Ethics, Equity, and Algorithmic Bias

Module 6

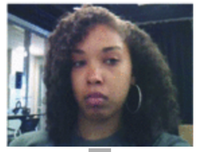
SHAMYA KARUMBAIAH, PhD
Assistant Professor
Department of Educational Psychology
University of Wisconsin-Madison

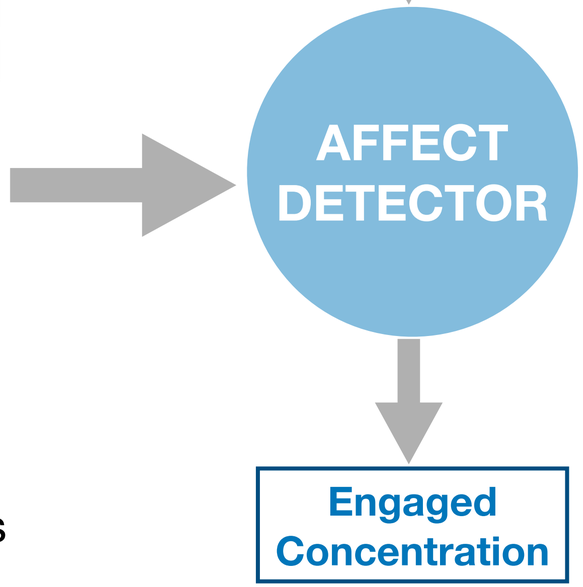
Learner Context

Li & Deng (2015)

Annotated data of student emotions

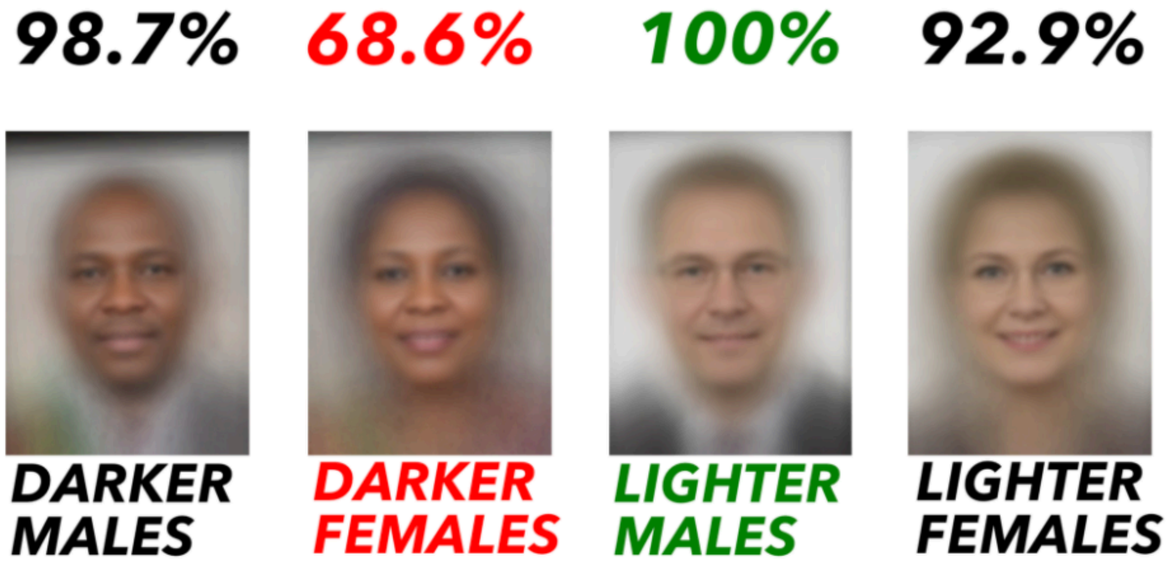
Nye et al. (2018)





In what ways do ignoring learner context introduce harmful biases in adaptive learning systems?

Learner Context



Performance Disparity in Gender Classification by Amazon Recognition
Joy Buolamwini (2019)

Madani et al. (2017); Lohr (2018); Sap et al. (2019)

Learner Context

Please write down 1-2 examples of predictive models from your context and the student populations they are expected to serve.

