Upstream Sources of Bias in Adaptive Learning System

SHAMYA KARUMBAIAH
Postdoctoral Fellow
Carnegie Mellon University

The Search for Context

Bias in Adaptive Learning Systems

Upstream Sources of Bias

Contextualizing Theoretical Model of Affect

Contextualizing Origins of Bias

Continued Search For Context

The Search for Context

PENN

Learning Sciences, Learning Analytics
2017-2021

Machine Learning, Learning Analytics
2015-2017

Computer Science, Software Engineering
2007-2015

Affect (AIED18), Persistence (EDM18), Cognition (LDK19), Self-Identity (LAK19), Help-Seeking (EDM19)

Li & Deng (2015)

Annotated data of student emotions

Nye et al. (2018)

AFFECT DETECTOR

Engaged Concentration

Nye, B. D., **Karumbaiah, S.,** et al (2018) Engaging with the Scenario: Affect and Facial Patterns from a Scenario-Based Intelligent Tutoring System [AIED18]

The Search for Context

Affect (AIED18), Persistence (EDM18), Cognition (LDK19), Self-Identity (LAK19), Help-Seeking (EDM19)

Learning Sciences, Learning Analytics
2017-2021

"Although the **learning sciences** is continually evolving, what remains true of the tenets of this educational field is that learning happens through mediated processes that most often require collaboration with others whereby learning is inextricably linked to **context** and culture" - Dr. Yoon, EDUC 545, Penn GSE

Machine Learning, Learning Analytics
2015-2017

"**Learning analytics** is the measurement, collection, analysis and reporting of data about learners and their **contexts**, for purposes of understanding and optimizing learning and the environments in which it occurs." - SOLAR

Computer Science, Software Engineering
2007-2015

The Search for Context

Learning Sciences, Learning Analytics
2017-2021

Machine Learning, Learning Analytics
2015-2017

Computer Science, Software Engineering
2007-2015

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

Karumbaiah, S., Ocumpaugh, J., & Baker, R. S. (2021). Context Matters: Differing Implications of Motivation and Help-Seeking in Educational Technology. *International Journal of Artificial Intelligence in Education*.

Karumbaiah, S., Lan, A., Nagpal, S., Baker, R. S., Botelho, A., & Heffernan, N. (2021). Using Past Data to Warm Start Active Machine Learning: Does Context Matter? *Learning Analytics and Knowledge*.

Karumbaiah, S., Baker, R. B., Ocumpaugh, J., & Andres, A. (2021). A Re-Analysis and Synthesis of Data on Affect Dynamics in Learning. *IEEE Transactions on Affective Computing*.

The Search for Context

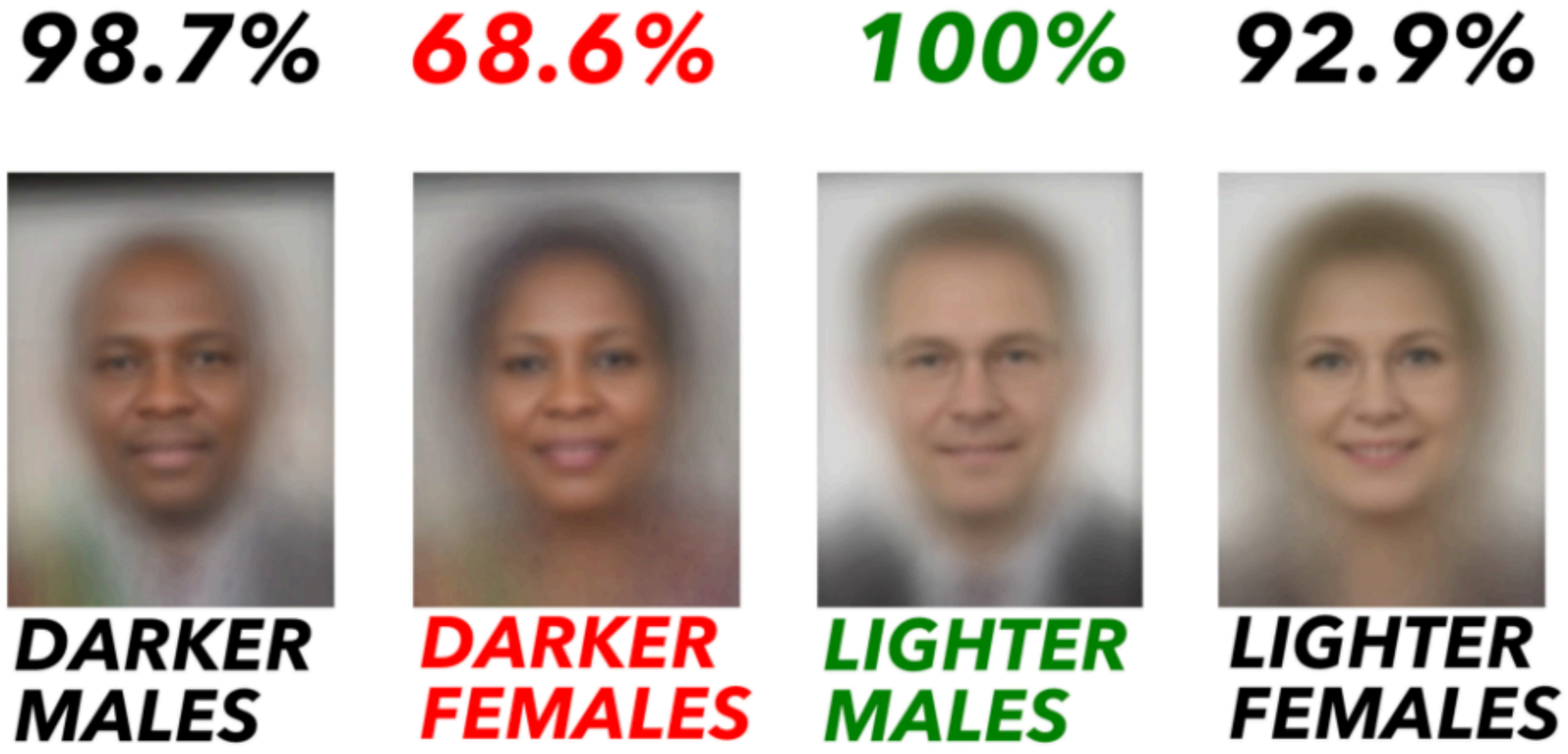
Learning Sciences, Learning Analytics
2017-2021

Machine Learning, Learning Analytics
2015-2017

Computer Science, Software Engineering
2007-2015

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

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



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

The Search for Context

Bias in Adaptive Learning Systems

Upstream Sources of Bias

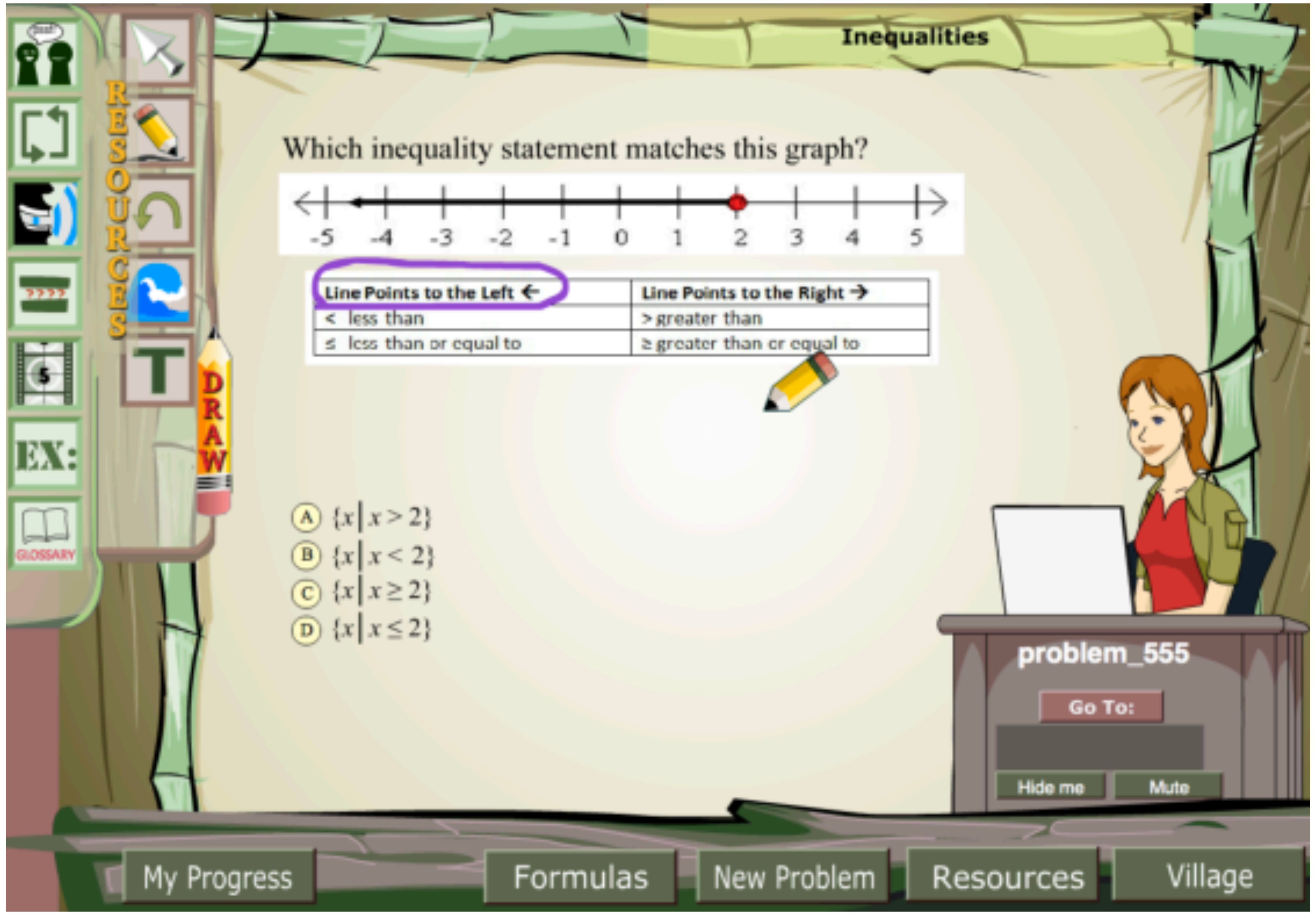
Contextualizing Theoretical Model of Affect

Contextualizing Origins of Bias

Continued Search For Context

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]

