A Novel Video Recommendation System for Algebra: An Effectiveness Evaluation Study

Presenter: Walter L. Leite, Director of the Virtual Learning Lab

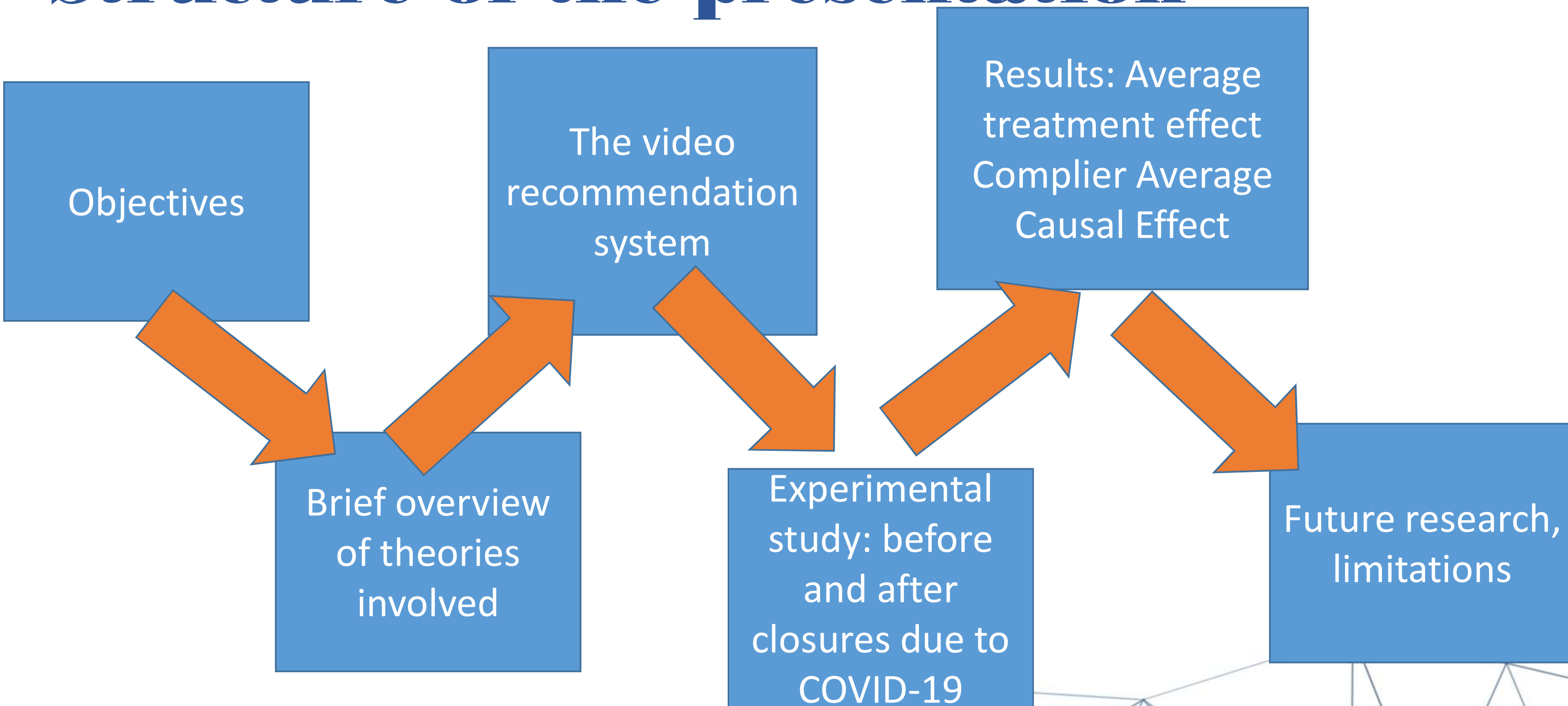
Authors: Walter L. Leite, Samrat Roy, Nilanjana Chakraborty, George Michailidis, A. Corinne Huggins-Manley, Sidney K. D'Mello, Mohamad Kazem Shirani Faradonbeh, Emily Jensen, Huan Kuang and Zeyuan Jing

UF | College of Education

Acknowledgements:

- The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305C160004 to the University of Florida. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