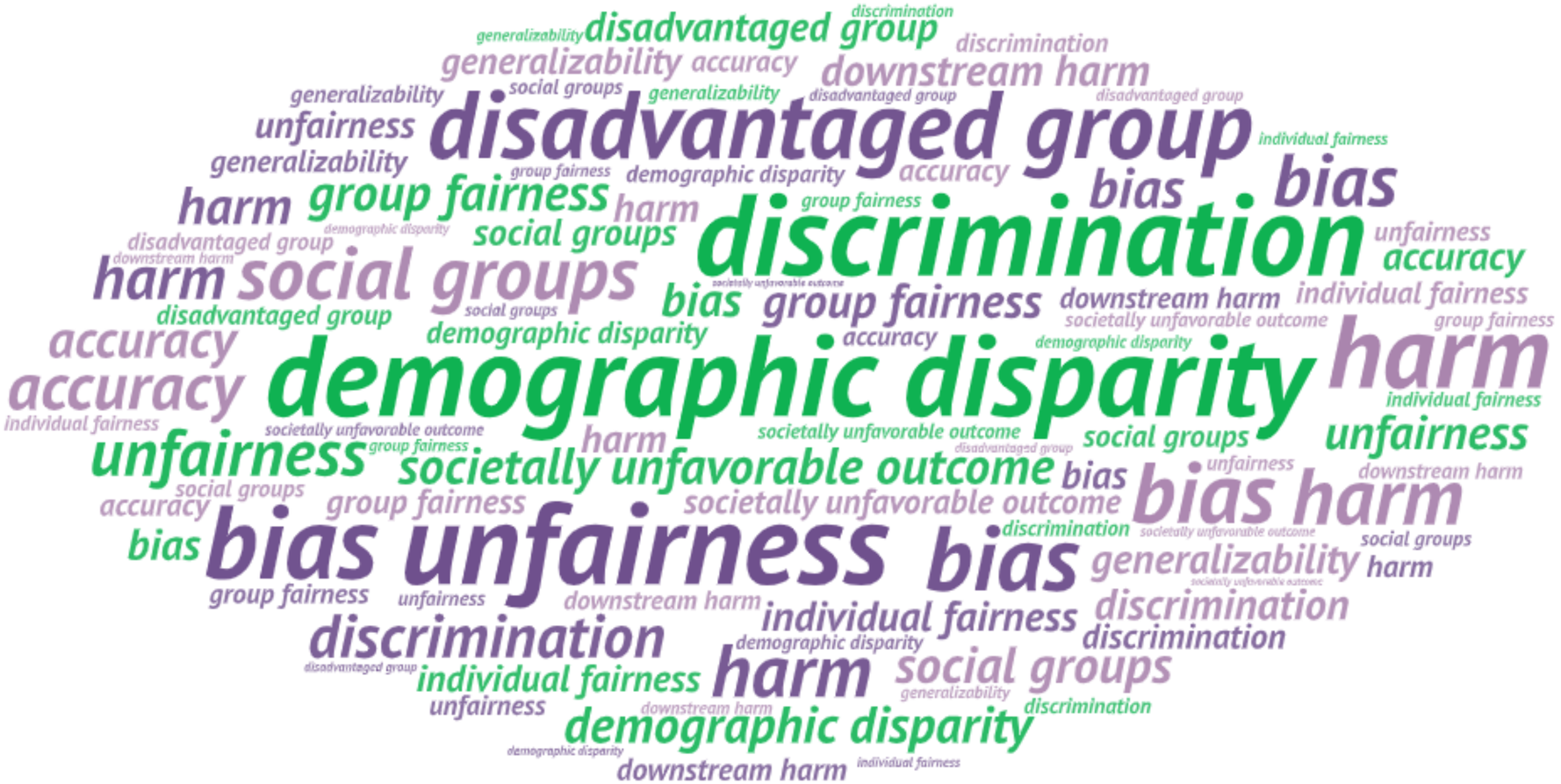
Discussion Board Question

What is Bias? Why Study Bias?

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

What is Bias in an Adaptive System?



Baker, R. S., & Hawn, A. (2021). Algorithmic Bias in Education.

Biases as the possible *sources of downstream harm* that can lead to *societally unfavorable outcomes* in specific student subpopulations.

Suresh and Guttag (2020)

Why Study Bias in Learning Analytics?

- Downstream harm of inequitable student outcomes
- Allocative harm

  e.g., bias in standardized test leading to denied college admissions (Dorans, 2010)

- Representational harm

  e.g., African American English tagged as hate speech in discussion forum posts (Sap et al, 2019)