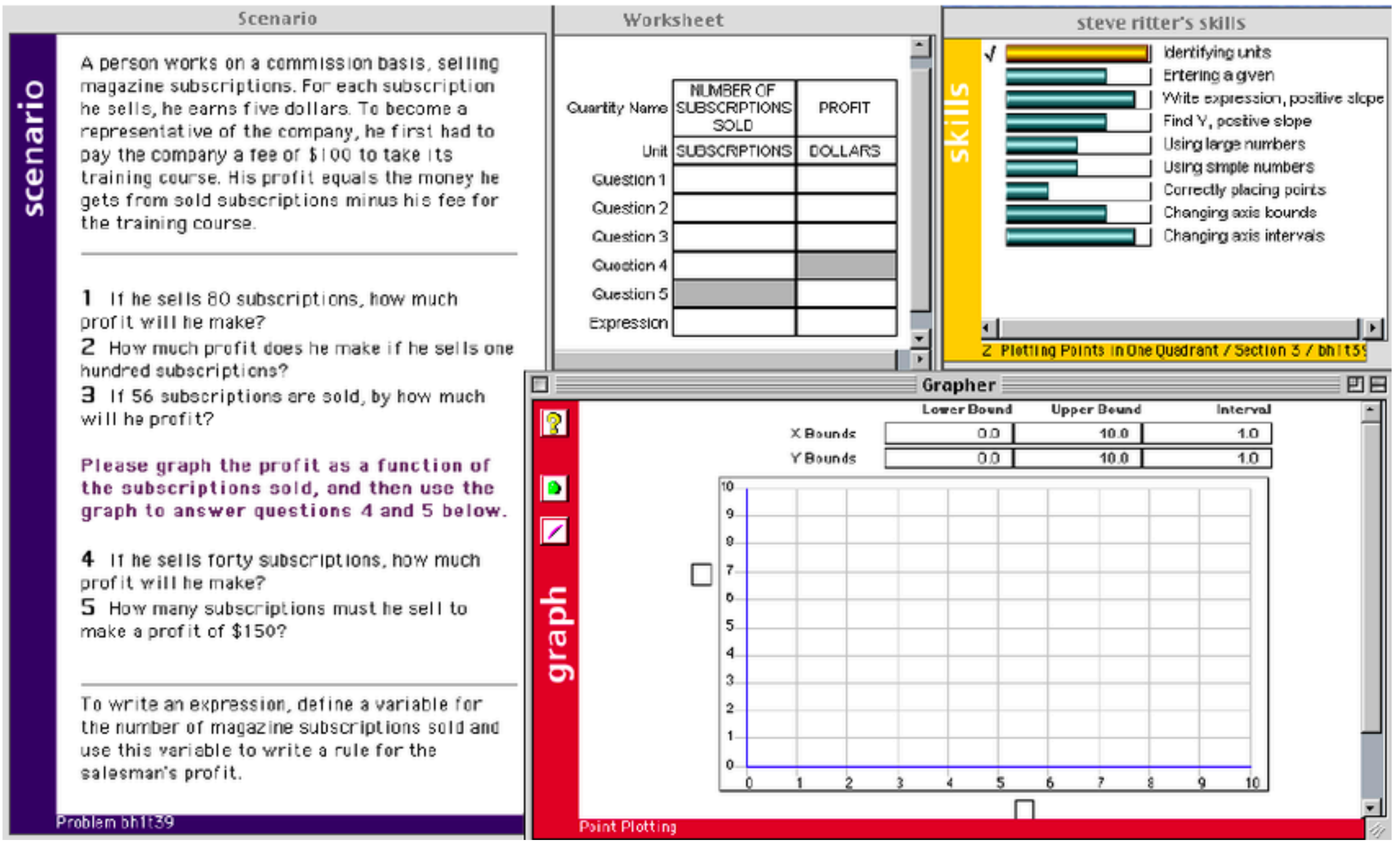
Adaptive Learning Systems

aleks.com

ALEKS

~600,000 students

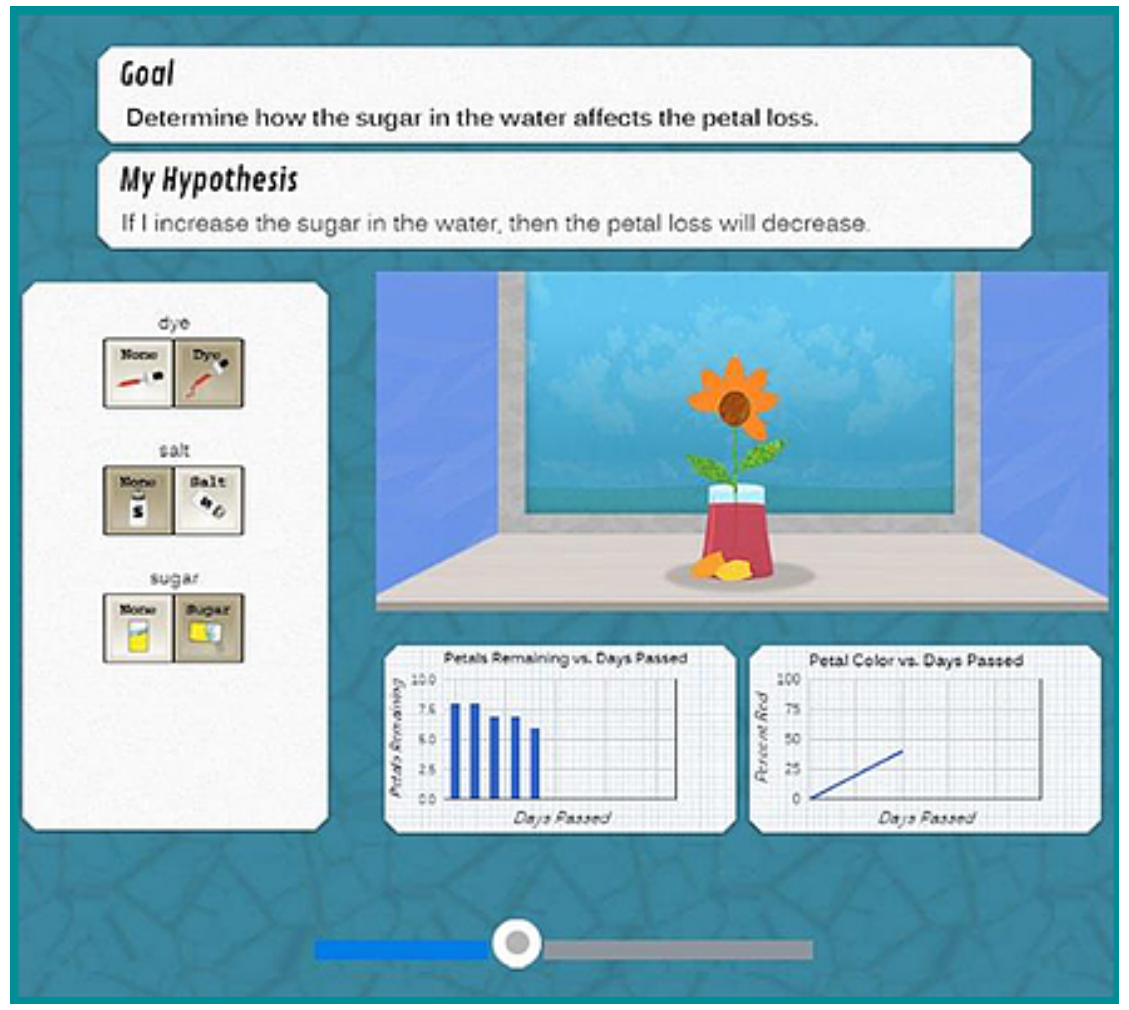
Ritter et al. (2007)



Cognitive Tutor aka Mathia

~500,000 students

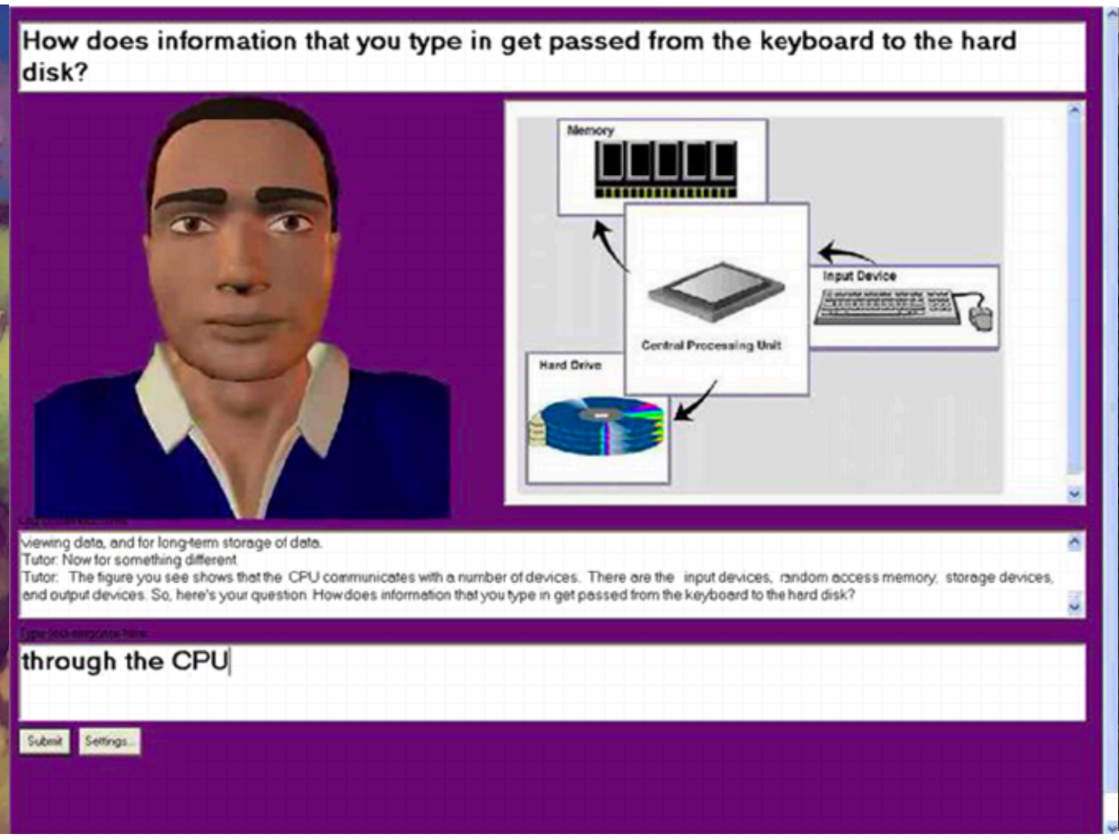
inqits.com



Inq-ITS

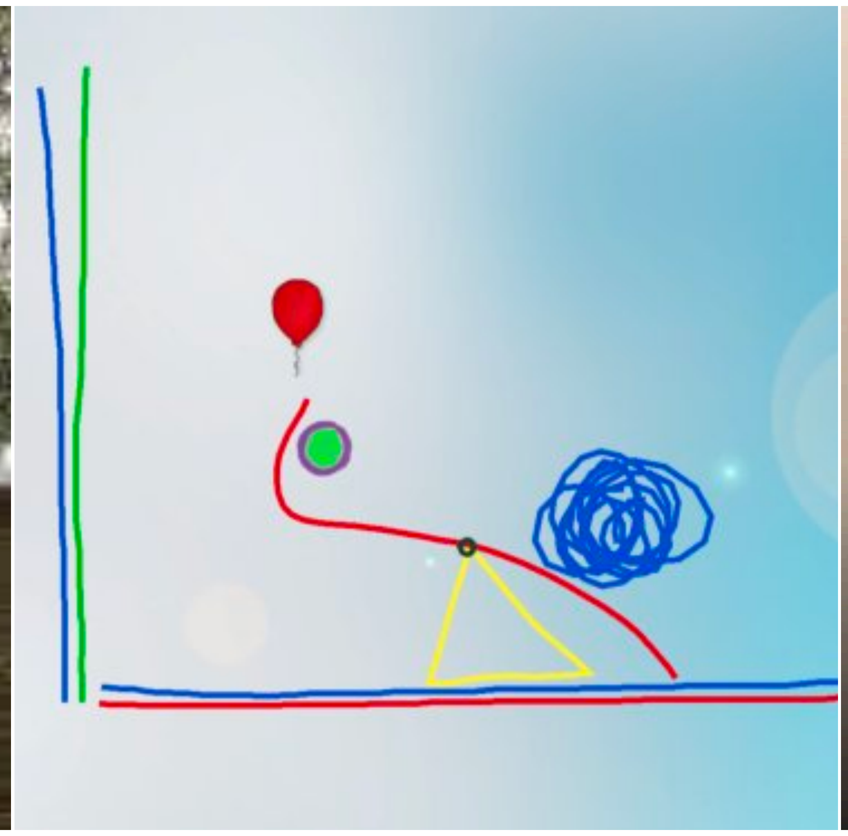
~100,000 students

Zoombinis



AutoTutor

Crystal Island



Physics Playground

TransfrVR

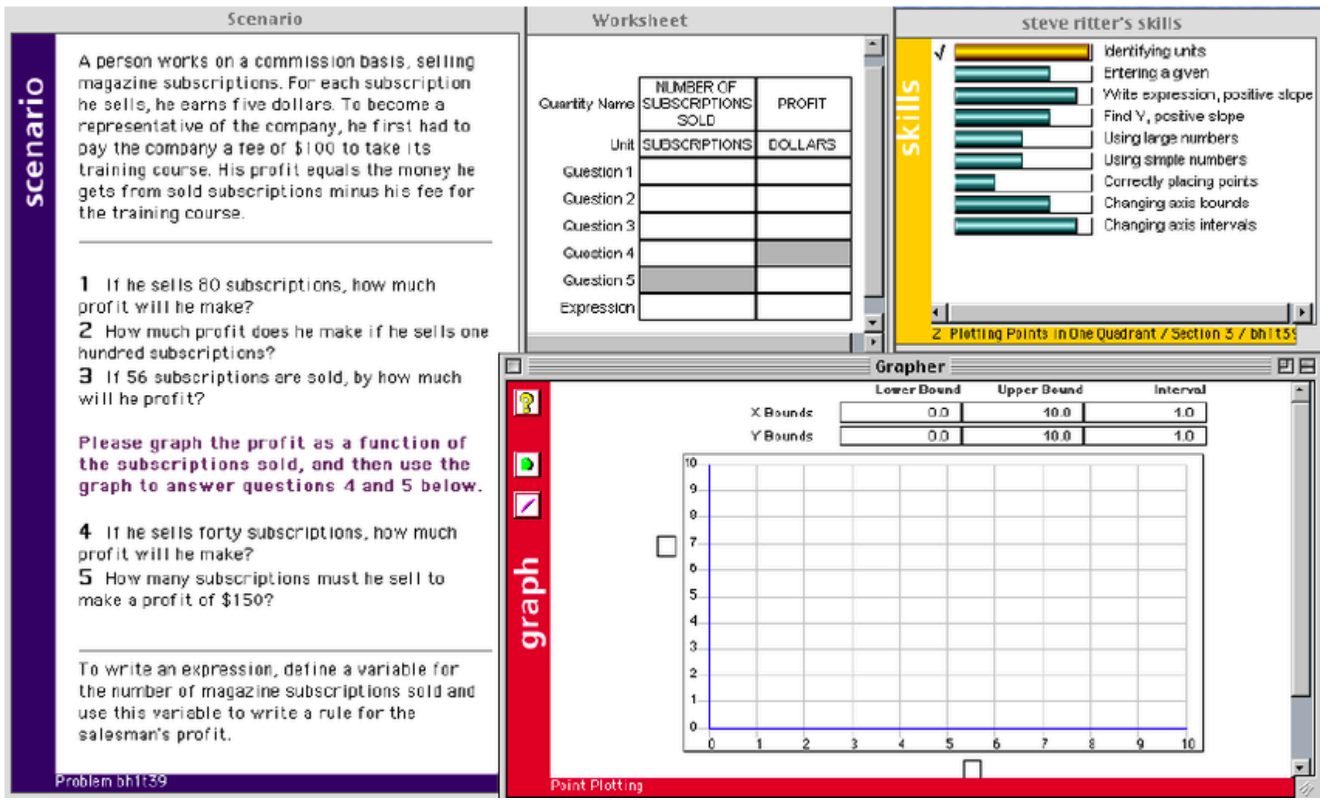
Studying Bias: important, yet challenging

