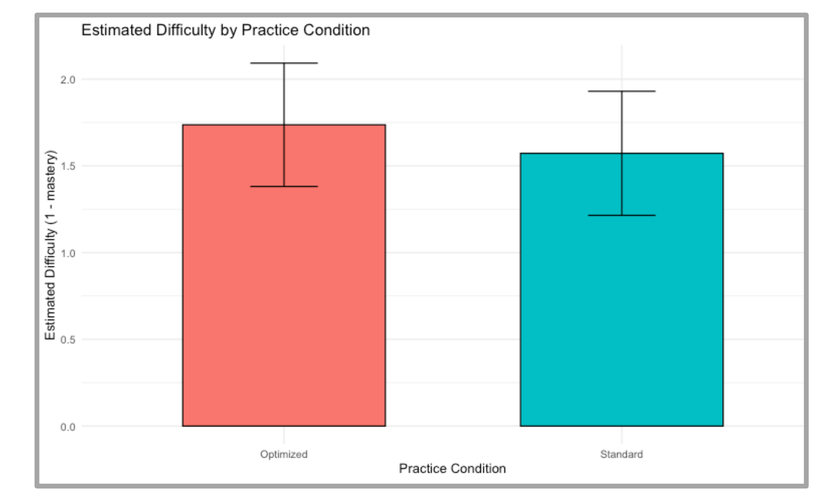
- **No statistically significant differences** in learning rates
- Students in the optimized condition required **significantly** more instructional time ($p = .01$)



Difficulty

DEFINITION

- **1 – mastery:** $d_q = \sum_{kc_i}(1 - \text{predicted\_mastery}(kc_i))$



Problems in the optimized tutor were **significantly** more difficult ($p = .01$)

COMPLETION

49% of students in the **optimized** condition completed all practice units

70% of students in the **standard** condition completed all practice units

Key Takeaways

- The data-driven redesign might have made problems *too* difficult to complete in the time allotted
    - The optimized tutor selected more difficult problems on average
    - Fewer students were able to complete optimized units

Discussion

Data-driven tutors have been proven to make practice more efficient by reducing over and under practice (Xia et al., 2025)

This work suggests the need to balance problem utility with learner accessibility

Future iterations should be careful to integrate focused practice with difficulty calibration to challenge learners without overwhelming them

Acknowledgements: *This research was primarily funded by the Institute of Education Sciences (IES) of the U.S. Department of Education under Award No. R305A220386 and partially supported by the National Science Foundation under Award No. 2349558.*

Wellesley Carnegie Mellon University

Questions?
kx100@wellesley.edu