Francesco Bonchi

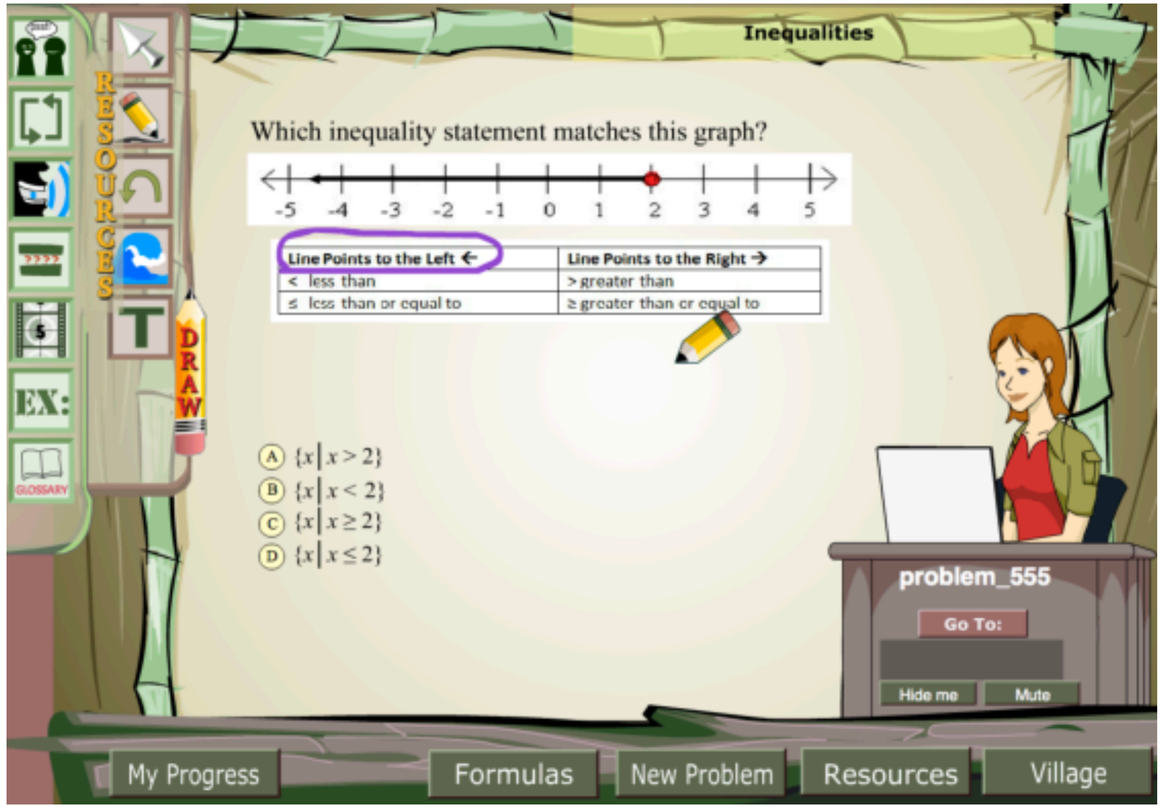
Crawford, 2017

Adaptive Learning Systems

Self, 1999; Shute & Psotka, 1994; Corbett et al., 1997; Koedinger et al., 1997; VanLehn, 2011; Luckin et al., 2016

Adaptive Learning Systems

Self, 1999; Shute & Psotka, 1994; Corbett et al., 1997; Koedinger et al., 1997; VanLehn, 2011; Luckin et al., 2016