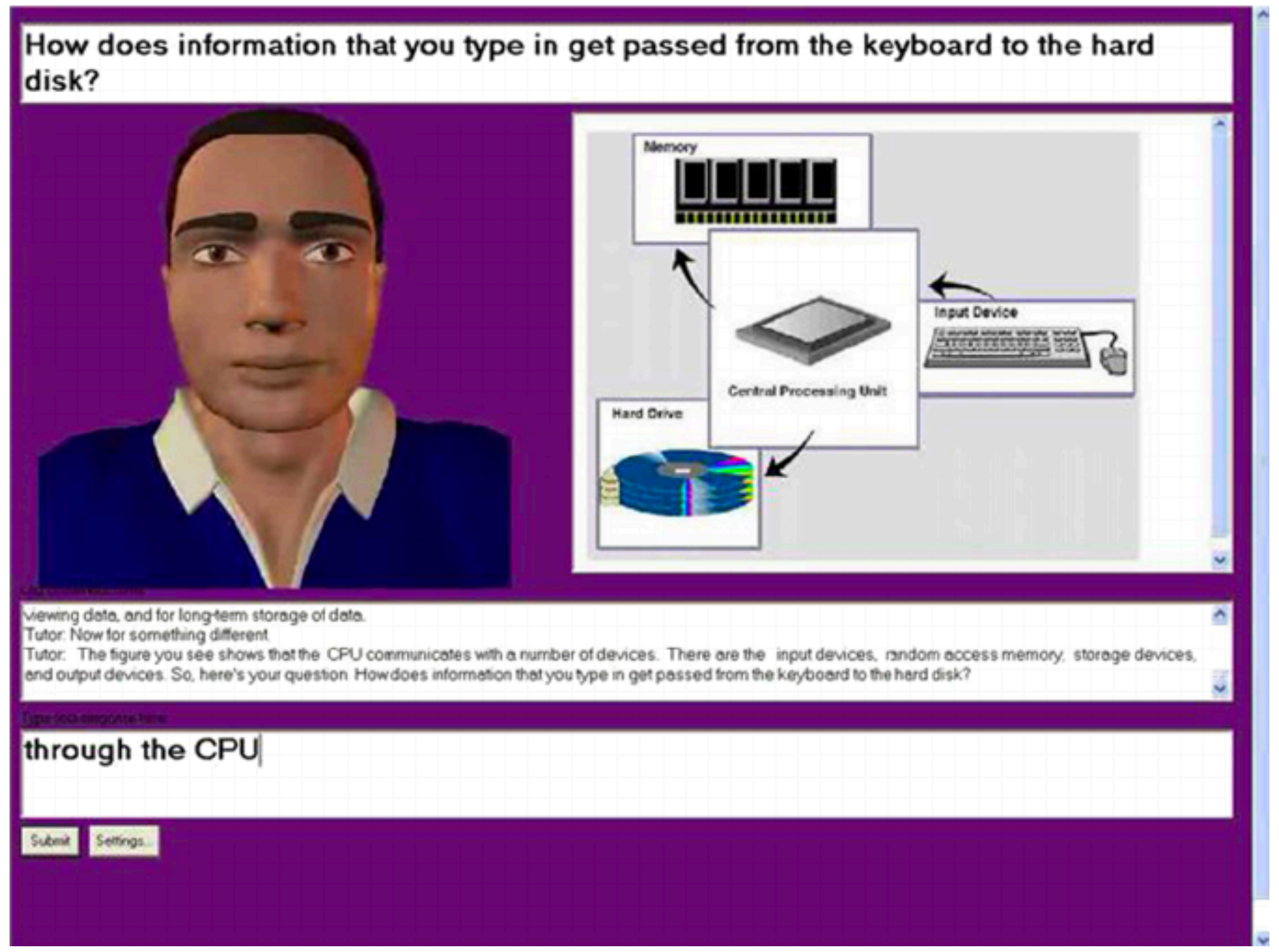
1. Makes real-time decisions that impact students' learning and experiences closely

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

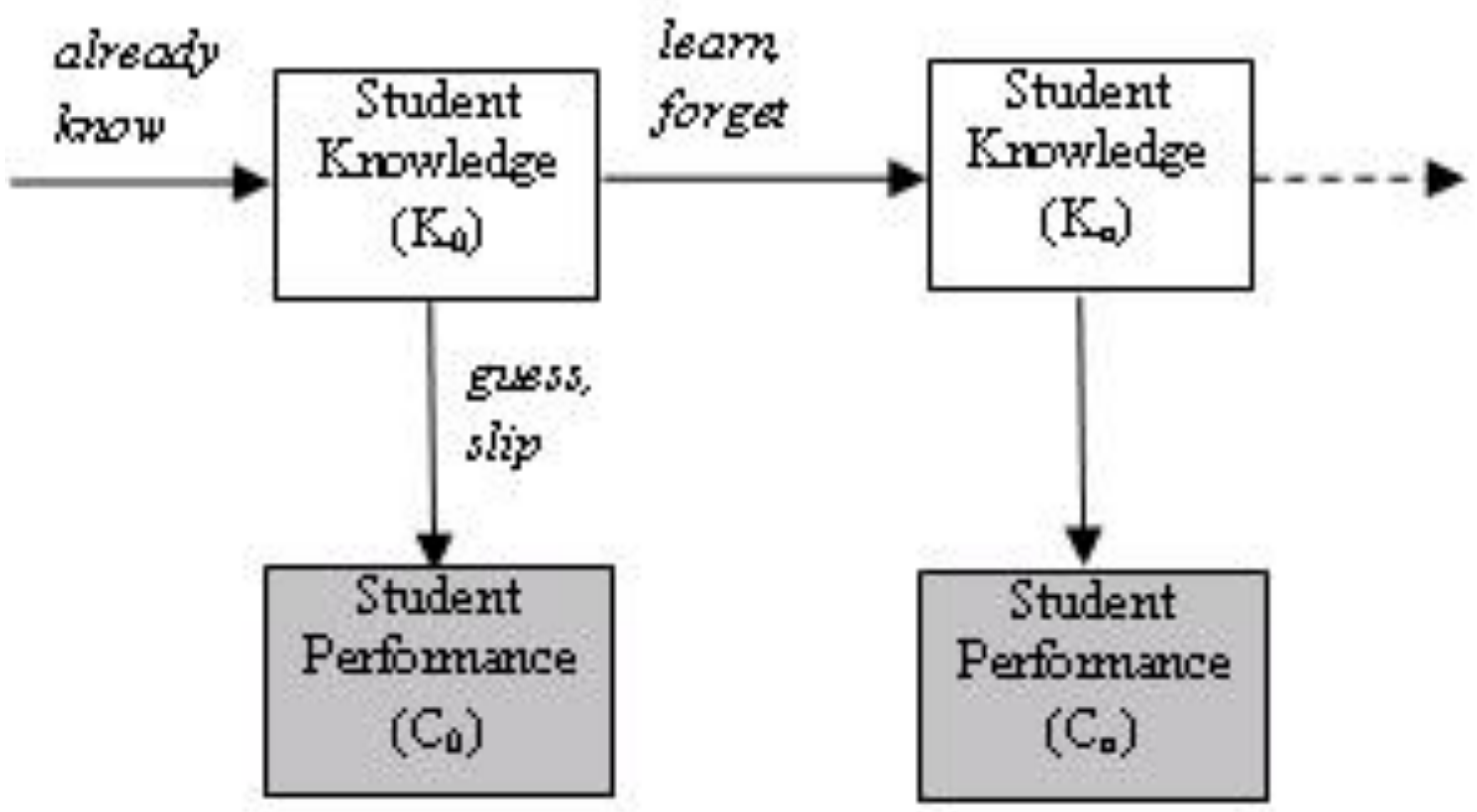




AutoTutor

An Example

Knowledge Tracing



cs.cmu.edu

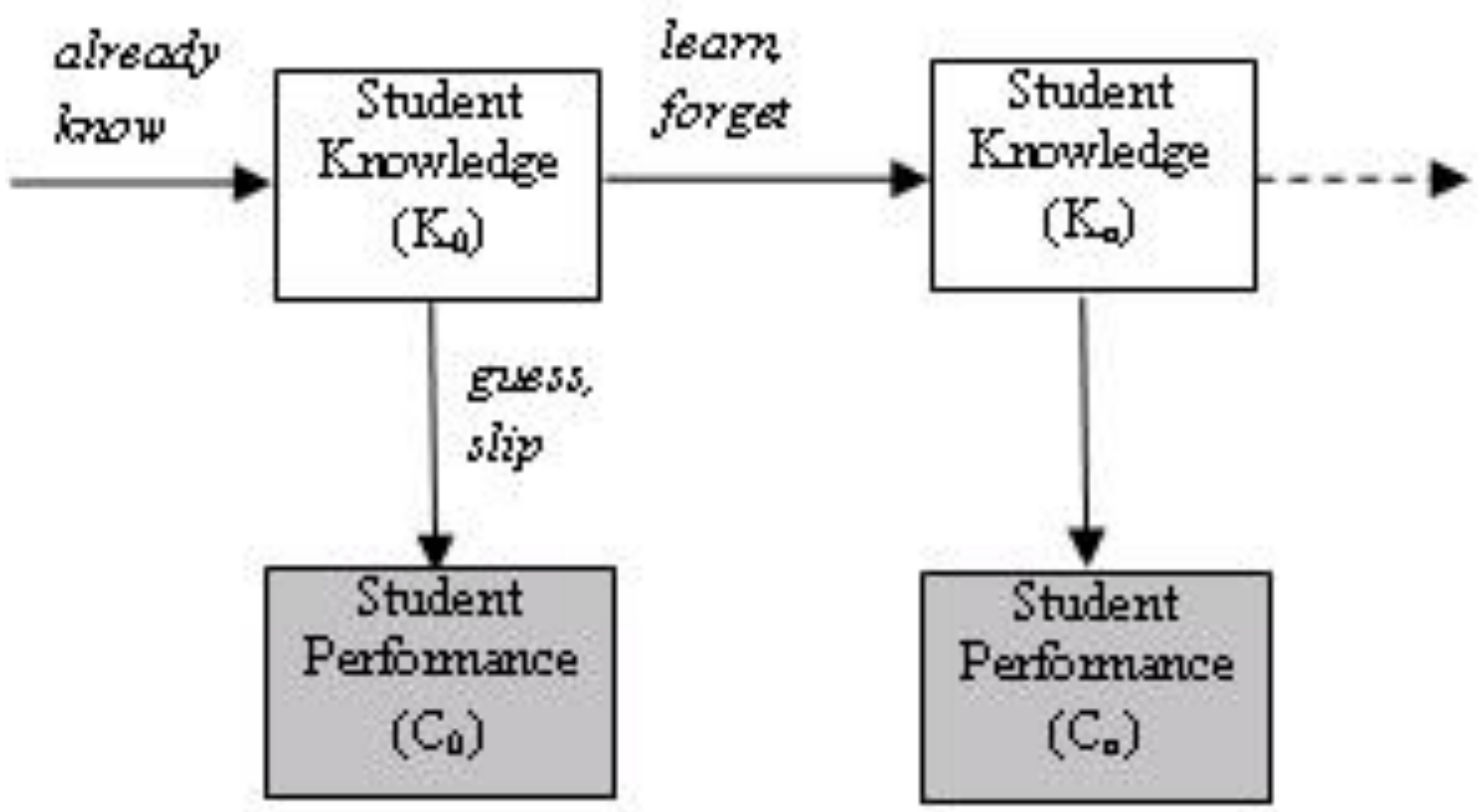
Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge.

Learnta

Baker et al. (2020)