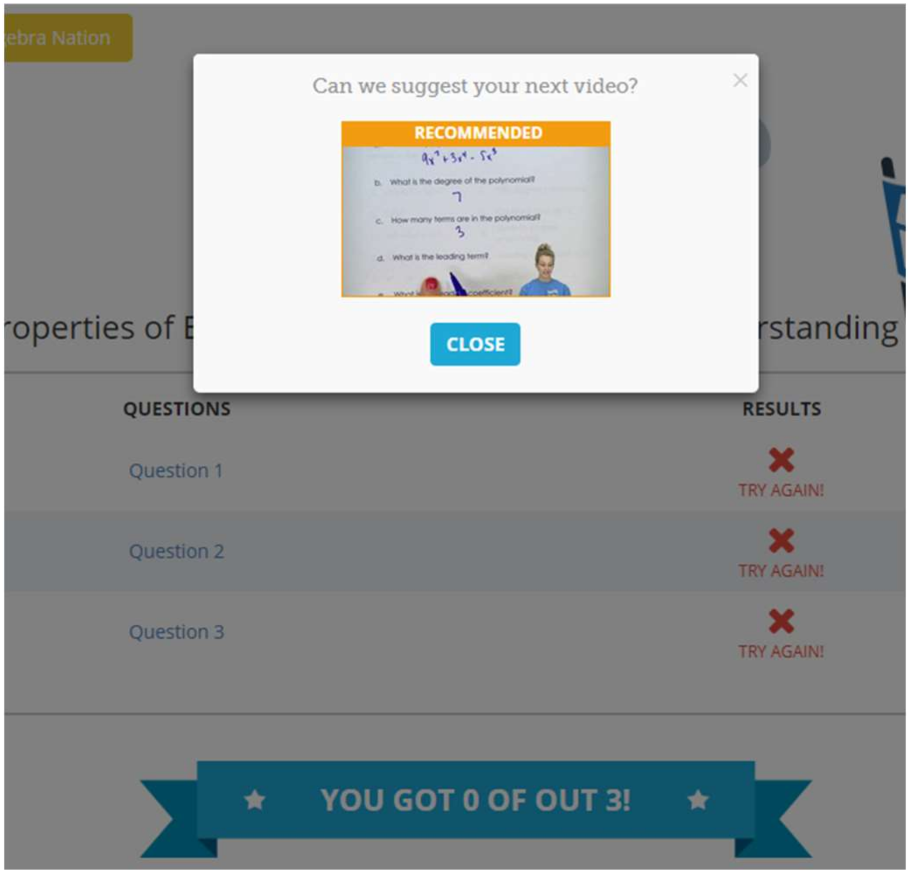
Structure of the presentation



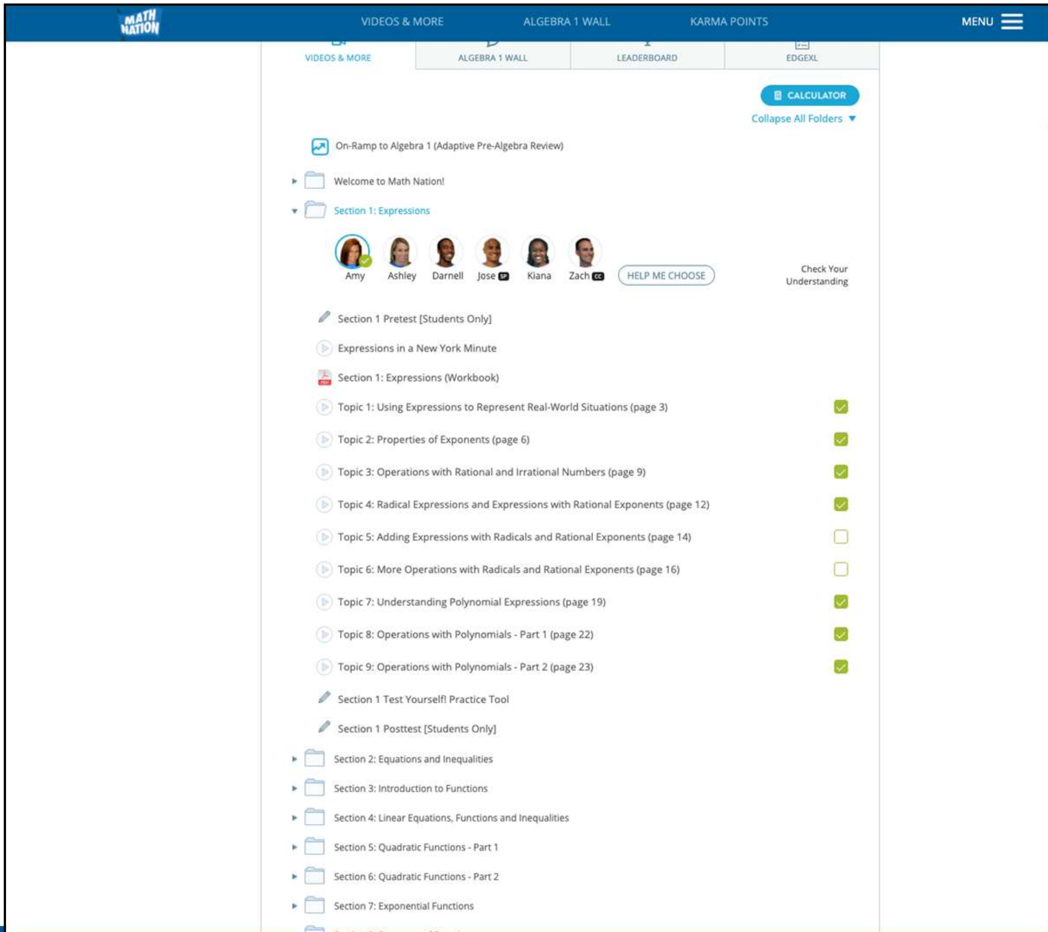
Objective of the Video Recommendation System