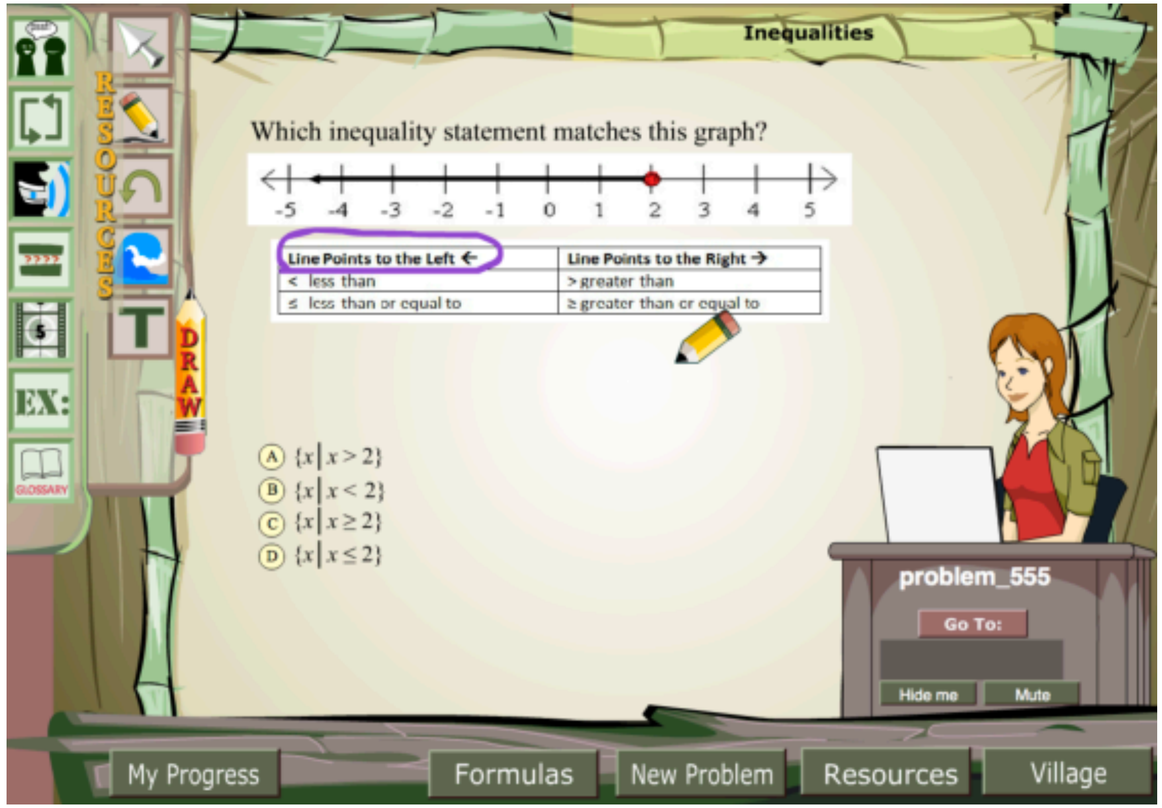


MathSpring

Karumbaiah, S. et al. (2017) Addressing Student Behavior and Affect with Empathy and Growth Mindset. [EDM17]

Adaptive Learning Systems

Self, 1999; Shute & Psotka, 1994; Corbett et al., 1997; Koedinger et al., 1997; VanLehn, 2011; Luckin et al., 2016



MathSpring

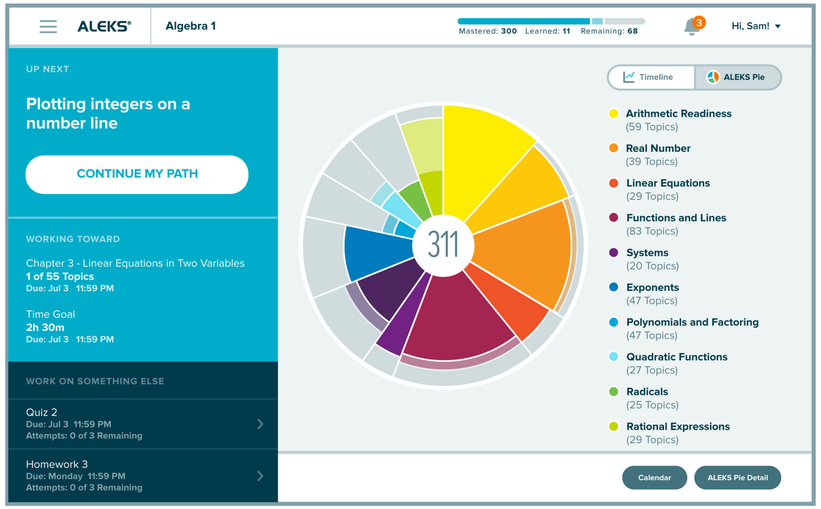
AFFECT DETECTOR

MathSpring

Karumbaiah, S. et al. (2017) Addressing Student Behavior and Affect with Empathy and Growth Mindset. [EDM17]