Allocative Harms of Bias

Knowledge Tracing



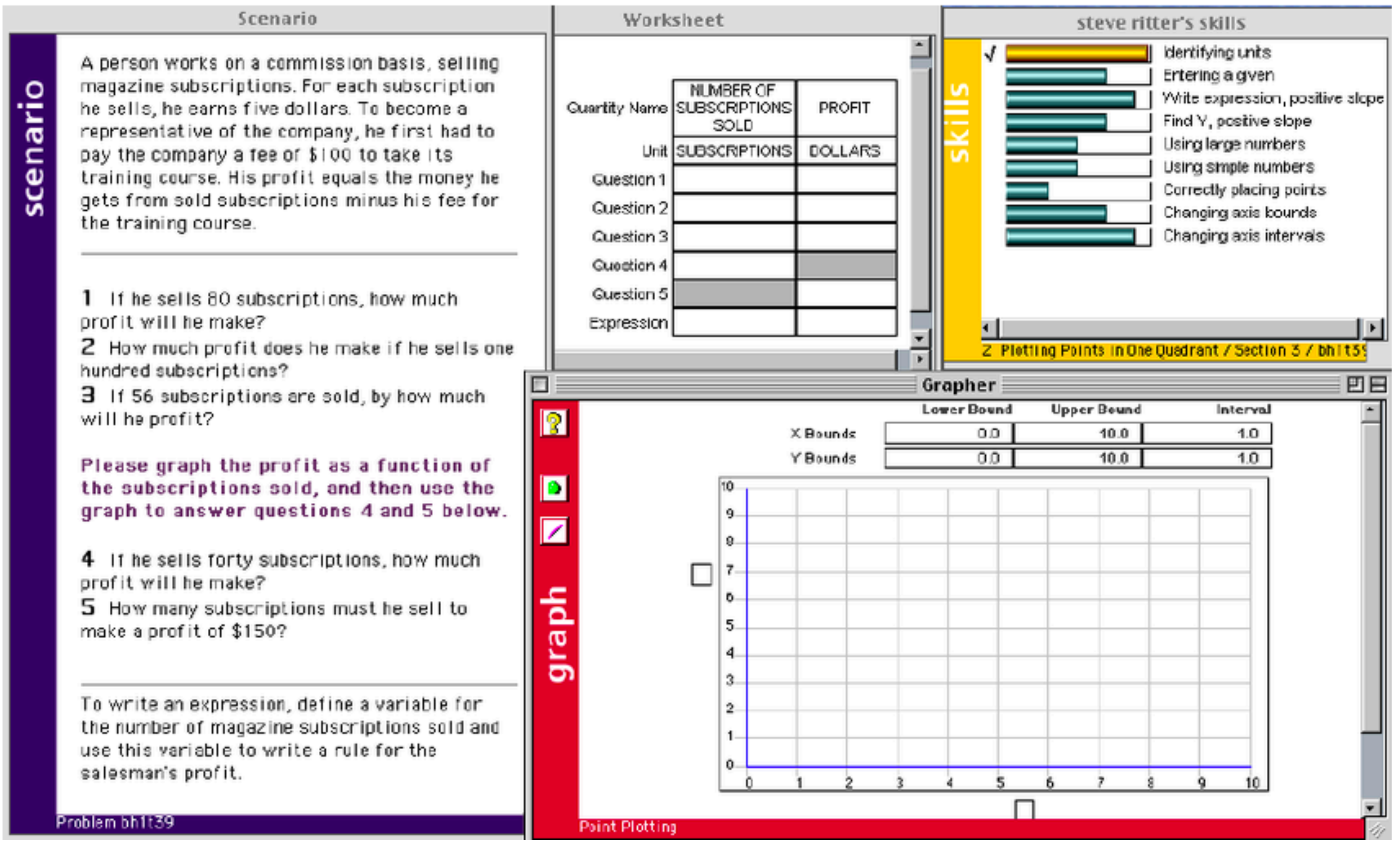
cs.cmu.edu

Francesco Bonchi

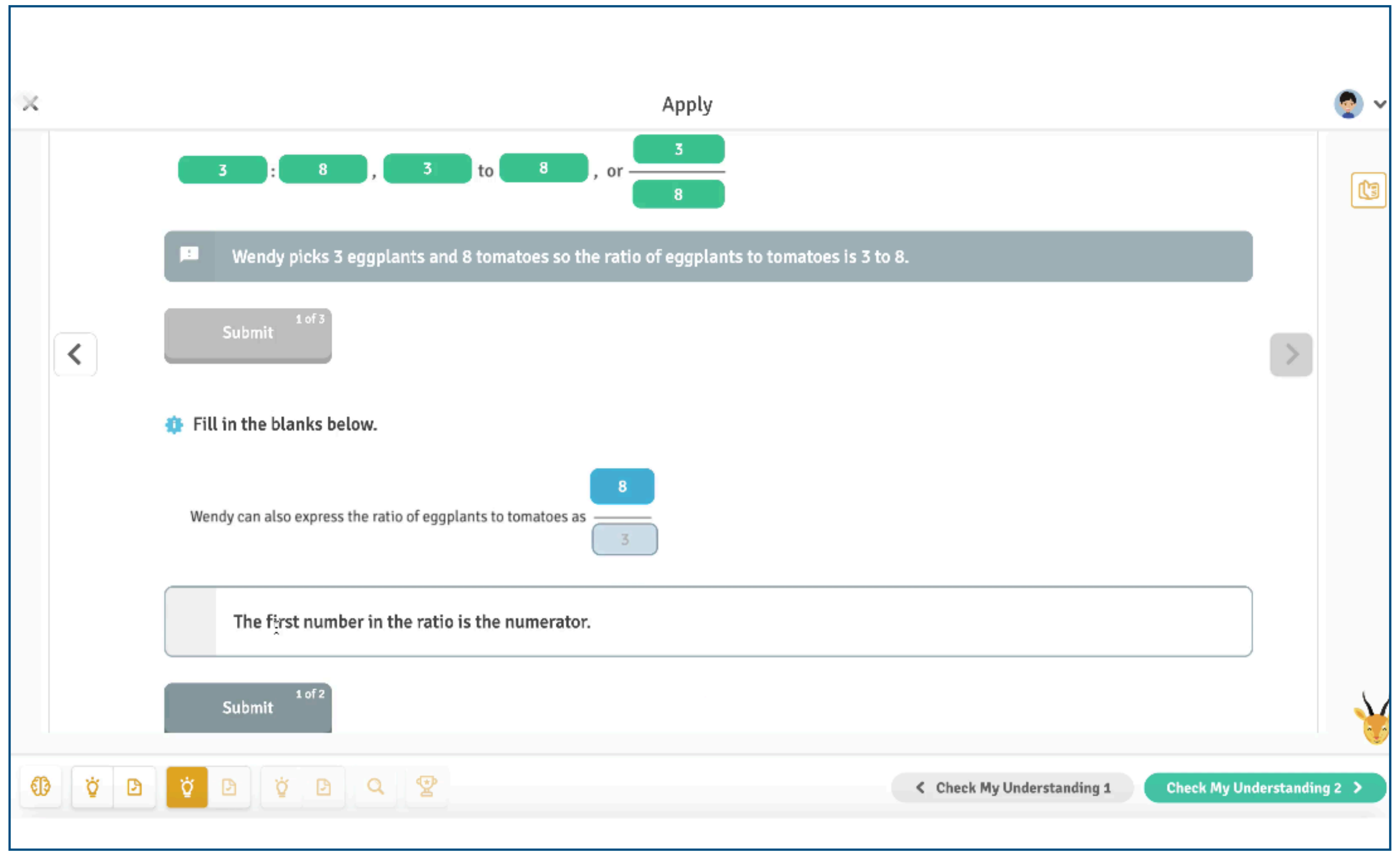
Crawford, K. (2017). The Trouble with Bias. NIPS Keynote.

Diversity in Learner Context

Ritter et al. (2007)



nexgenedu.com



Cognitive Tutor aka Mathia

Used in the United States and Chile

Alef NexGen

Serves students from elementary through high school

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

The Search for Context

Bias in Adaptive Learning Systems

Upstream Sources of Bias

Contextualizing Theoretical Model of Affect

Continued Search For Context

Machine Learning Workflow

Li & Deng (2015)

Annotated data of student emotions

Nye et al. (2018)

DETECTOR

Engaged Concentration

Kearns, CIS 399

- Gather sample S from P $\{<x_1,y_1>, <x_2, y_2>, \ldots <x_n, y_n>\}$
- Design/choose model class H
- Use algorithm/heuristic to find $h \in H$ with small error $\hat{e}(h)$
- Estimate e(h) on **new** data **also from P**

Learning = **generalization**

Fundamental Theorem of Machine Learning

Learning = **generalization**

Kearns, CIS 399

No matter what P looks like…

…and for any reasonable H…

…if we have **enough data** S…

…then for **every h $\in$ H**, we have

$$\hat{e}_S(h) \approx e_P(h)$$

DETECTOR

- Gather sample S from P $\{<x_1,y_1>, <x_2, y_2>, \ldots <x_n, y_n>\}$
- Design/choose model class H
- Use algorithm/heuristic to find h $\in$ H with small error $\hat{e}(h)$
- Estimate e(h) on **new** data **also from P**

Fundamental Theorem of Machine Learning

Learning = **generalization**

Kearns, CIS 399

No matter what P looks like…

…and for any reasonable H…

…if we have **enough data** S…

…then for **every h ∈ H**, we have

minimizing error on **data**
$\approx$
minimizing **true/future** error

DETECTOR

- Gather sample S from P $\{\langle x_1,y_1\rangle, \langle x_2, y_2\rangle,\ldots.\langle x_n, y_{n}\rangle\}$
- Design/choose model class H
- Use algorithm/heuristic to find h ∈ H with small error $\hat{e}(h)$
- Estimate e(h) on **new** data **also from P**

But Learning Context Varies Widely

Generalization to Student Subgroups

Kearns, CIS 399

No matter what P looks like…

…and for any reasonable H…

…if we have **enough data** S…

…then for **every h ∈ H**, we have

minimizing error on **data**

≈

minimizing **true/future** error

DETECTOR

- Gather sample S from P $\{<x_1,y_1>, <x_2, y_2>, \ldots <x_n, y_{n}>\}$
- Design/choose model class H
- Use algorithm/heuristic to find h ∈ H with small error ê(h)
- Estimate e(h) on new data also from P

Student population not reported with generalization estimates (Paquette et al., 2020)

The Problem of Bias

No matter what P looks like…

…and for any reasonable H…

…if we have **enough data** S…

…then for every **h** $\in$ **H**, we have

minimizing error on **data**
$\approx$
minimizing **true/future** error

DETECTOR

"All models are wrong
but some are useful"
- George Box

The Problem of Bias

No matter what P looks like…

…and for any reasonable H…

…if we have **enough data** S…

…then for **every h** $\in$ **H**, we have

minimizing error on **data**

$\approx$

minimizing **true/future** error

DETECTOR

Politico

Haley Falcon

The Wire

Karumbaiah, S., & Brooks, J. (2021) How Colonial Continuities Underlie Algorithmic Injustices in Education. [IEEE RESPECT21]

Current Downstream Efforts

Focus on Model Development and Evaluation

No matter what P looks like…

…and for any reasonable H…

…if we have **enough data** S…

…then for **every h ∈ H**, we have

minimizing error on **data**

$\approx$

minimizing **true/future** error

DETECTOR

Study	Subgroups	Prediction Task
Hu & Rangwala, 2020	Gender, Race	At-Risk (course)
Yu et al., 2020	Gender, Race	College success
Lee & Kizilcec, 2020	Gender, Race	Course grade
Anderson et al., 2019	Gender, Race	Graduation
Kai et al., 2017	Gender, Race	Online college retention
Bridgeman et al., 2009, 2012	Gender, Nationality	Essay scoring
Ogan et al., 2015	Nationality	Learning Outcome

Kizilcec, R. F., & Lee, H. (2020). Algorithmic Fairness in Education.
Baker, R. S., & Hawn, A. (2021). Algorithmic Bias in Education.

Need to Move Upstream

No matter what P looks like…

…and for any reasonable H…

…if we have **enough data** S…

…then for **every h** ∈ **H**, we have

minimizing error on **data**

≈

minimizing **true/future** error

DETECTOR

What upstream sources shape data collection, modeling, and adaptive decision making?

Are they context aware?

Karumbaiah, S., & Brooks, J. (2021) How Colonial Continuities Underlie Algorithmic Injustices in Education. [IEEE RESPECT21]

My dissertation - Upstream Sources of Bias

Focus on Data Collection Method, System Design, and Theory

No matter what P looks like…

…and for any reasonable H…

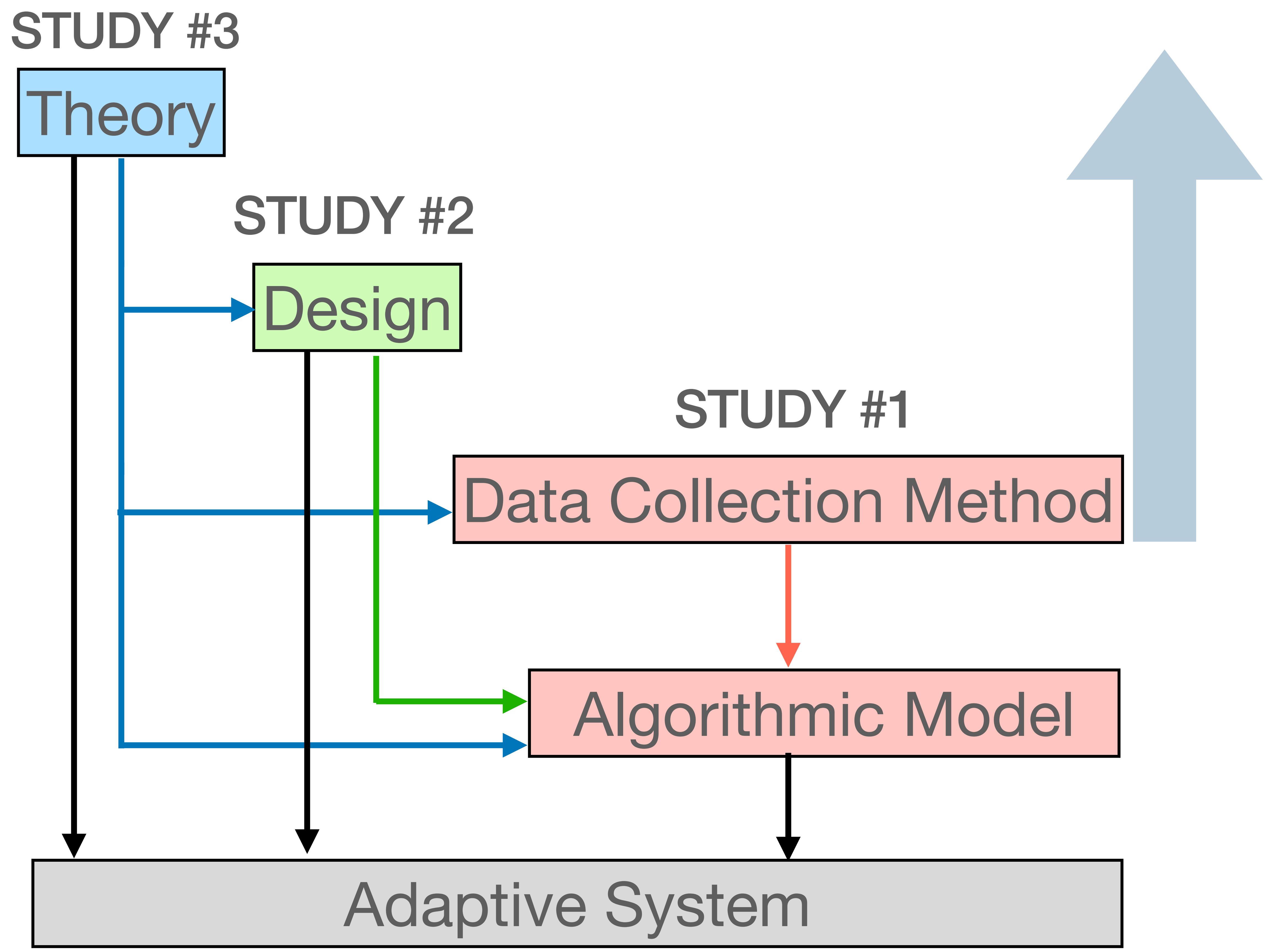
…if we have **enough data** S…

…then for **every h ∈ H**, we have

minimizing error on **data**

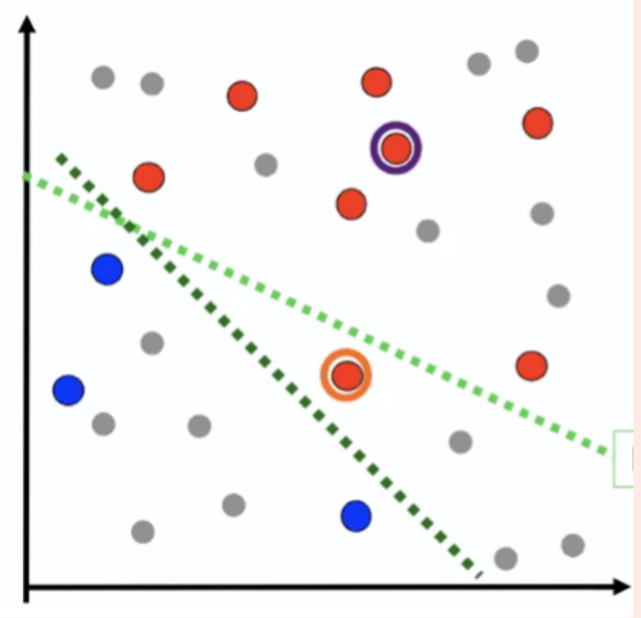
$\approx$

minimizing **true/future** error



My dissertation - Upstream Sources of Bias

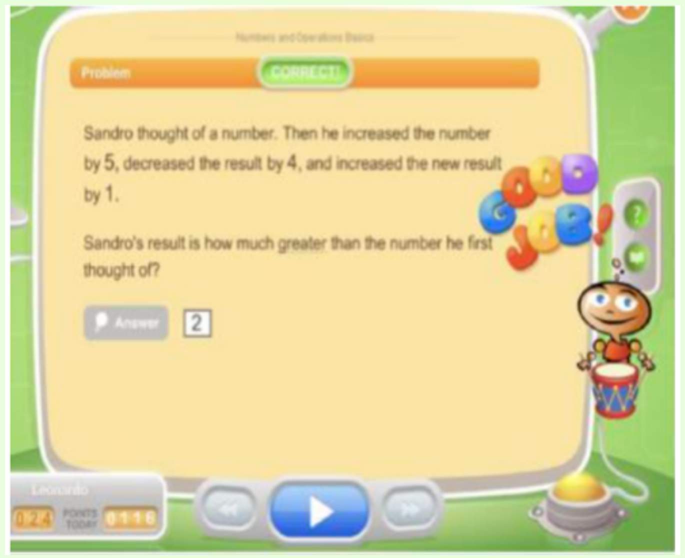
STUDY #1
ANNOTATED DATA COLLECTION



- Active machine learning to improve annotated data collection and cold start problem
- Varying effectiveness of methodological improvements