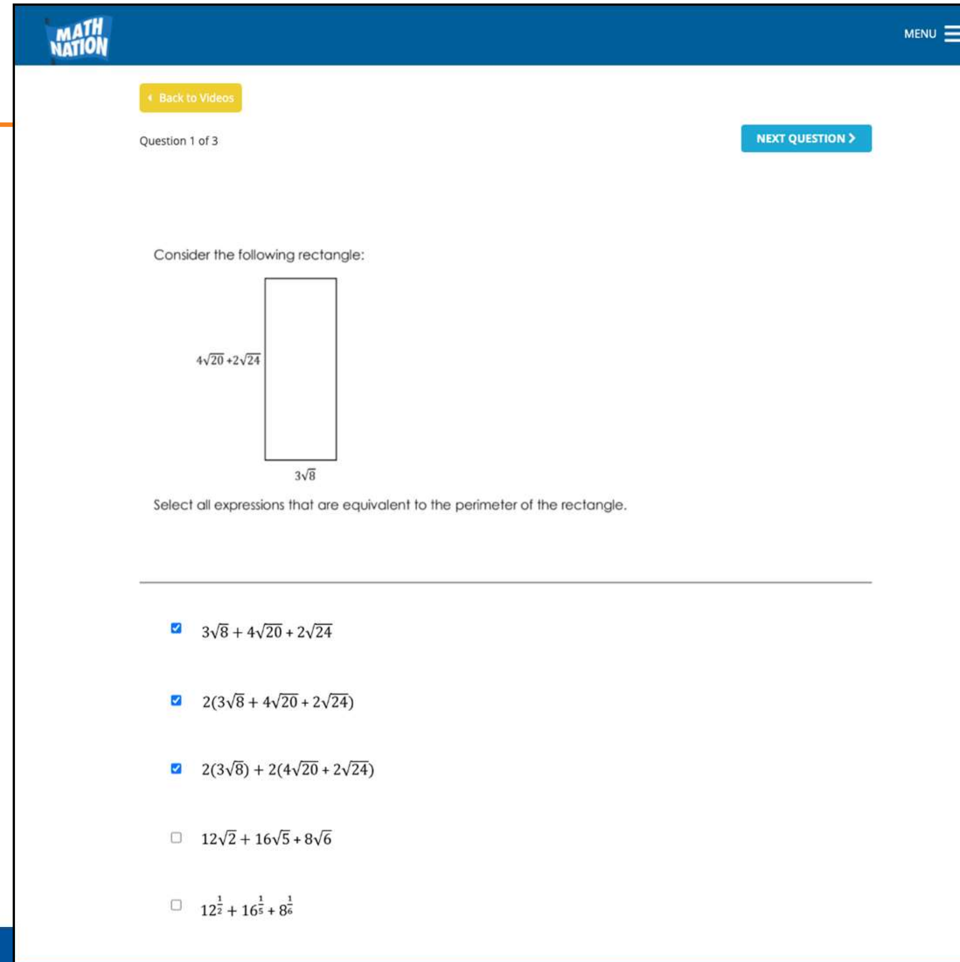
Provide students with a personalized video recommendation that takes into account their current **knowledge** as well as their **engagement** with the system.