Adaptive Learning Systems

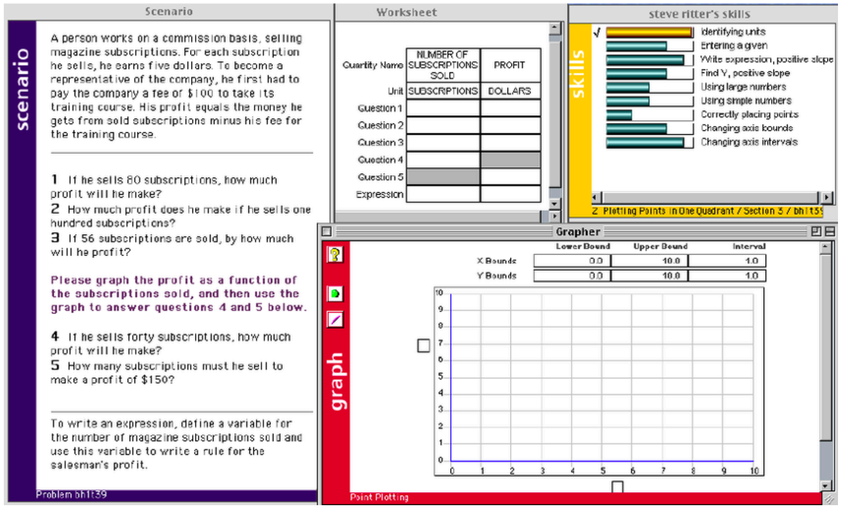
aleks.com



ALEKS

~600,000 students

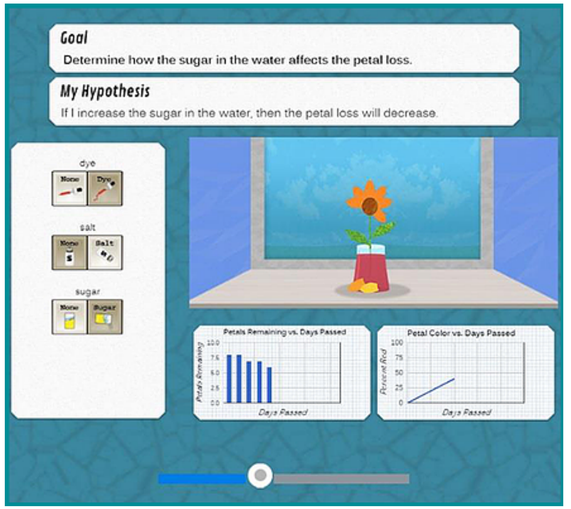
Ritter et al. (2007)



Cognitive Tutor aka Mathia

~500,000 students

inqits.com

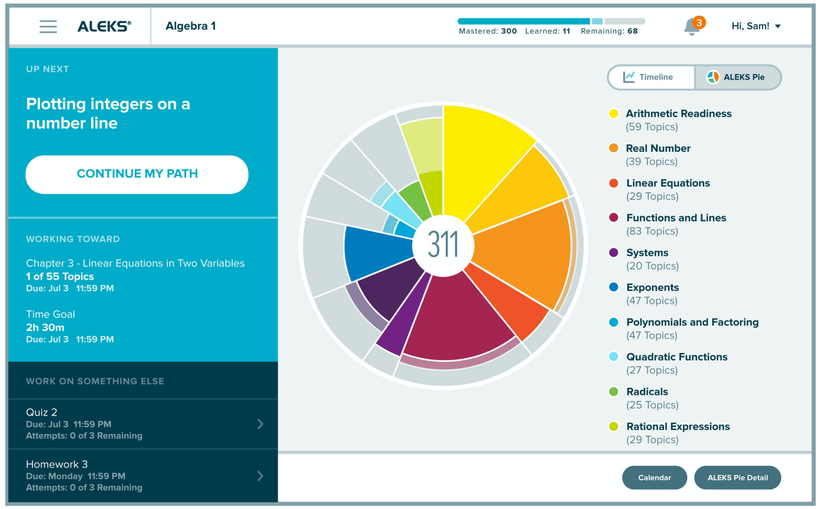


Inq-ITS

~100,000 students

Adaptive Learning Systems

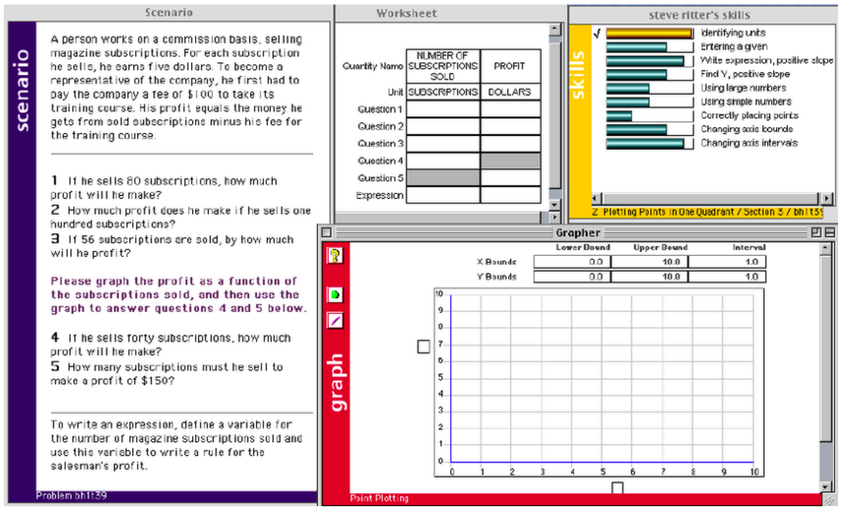
aleks.com



ALEKS

~600,000 students

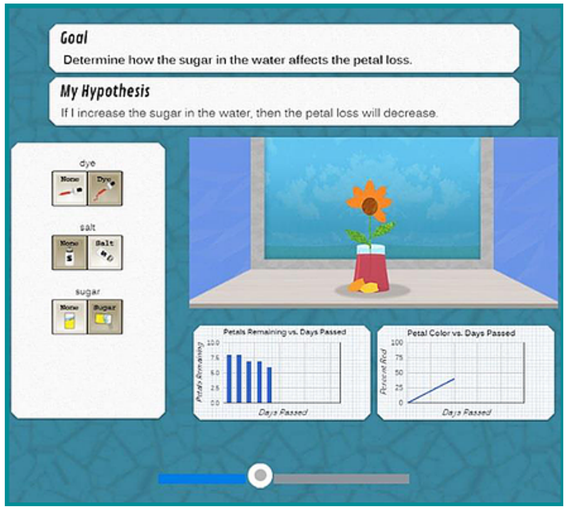
Ritter et al. (2007)