Karumbaiah, S. et al. (2021)
Using Past Data to Warm Start Active Machine Learning: Does Context Matter? [ACM LAK21]

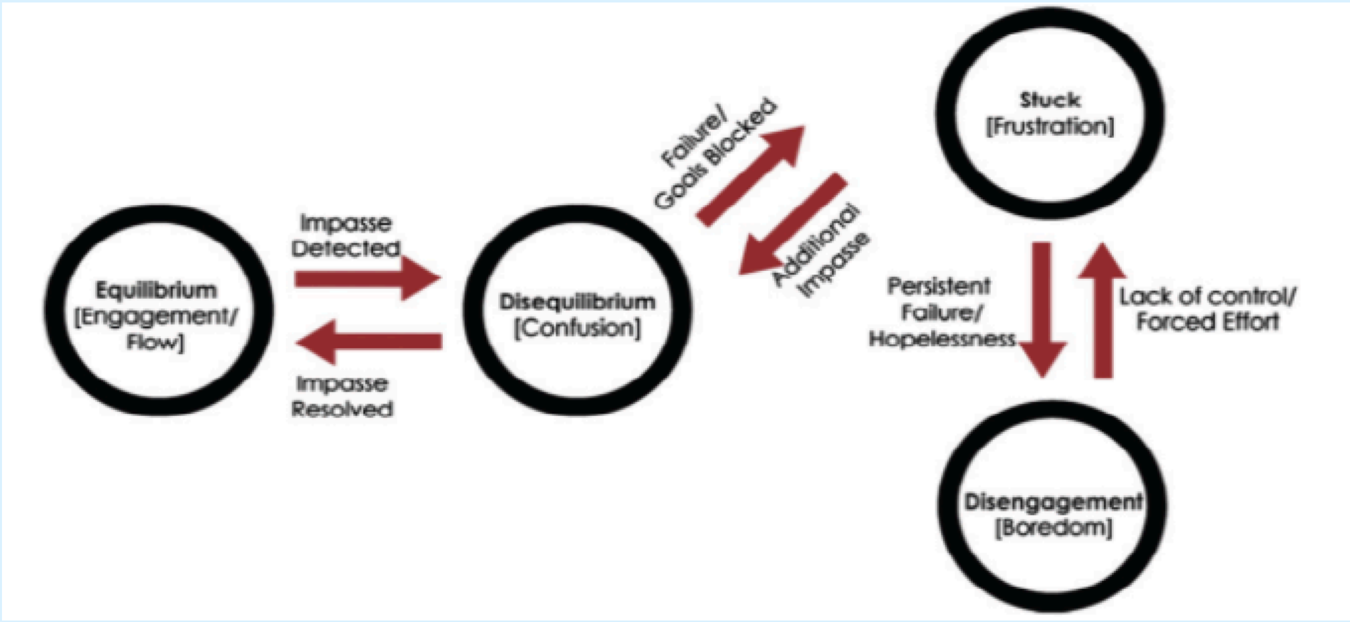
STUDY #2
EDTECH DESIGN



- Differing implications of technology design on student outcomes
- Use of publicly-available, school-level demographics for bias research

Karumbaiah, S. et al. (2021)
Context Matters: Differing Implications of Motivation and Help-Seeking in Educational Technology. [IJAIED21]

STUDY #3
THEORETICAL MODEL



- Generalizability of a widely accepted theory
- Country-level differences in the non-conformance of empirical data

Karumbaiah, S. et al. (2021)
A Re-Analysis and Synthesis of Data on Affect Dynamics in Learning.
[IEEE TAC21]

My dissertation - Upstream Sources of Bias

Graduating December 2021

STUDY #1
ANNOTATED DATA COLLECTION

- Active machine learning to improve annotated data collection and cold start problem
- Varying effectiveness of methodological improvements

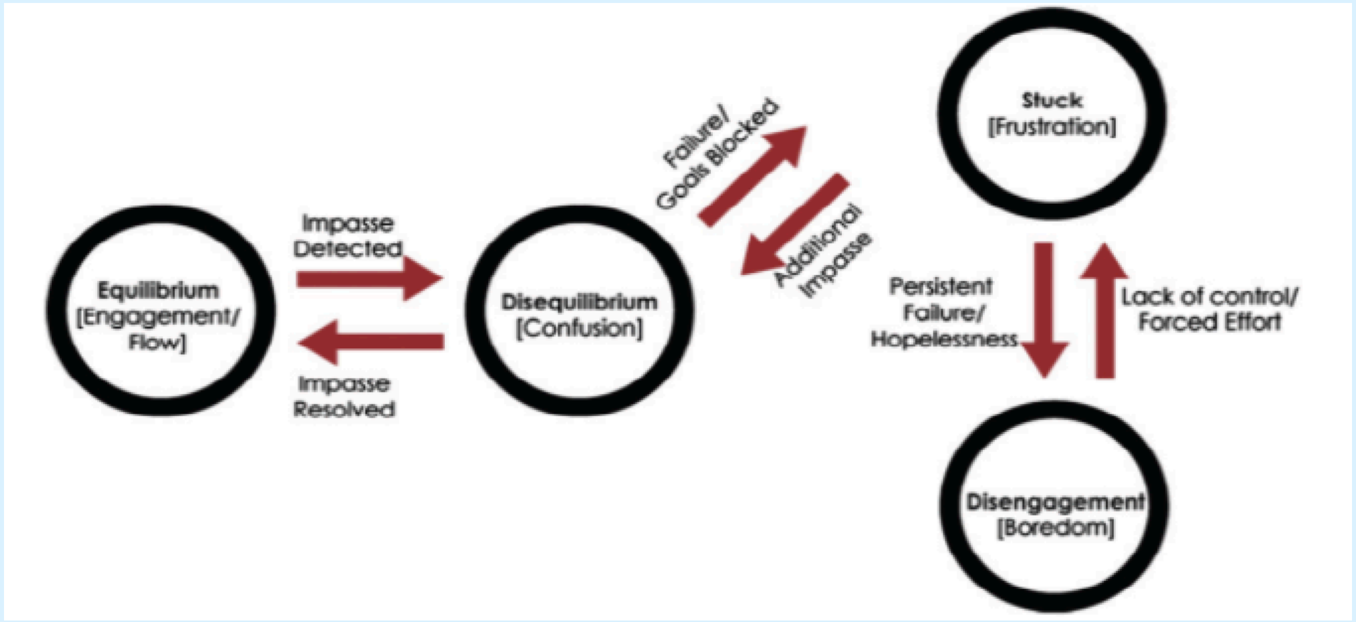
Karumbaiah, S. et al. (2021)
Using Past Data to Warm Start Active Machine Learning: Does Context Matter? [ACM LAK21]

STUDY #2
EDTECH DESIGN

- Differing implications of technology design on student outcomes
- Use of publicly-available, school-level demographics for bias research

Karumbaiah, S. et al. (2021)
Context Matters: Differing Implications of Motivation and Help-Seeking in Educational Technology. [IJAIED21]

STUDY #3
THEORETICAL MODEL



- Generalizability of a widely accepted theory
- Country-level differences in the non-conformance of empirical data

Karumbaiah, S. et al. (2021)
A Re-Analysis and Synthesis of Data on Affect Dynamics in Learning.
[IEEE TAC21]

The Search for Context

Bias in Adaptive Learning Systems

Upstream Sources of Bias

Contextualizing Theoretical Model of Affect

Contextualizing Origins of Bias

Continued Search For Context

Dissertation Study #3: Acknowledgements

- Advisor and dissertation chair: Dr. Ryan Baker (UPenn)
- Committee members: Dr. Rand Quinn (UPenn), Dr. Rene Kizilcec (Cornell)
- Other collaborators: Juliana Ma. Alexandra L. Andres, Dr. Jaclyn Ocumpaugh
- Data: Dr. Douglas DiStefano, Dr. Anna Fisher, Dr. Karrie E. Godwin, Dr. Thea Faye Guia, Dr. Juan Miguel Andres-Bray, Dr. Ryan Baker, Dr. Anthony Botelho, Dr. James Lester, Dr. Ma. Mercedes Rodrigo, and Dr. Jennifer Sabourin

Data Sharing: See BROMPository

Influence of Theory on Adaptive Decisions

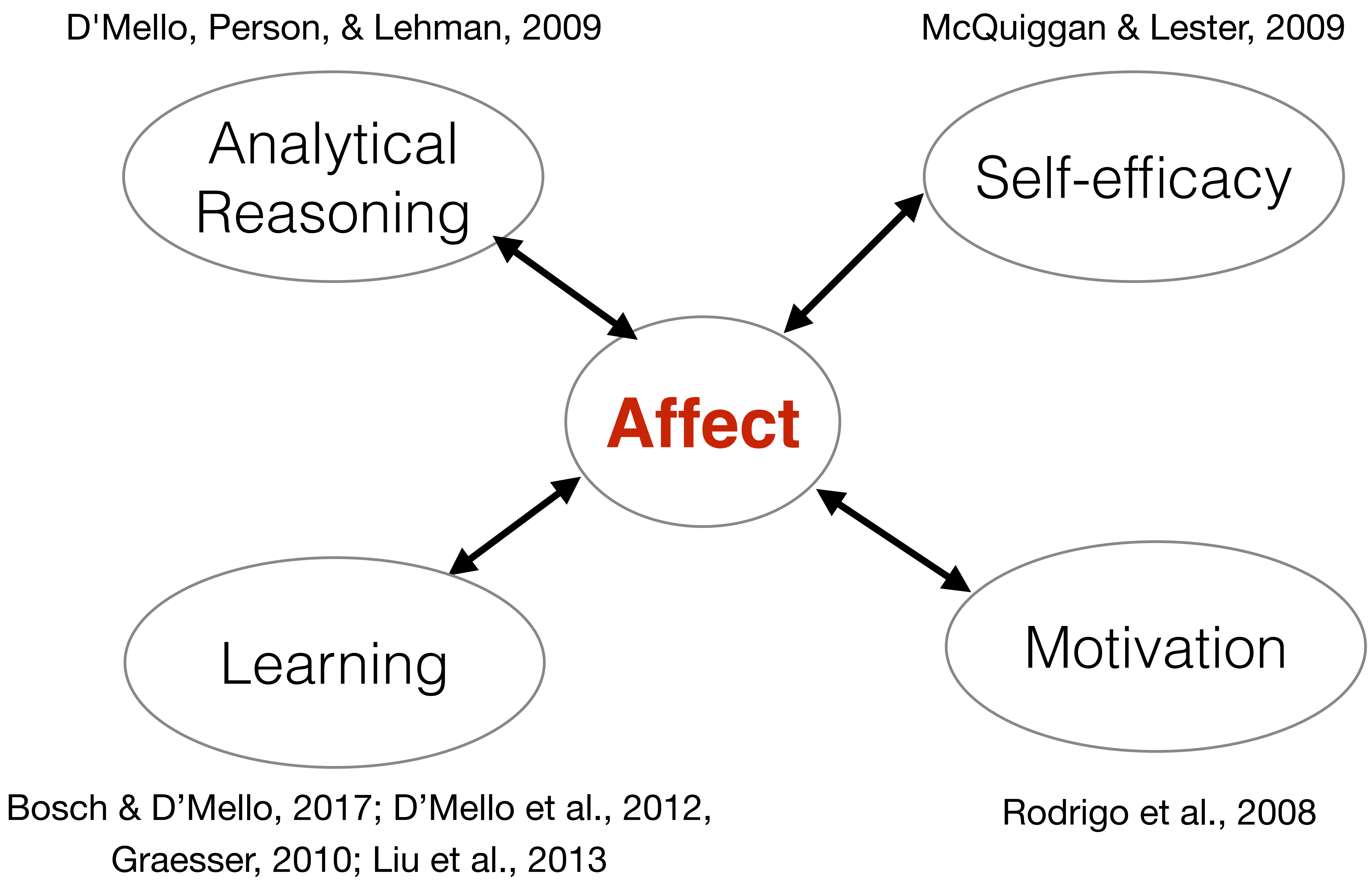
1. Assumptions around the conceptualization of the construct
2. Interpretation of student behaviors in the data collected
3. Construction of variables used in predictive modeling
4. Design of interventions in adaptive systems

Wise, A. F., & Shaffer, D. W. (2015). Why theory matters more than ever in the age of big data. Journal of Learning Analytics.