Video Recommendation Screen in Math Nation



Context: Math Nation

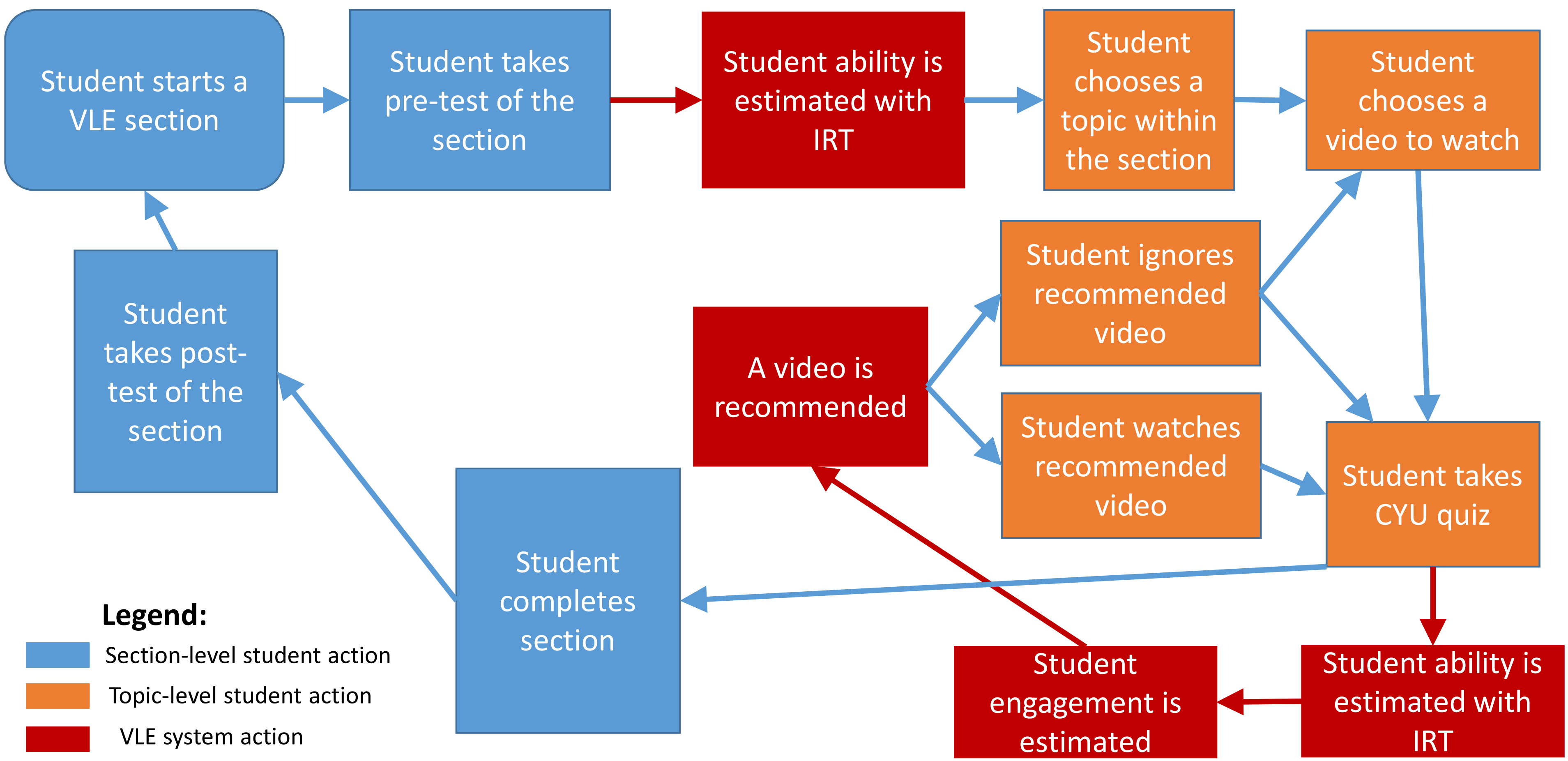




Theoretical Background

- Vygotsky's theory of Zone of Proximal Development
- D'Mello, Dieterle and Duckworth's advanced, analytic, automated (AAA) approach to measure engagement for interactions with digital learning technologies
- Item Response Theory