Cognitive Tutor aka Mathia

~500,000 students

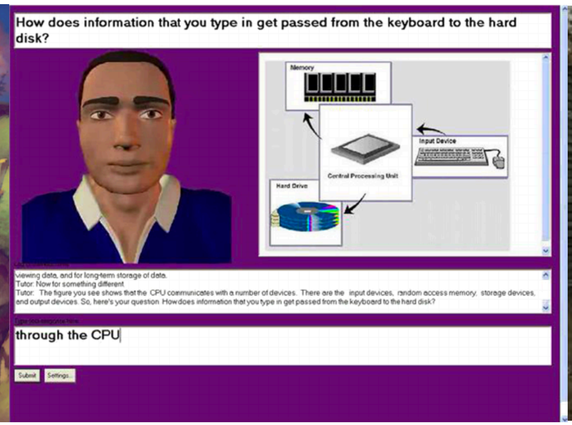
inqits.com



Inq-ITS

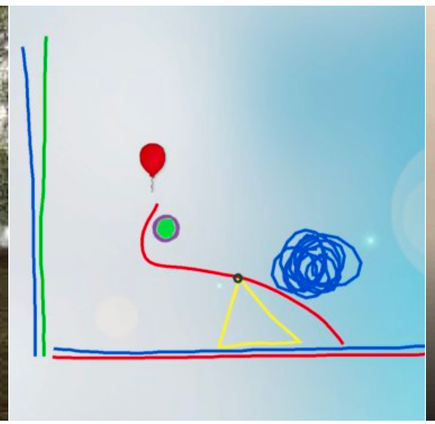
~100,000 students

Zoombinis



AutoTutor

Crystal Island



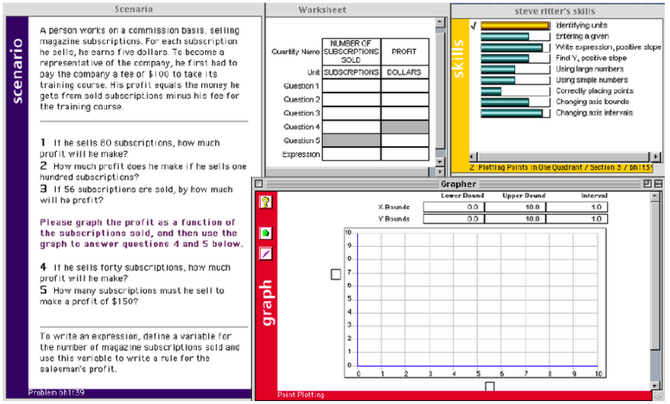
Physics Playground

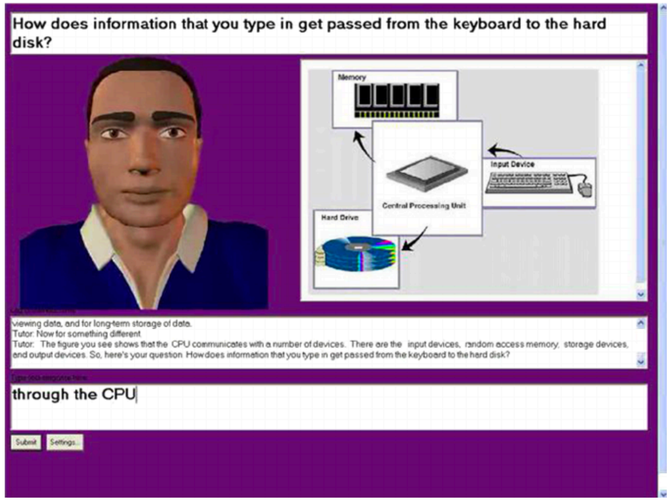
TransfrVR

Studying Bias: important, yet challenging

1. Makes real-time decisions that impact students' learning and experiences closely
2. Involves models of complex educational constructs that utilize fine-grained interaction data
3. Despite wide usage, biases not yet studied thoroughly