Role of Affect in Deep Learning

- **Signaling** - draw attention to learning challenges (Schwarz, 2012)
- **Evaluation** - appraise learning (Izard, 2010)
- **Modulation** - guide cognitive focus (Barth & Funke, 2010; D'Mello & Graesser, 2015; Fredrickson & Branigan, 2005; Schwarz, 2012)

Affect in Adaptive Learning Systems

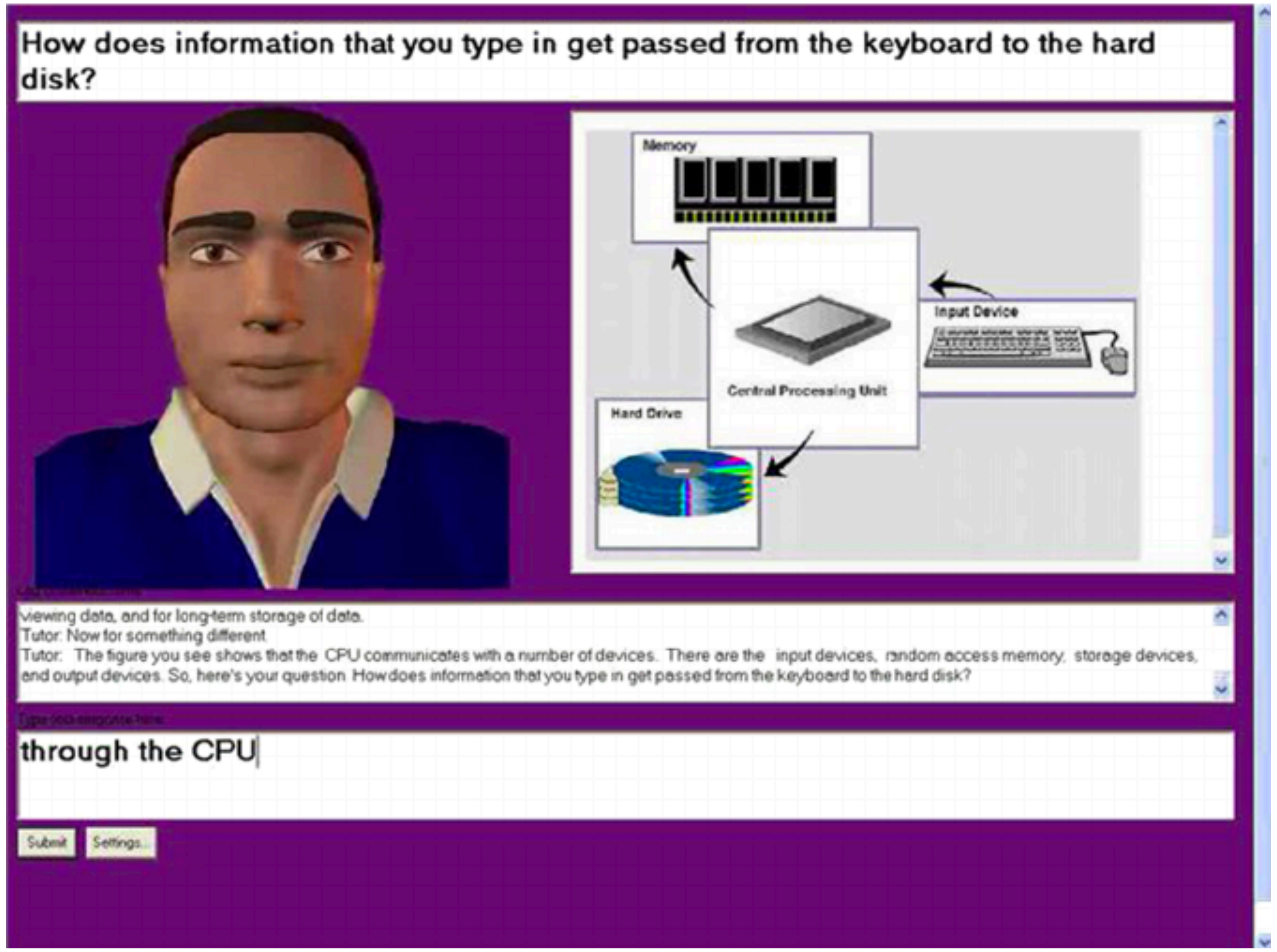


Educational Affective Computing

Recognize, measure, analyze, and respond to student affect to "narrow the communicative gap between the highly emotional human and the emotionally-challenged computer"

Calvo, R. A., & D'Mello, S. (2010). Affect detection: an interdisciplinary review of models, methods, and their applications. IEEE Transactions in Affective Computing.

An Example: Affect Aware Tutors



Affective AutoTutor

D'Mello, S., & Graesser, A. (2013). AutoTutor and affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. ACM Transactions on Interactive Intelligent Systems.

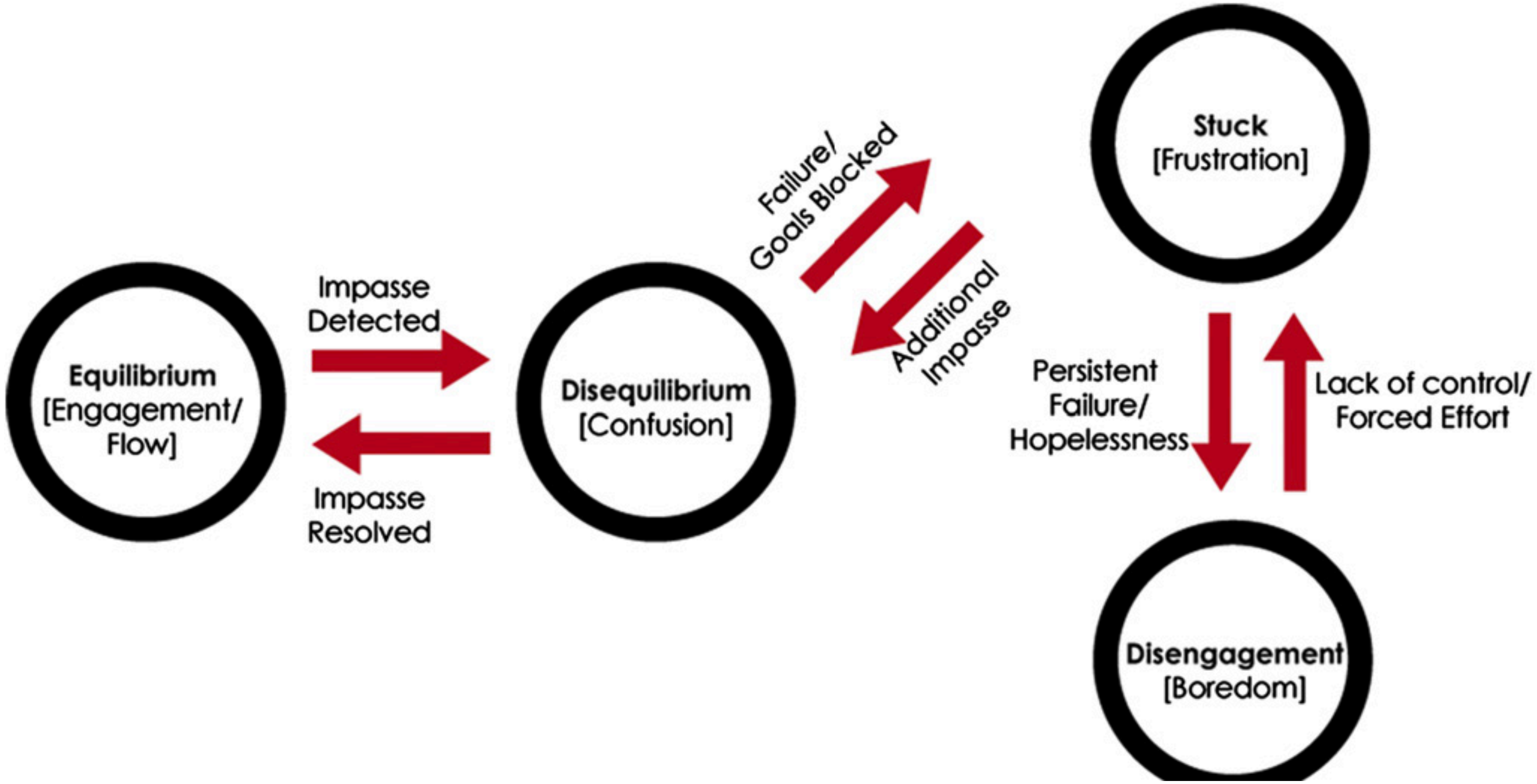
Woolf, B., et al. (2009). Affect-aware tutors: recognising and responding to student affect. International Journal of Learning Technology.

Affect Dynamics

Developing effective interventions that occur in real-time depends on understanding how affect develops and manifests over time.

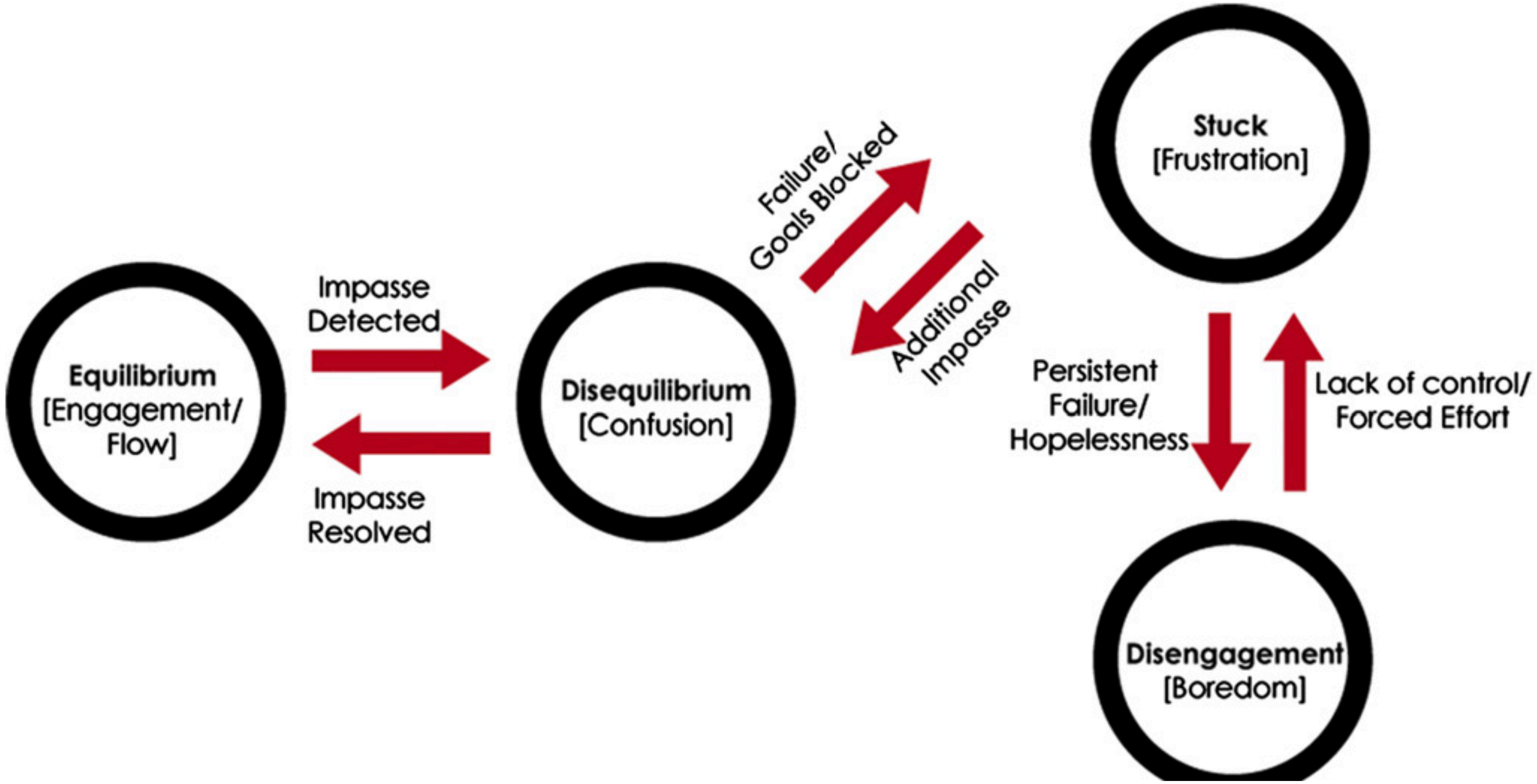
Kuppens, P. (2015). It's about time: A special section on affect dynamics. Emotion Rev.

The Theoretical Model of Affect Dynamics



D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. Learning and Instruction.

Why Study Bias?