The video recommendation flowchart



The algorithm for new videos

ALGORITHM 1. New Video Recommendation Policy for Student *i*

Inputs: *initial ability estimates* $\{a_{ij}(0)\}$, $1 \leq i \leq n, 1 \leq j \leq r$.

Output: *sequence of recommended videos* $\hat{\jmath}\,(t) \in \{1,\ldots,r\}, t \geq 0$

for $t = 0, 1, \ldots$ **do**

Compute peer ability-estimates

$$b_j(t) = n^{-1} \sum_{i=1}^{n} a_{ij}(t).$$

Compute the probability distribution $\{p_j(t)\}, j = 1,2,\ldots,r$,

$$p_j(t) = \frac{exp\left[-w_j\,(a_{ij}(t) - b_j(t))\right]}{\sum_{j=1}^{r} exp\left[-w_j\,(a_{ij}(t) - b_j(t))\right]}.$$

Sample $\hat{\jmath}\,(t)$ *from the distribution* $\{p_j(t)\}_{1 \leq j \leq r}$.

Read $\{a_{ij}(t+1)\}$, $1 \leq i \leq n$, $1 \leq j \leq r$ *from the database.*

end for

The video recommendation system

Five categories of video recommendation:

1) View new video;
2) 2) Review current topic video with a new tutor;
3) 3) Review segment of current video with current tutor;
4) 4) Review segment of current video with a new tutor;
5) 5) View next video in curriculum sequence.

CYU score	Engagement Threshold	Probability of Recommendation of Category C
0	< 3.5	p(C=1) = 0.7 p(C=2) = 0.3
0	>= 3.5	p(C=1) = 0.3 p(C=2) = 0.7
1	< 3.5	p(C=1) = 0.3 p(C=4) = 0.7
1	>= 3.5	p(C=1) = 0.3 p(C=3) = 0.7
2	Any	p(C=3) = 1
3	Any	p(C=5) = 1

Research Questions

- Did the students, who were offered video recommendations perform better on the post-test assessments than the students who were not offered such recommendations?
- What is the causal effect of video recommendations on the achievement of those students who watched the recommended videos when offered?

Field Experiment

- Three large school districts in Florida:
  - ◇18,925 students from 152 teachers in 149 schools
- The study lasted for 17 weeks during the Spring 2020 semester (i.e., February 3rd to May 31st)
- Transition on March 17th when all schools were closed due to the COVID-19 pandemic and instruction resumed online.

Average Treatment Effects (Intention to Treat)

	Before school closure			After school closure		
	Coefficient	SE	p-value	Coefficient	SE	p-value
(Intercept)	-0.752	0.176	0.000	-0.339	0.602	0.573
ITT	**0.054**	0.027	0.043	-0.009	0.030	0.775
Pretest	-0.012	0.040	0.764	0.025	0.112	0.825
Engagement	-0.021	0.022	0.349	-0.007	0.028	0.805

Complier Average Causal Effect

- Before schools closed, the proportion of compliers among the students who were assigned to the treatment group was = 0.15921, SE = 0.0188, CI = [0.122, 0.196].
- Before-closure period, the final CACE standardized estimate is 0.34 (SE = 0.17, $p = 0.043$)
- After schools closed, the proportion of compliers was = 0.1123, SE = 0.0112, CI = [0.090, 0.134].
- The CACE was not statistically significant for the period after schools closed (CACE = -0.076, SE = 0.266, $p = 0.775$).

Where do we go from here?

- Completed a longer replication of the experiment (November 24th 2020 to June 1st 2021) to be presented at L@S 2022

- Key questions:
  - Are the effects larger with a longer exposure?
  - Do students who use the system more extensively benefit more?
  - Are there certain groups of students who benefit more?
  - Are there certain teaching strategies that moderate the impact of the video recommendation system?
  - What are the fairness and equity implications of personalization?