Cognitive Tutor Algebra

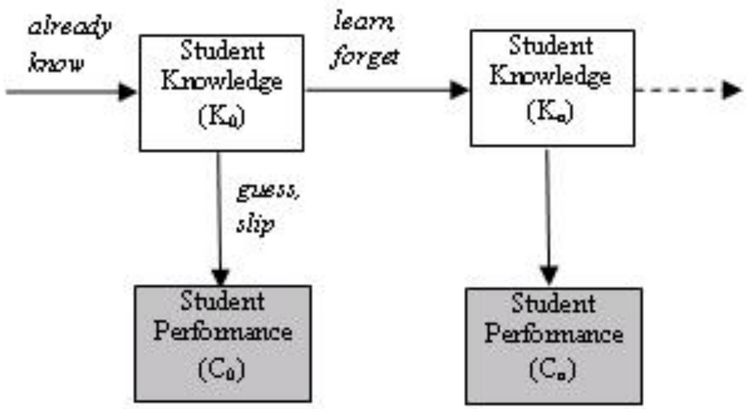




AutoTutor

An Example

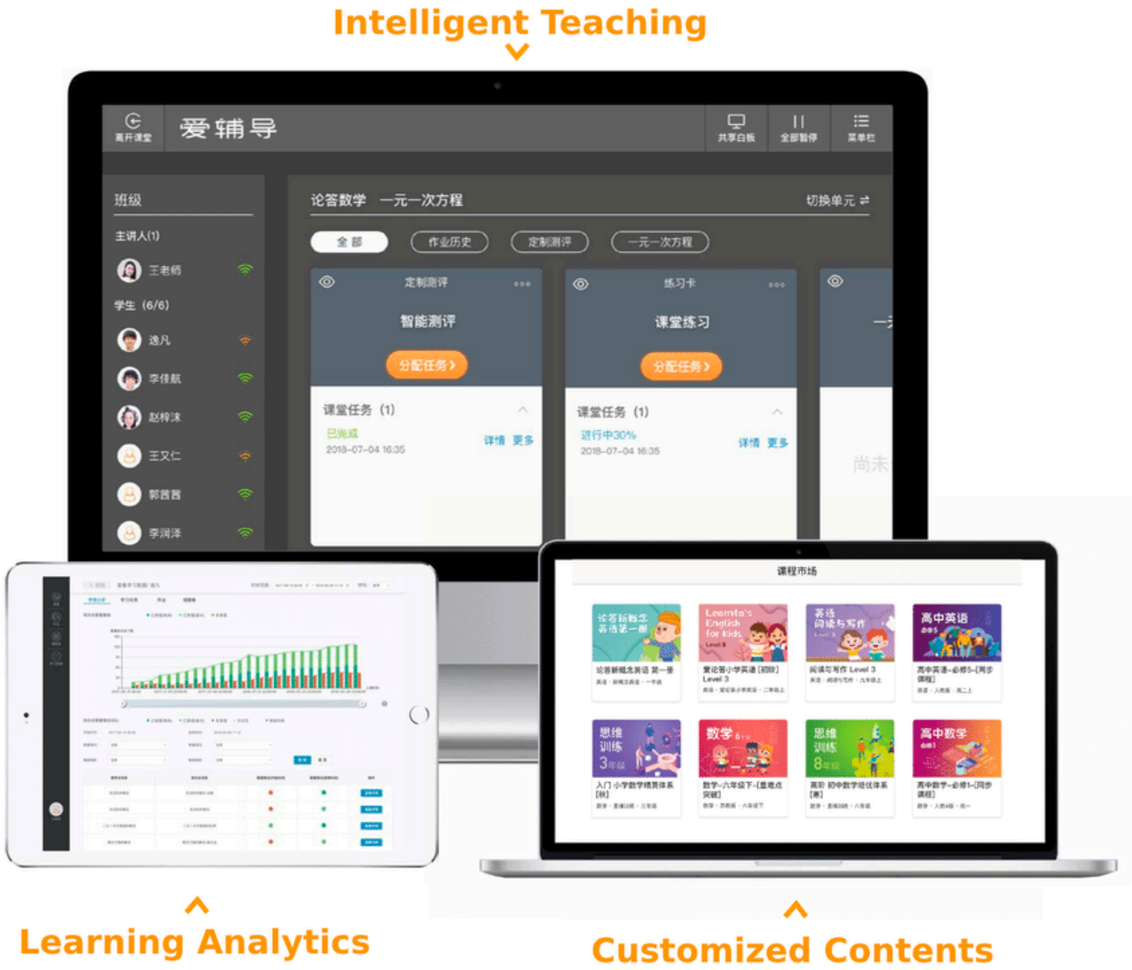
Knowledge Tracing



cs.cmu.edu

Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge.