A highly cited and widely accepted theoretical model in the field used in diverse learning contexts

#Citations: 666 (as of Nov, 2021)

Goal of the Systematic Review

Investigate the impact of the
methodological and contextual differences
that may be contributing to divergence between the
theoretical model and the empirical results

Karumbaiah, S. et al. (2018) The Implications of a Subtle Difference in the Calculation of Affect Dynamics. [ICCE18]

Karumbaiah et al., 2018

Studies that Show Some Evidence for the Model

	Region	Age	N	School/Grade Population	Learning System	Class v. Lab	Obs. Type/ Grain Size	Obs. Session	Self-trans	Aligned Transitions
Andres & Rodrigo, 2014	Quezon City, PH	13-16	60	Public school	Physics Playground	C	QFO	2hrs	Inc	0
Baker et al., 2007	Manila, PH	14-19	36	High school	Inc. Machine	C	QFO ev. 60s	10min	Inc	0
Bosch & D'Mello, 2013	US	--	29	Undergrads	Unnamed	L	RJP on 100 fixed points	25min	Exc	3
Bosch, & D'Mello, 2017	Midwestern US	17-21	99	Undergrads	Unnamed	L	RJP on 100 fixed points	25min	Exc	5
D'Mello & Graesser, 2012	Southern US	--	28; 30	Undergrads	Auto-Tutor	L	RJP every 20s; fixed points	32min; 35min	Exc	4;5
D'Mello et al., 2007	Southern US	--	28	Undergrads	Auto-Tutor	L	RJP ev. 20s	32min	Inc	2
D'Mello et al., 2009	Southern US	--	41	Undergrads	Unnamed	L	RJP on fixed points	35min	Exc	1
D'Mello & Graesser, 2010	Southern US	--	28; 30	Undergrads	Auto-Tutor	L	RJP ev. 20s; fixed points	32min; 35min	Exc	3;3
Guia et al., 2011; 2013	Quezon City, PH	18-20	60	Undergrads	SQL Tutor	C	QFO ev. 200s	1hr	Inc	0
McQuiggan et al., 2008; 2010	US	21-60	35	Grad students	Crystal Island	L	SRI	35min	Inc	1
Ocumpaugh et al., 2017	New York, US	18-22	108	West Point	vMedic	C	QFO ev.122s	--	Inc	2
Rodrigo et al., 2008	Quezon City & Cavite Prov., PH	9-13	180	Private school	Ecolab	C	QFO	40min	Inc	1
Rodrigo et al., 2011; 2012	Quezon City, PH	12-14	126	High school	Scatterplot Tutor	C	QFO ev. 200s	80min	Inc	1

* PH: Philippines, QFO: Qualitative field observation, RJP: Retrospective judgment protocol, SRI: self-report based on interactions, Inc: self transitions included, Exc: self transitions excluded

1. Student Demographics

	Region	Age	N	School/Grade Population	Learning System	Class v. Lab	Obs. Type/ Grain Size	Obs. Session	Self-trans	Aligned Transitions
Andres & Rodrigo, 2014	Quezon City, PH	13-16	60	Public school	Physics Playground	C	QFO	2hrs	Inc	0
Baker et al., 2007	Manila, PH	14-19	36	High school	Inc. Machine	C	QFO ev. 60s	10min	Inc	0
Bosch & D'Mello, 2013	US	--	29	Undergrads	Unnamed	L	RJP on 100 fixed points	25min	Exc	3
Bosch, & D'Mello, 2017	Midwestern US	17-21	99	Undergrads	Unnamed	L	RJP on 100 fixed points	25min	Exc	5
D'Mello & Graesser, 2012	Southern US	--	28; 30	Undergrads	Auto-Tutor	L	RJP every 20s; fixed points	32min; 35min	Exc	4;5
D'Mello et al., 2007	Southern US	--	28	Undergrads	Auto-Tutor	L	RJP ev. 20s	32min	Inc	2
D'Mello et al., 2009	Southern US	--	41	Undergrads	Unnamed	L	RJP on fixed points	35min	Exc	1
D'Mello & Graesser, 2010	Southern US	--	28; 30	Undergrads	Auto-Tutor	L	RJP ev. 20s; fixed points	32min; 35min	Exc	3;3
Guia et al., 2011; 2013	Quezon City, PH	18-20	60	Undergrads	SQL Tutor	C	QFO ev. 200s	1hr	Inc	0
McQuiggan et al., 2008; 2010	US	21-60	35	Grad students	Crystal Island	L	SRI	35min	Inc	1
Ocumpaugh et al., 2017	New York, US	18-22	108	West Point	vMedic	C	QFO ev.122s	--	Inc	2
Rodrigo et al., 2008	Quezon City & Cavite Prov., PH	9-13	180	Private school	Ecolab	C	QFO	40min	Inc	1
Rodrigo et al., 2011; 2012	Quezon City, PH	12-14	126	High school	Scatterplot Tutor	C	QFO ev. 200s	80min	Inc	1

* PH: Philippines, QFO: Qualitative field observation, RJP: Retrospective judgment protocol, SRI: self-report based on interactions, Inc: self transitions included, Exc: self transitions excluded

2. Learning Settings

	Region	Age	N	School/Grade Population	Learning System	Class v. Lab	Obs. Type/ Grain Size	Obs. Session	Self-trans	Aligned Transitions
Andres & Rodrigo, 2014	Quezon City, PH	13-16	60	Public school	Physics Playground	C	QFO	2hrs	Inc	0
Baker et al., 2007	Manila, PH	14-19	36	High school	Inc. Machine	C	QFO ev. 60s	10min	Inc	0
Bosch & D'Mello, 2013	US	--	29	Undergrads	Unnamed	L	RJP on 100 fixed points	25min	Exc	3
Bosch, & D'Mello, 2017	Midwestern US	17-21	99	Undergrads	Unnamed	L	RJP on 100 fixed points	25min	Exc	5
D'Mello & Graesser, 2012	Southern US	--	28; 30	Undergrads	Auto-Tutor	L	RJP every 20s; fixed points	32min; 35min	Exc	4;5
D'Mello et al., 2007	Southern US	--	28	Undergrads	Auto-Tutor	L	RJP ev. 20s	32min	Inc	2
D'Mello et al., 2009	Southern US	--	41	Undergrads	Unnamed	L	RJP on fixed points	35min	Exc	1
D'Mello & Graesser, 2010	Southern US	--	28; 30	Undergrads	Auto-Tutor	L	RJP ev. 20s; fixed points	32min; 35min	Exc	3;3
Guia et al., 2011; 2013	Quezon City, PH	18-20	60	Undergrads	SQL Tutor	C	QFO ev. 200s	1hr	Inc	0
McQuiggan et al., 2008; 2010	US	21-60	35	Grad students	Crystal Island	L	SRI	35min	Inc	1
Ocumpaugh et al., 2017	New York, US	18-22	108	West Point	vMedic	C	QFO ev.122s	--	Inc	2
Rodrigo et al., 2008	Quezon City & Cavite Prov., PH	9-13	180	Private school	Ecolab	C	QFO	40min	Inc	1
Rodrigo et al., 2011; 2012	Quezon City, PH	12-14	126	High school	Scatterplot Tutor	C	QFO ev. 200s	80min	Inc	1

* PH: Philippines, QFO: Qualitative field observation, RJP: Retrospective judgment protocol, SRI: self-report based on interactions, Inc: self transitions included, Exc: self transitions excluded

3. Data Collection Procedure

	Region	Age	N	School/Grade Population	Learning System	Class v. Lab	Obs. Type/ Grain Size	Obs. Session	Self-trans	Aligned Transitions
Andres & Rodrigo, 2014	Quezon City, PH	13-16	60	Public school	Physics Playground	C	QFO	2hrs	Inc	0
Baker et al., 2007	Manila, PH	14-19	36	High school	Inc. Machine	C	QFO ev. 60s	10min	Inc	0
Bosch & D'Mello, 2013	US	--	29	Undergrads	Unnamed	L	RJP on 100 fixed points	25min	Exc	3
Bosch, & D'Mello, 2017	Midwestern US	17-21	99	Undergrads	Unnamed	L	RJP on 100 fixed points	25min	Exc	5
D'Mello & Graesser, 2012	Southern US	--	28; 30	Undergrads	Auto-Tutor	L	RJP every 20s; fixed points	32min; 35min	Exc	4;5
D'Mello et al., 2007	Southern US	--	28	Undergrads	Auto-Tutor	L	RJP ev. 20s	32min	Inc	2
D'Mello et al., 2009	Southern US	--	41	Undergrads	Unnamed	L	RJP on fixed points	35min	Exc	1
D'Mello & Graesser, 2010	Southern US	--	28; 30	Undergrads	Auto-Tutor	L	RJP ev. 20s; fixed points	32min; 35min	Exc	3;3
Guia et al., 2011; 2013	Quezon City, PH	18-20	60	Undergrads	SQL Tutor	C	QFO ev. 200s	1hr	Inc	0
McQuiggan et al., 2008; 2010	US	21-60	35	Grad students	Crystal Island	L	SRI	35min	Inc	1
Ocumpaugh et al., 2017	New York, US	18-22	108	West Point	vMedic	C	QFO ev.122s	--	Inc	2
Rodrigo et al., 2008	Quezon City & Cavite Prov., PH	9-13	180	Private school	Ecolab	C	QFO	40min	Inc	1
Rodrigo et al., 2011; 2012	Quezon City, PH	12-14	126	High school	Scatterplot Tutor	C	QFO ev. 200s	80min	Inc	1

* PH: Philippines, QFO: Qualitative field observation, RJP: Retrospective judgment protocol, SRI: self-report based on interactions, Inc: self transitions included, Exc: self transitions excluded

4. Exclusion of Self-Transitions

	Region	Age	N	School/Grade Population	Learning System	Class v. Lab	Obs. Type/ Grain Size	Obs. Session	Self-trans	Aligned Transitions
Andres & Rodrigo, 2014	Quezon City, PH	13-16	60	Public school	Physics Playground	C	QFO	2hrs	Inc	0
Baker et al., 2007	Manila, PH	14-19	36	High school	Inc. Machine	C	QFO ev. 60s	10min	Inc	0
Bosch & D'Mello, 2013	US	--	29	Undergrads	Unnamed	L	RJP on 100 fixed points	25min	Exc	3
Bosch, & D'Mello, 2017	Midwestern US	17-21	99	Undergrads	Unnamed	L	RJP on 100 fixed points	25min	Exc	5
D'Mello & Graesser, 2012	Southern US	--	28; 30	Undergrads	Auto-Tutor	L	RJP every 20s; fixed points	32min; 35min	Exc	4;5
D'Mello et al., 2007	Southern US	--	28	Undergrads	Auto-Tutor	L	RJP ev. 20s	32min	Inc	2
D'Mello et al., 2009	Southern US	--	41	Undergrads	Unnamed	L	RJP on fixed points	35min	Exc	1
D'Mello & Graesser, 2010	Southern US	--	28; 30	Undergrads	Auto-Tutor	L	RJP ev. 20s; fixed points	32min; 35min	Exc	3;3
Guia et al., 2011; 2013	Quezon City, PH	18-20	60	Undergrads	SQL Tutor	C	QFO ev. 200s	1hr	Inc	0
McQuiggan et al., 2008; 2010	US	21-60	35	Grad students	Crystal Island	L	SRI	35min	Inc	1
Ocumpaugh et al., 2017	New York, US	18-22	108	West Point	vMedic	C	QFO ev.122s	--	Inc	2
Rodrigo et al., 2008	Quezon City & Cavite Prov., PH	9-13	180	Private school	Ecolab	C	QFO	40min	Inc	1
Rodrigo et al., 2011; 2012	Quezon City, PH	12-14	126	High school	Scatterplot Tutor	C	QFO ev. 200s	80min	Inc	1

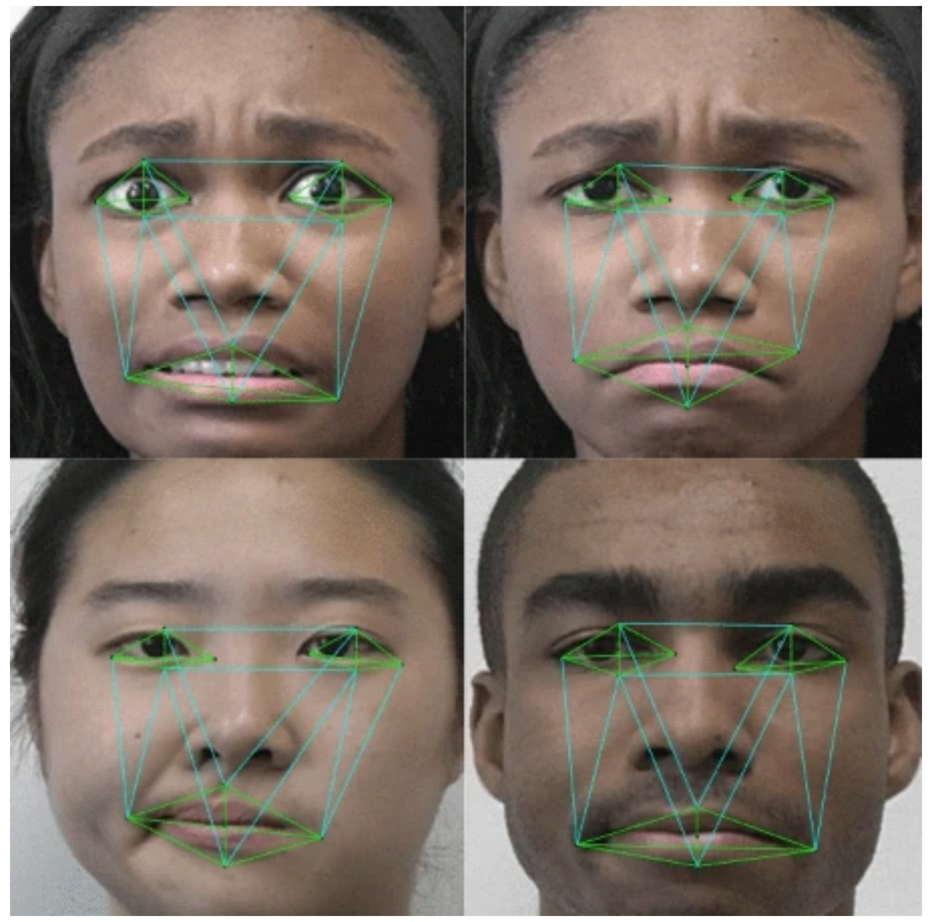
* PH: Philippines, QFO: Qualitative field observation, RJP: Retrospective judgment protocol, SRI: self-report based on interactions, Inc: self transitions included, Exc: self transitions excluded