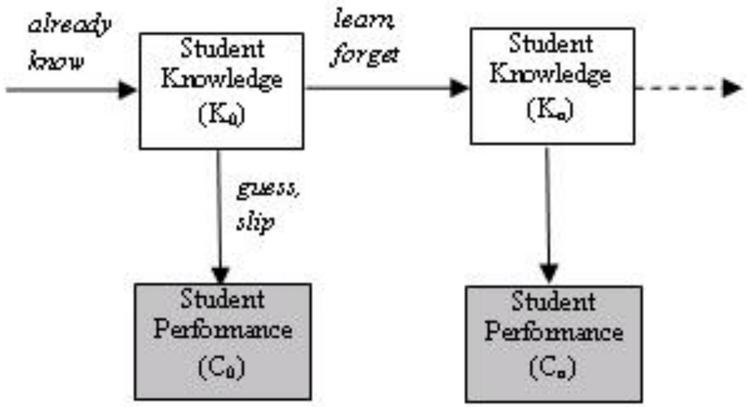
Learnta



Baker et al. (2020)

Allocative Harms of Bias

Knowledge Tracing



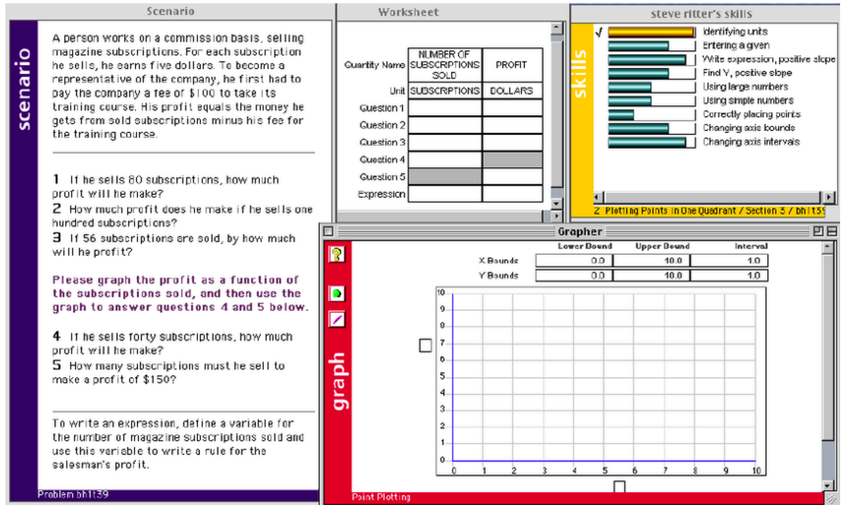
cs.cmu.edu

Francesco Bonchi

Crawford, K. (2017). The Trouble with Bias. NIPS Keynote.

Diversity in Learner Context

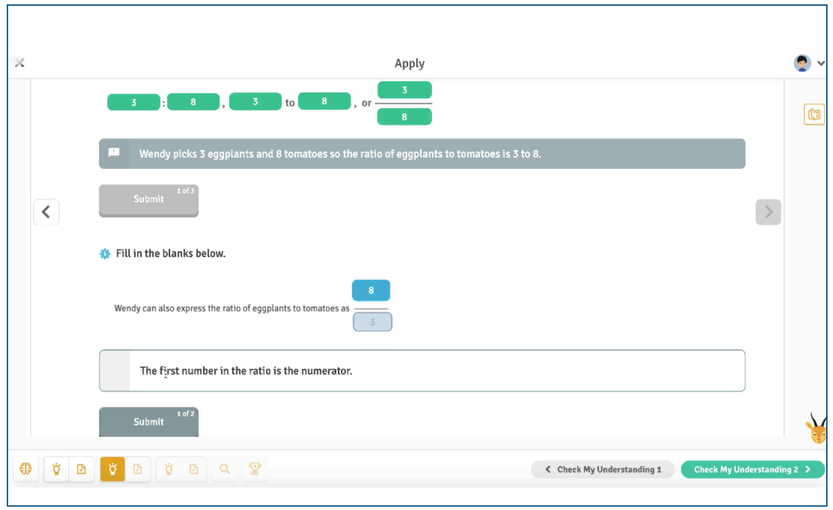
Ritter et al. (2007)



Cognitive Tutor aka Mathia

Used in the United States and Chile

nexgenedu.com



Alef NexGen

Serves students from elementary through high school

Kizilcec, R. F., & Lee, H. (2020). Algorithmic Fairness in Education.

Why Study Bias in Learning Analytics?

- Downstream harm of inequitable student outcomes
- Allocative harm

  e.g., bias in standardized test leading to denied college admissions (Dorans, 2010)

- Representational harm

  e.g., African American English tagged as hate speech in discussion forum posts (Sap et al, 2019)

For your example model, please share some potential harms to student populations if it were to be biased against them.

Discussion Board Question

Crawford, 2017