Affect Data

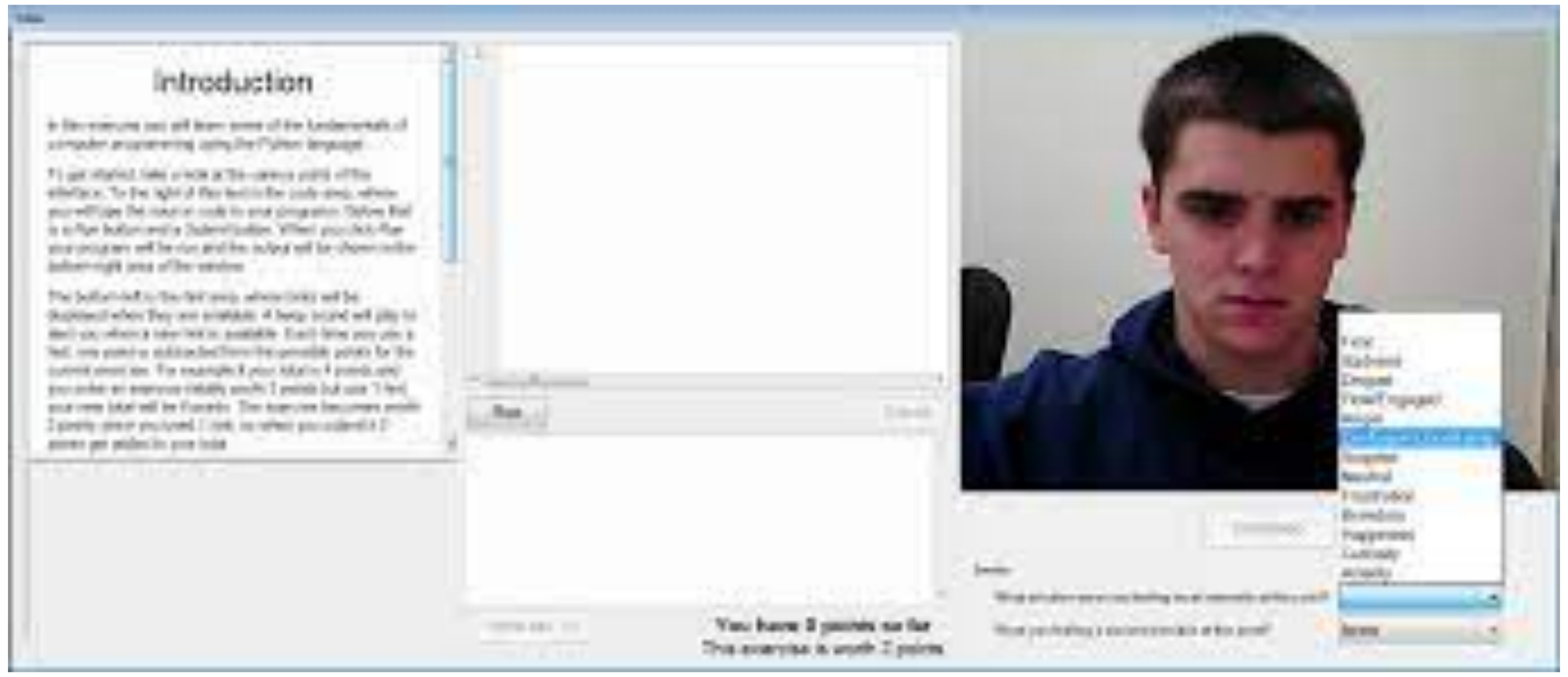
Time Step (seconds)	Affect
10	FLO
20	CON
30	CON
40	CON
50	FRU
60	FRU
70	BOR
….	
3600	FLO

FLOw, CONfusion, FRUstration, BORed

BROMP



Joseph & Geetha



Bosch

Affect Data

Time Step (seconds)	Affect
10	FLO
20	CON
30	CON
40	CON
50	FRU
60	FRU
70	BOR
….	
3600	FLO

FLOw, CONfusion, FRUstration, BORed

Affect Sequence

FLO, CON, CON, CON, FRU, FRU, BOR …. FLO

Studies with no evidence

Include Self Transition

Studies that show some evidence

FLO, **CON, CON, CON**, **FRU, FRU**, BOR …. FLO

Exclude Self Transition

FLO, **CON**, **FRU**, BOR …. FLO

Violation of Independence Assumption

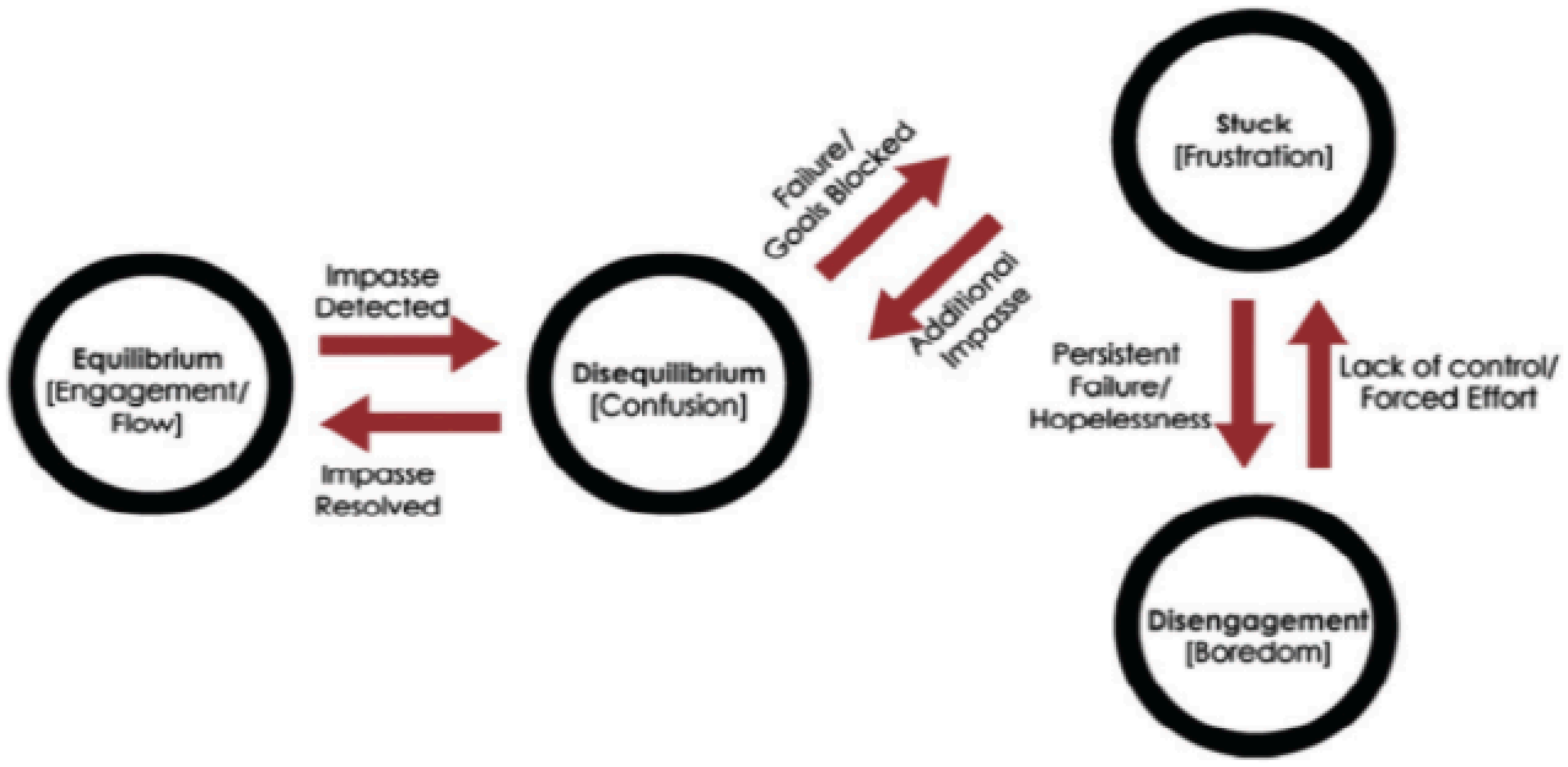
D'Mello & Graesser, 2012

perfect agreement) minus expected agreement. The likelihood metric proceeds in the same way, however, rather than agreement, we use conditional probability as a measure of association (see Eq. (2)). The expected degree of association is $\Pr(M_{t+1})$, because if M_{t+1} and M_t are independent, then $\Pr(M_{t+1}|M_t) = \Pr(M_{t+1})$. Therefore, the numerator of Eq. (2) equals the degree of association observed minus the degree of association expected under independence. If

$$L(prev \rightarrow next) = \frac{P(next|prev) - P(next)}{1 - P(next)} \qquad (1)$$

$$P(next \mid prev) = \frac{Count(prev \rightarrow next)}{Count(prev)} \qquad (2)$$

L Statistics



$$L(prev \rightarrow next) = \frac{P(next|prev) - P(next)}{1 - P(next)} \quad (1)$$

$$P(next \mid prev) = \frac{Count(prev \rightarrow next)}{Count(prev)} \quad (2)$$

Example L Calculation

ABBCAACCBA

(all transitions are equally likely)

$$P(next) = P(B\_next) = \frac{3}{9} = 0.33$$

$$P(next \mid prev) = P(B\_next \mid A\_prev) = \frac{1}{3} = 0.33$$

$$L(A \rightarrow B) = \frac{0.33 - 0.33}{1 - 0.33} = 0$$

Example L Calculation

L statistics calculation for an example sequence of *ABBCAACCBA* when self-transitions are included

Transition	Count	*P(next|prev)*	*P(next)*	*L*
A->A	1	0.33	0.33	0
A->B	1	0.33	0.33	0
A->C	1	0.33	0.33	0
B->A	1	0.33	0.33	0
B->B	1	0.33	0.33	0
B->C	1	0.33	0.33	0
C->A	1	0.33	0.33	0
C->B	1	0.33	0.33	0
C->C	1	0.33	0.33	0

What happens when self-transitions are removed?

If self-transitions are excluded, the sequence
ABBCAACCBA becomes **ABCACBA**.
(all transitions are equally likely)

$$P(next) = P(B\_next) = \frac{2}{6} = 0.33$$

$$P(next \mid prev) = P(B\_next \mid A\_prev) = \frac{1}{2} = \boxed{0.5}$$

$$L(A \rightarrow B) = \frac{0.5 - 0.33}{1 - 0.33} = \boxed{0.25}$$

What happens when self-transitions are removed?

L statistics calculation for an example sequence of *ABBCAACCBA* when self-transitions are excluded

Transition	Count	*P(next|prev)*	*P(next)*	*L*
A->B	1	0.5	0.33	0.25
A->C	1	0.5	0.33	0.25
B->A	1	0.5	0.33	0.25
B->C	1	0.5	0.33	0.25
C->A	1	0.5	0.33	0.25
C->B	1	0.5	0.33	0.25

Inconsistency in Results

From State	To State	D'Mello's L	p-value
Engaged Concentration	Engaged Concentration	—	—
	Boredom	0.260*	<0.001
	Confusion	0.004	0.136
	Frustration	-0.12*	0.012
	Neutral/Other	0.481*	<0.001
Boredom	Engaged Concentration	0.194*	<0.001
	Boredom	—	—
	Confusion	-0.004	0.208
	Frustration	0.036*	<0.001
	Neutral/Other	0.235*	<0.001
Confusion	Engaged Concentration	0.341*	0.006
	Boredom	-0.127*	<0.001
	Confusion	—	—
	Frustration	-0.026*	0.001
	Neutral/Other	-0.156	0.157
Frustration	Engaged Concentration	0.279*	<0.001
	Boredom	-0.107*	<0.001
	Confusion	0.008	0.391
	Frustration	—	—
	Neutral/Other	0.279*	<0.001
Neutral/Other	Engaged Concentration	0.753*	<0.001
	Boredom	-0.057*	<0.001
	Confusion	0.003	0.302
	Frustration	0.015*	0.007
	Neutral/Other	—	—

Botelho, A. F., Baker, R. S., Ocumpaugh, J., & Heffernan, N. T. (2018). Studying Affect Dynamics and Chronometry Using Sensor-Free Detectors. EDM.

Redefining L Value at Chance

$$L(prev \rightarrow next) = \frac{P(next|prev) - P(next)}{1 - P(next)} \quad (1)$$

$$P(next \mid prev) = \frac{Count(prev \rightarrow next)}{Count(prev)} \quad (2)$$

When Self-transitions **Included**

$$P(next) = \frac{n}{n^2} = \frac{1}{n}$$

$$P(next \mid prev) = \frac{1}{n}$$

$$L = 0$$

When Self-transitions **Excluded**

$$P(next) = \frac{n-1}{n^2 - n} = \frac{1}{n}$$

$$P(next \mid prev) = \frac{1}{n-1}$$

$$\boxed{L = \frac{1}{(n-1)^2}}$$

Karumbaiah, S., Baker, R.S., Ocumpaugh, J. (2019) The Case of Self-Transitions in Affective Dynamics. [AIED19]

Redefining L Value at Chance

When Self-transitions **Included**

$$P(next) = \frac{n}{n^2} = \frac{1}{n}$$

$$P(next \mid prev) = \frac{1}{n}$$

$$L = 0$$

When Self-transitions **Excluded**

$$P(next) = \frac{n-1}{n^2 - n} = \frac{1}{n}$$

$$P(next \mid prev) = \frac{1}{n-1}$$

$$\boxed{L = \frac{1}{(n-1)^2}}$$

n	3	4	5	6	7	8
chance L	0.25	0.11	0.0625	0.04	0.0277	0.0204

Karumbaiah, S., Baker, R.S., Ocumpaugh, J. (2019) The Case of Self-Transitions in Affective Dynamics. [AIED19]

L*

Definition 2. Let A and B be two affective states, and let

$$T_{\overline{A}} = \{\text{transitions where } next \neq A\}. \tag{2.1}$$

That is, $T_{\overline{A}}$ consists of all transitions where the next affective state is not equal to A. We can then define

$$L^*(A_{prev} \to B_{next}) := \frac{P(B_{next} \mid A_{prev}, T_{\overline{A}}) - P(B_{next} \mid T_{\overline{A}})}{1 - P(B_{next} \mid T_{\overline{A}})}. \tag{2.2}$$

Matayoshi, J., & **Karumbaiah, S.** (2020) Adjusting the L Statistic when Self-Transitions are Excluded in Affective Dynamics. [JEDM20]
Bosch, N., & Paquette, L. (2021). What's Next? Sequence Length and Impossible Loops in State Transition Measurement. Journal of Educational Data Mining.

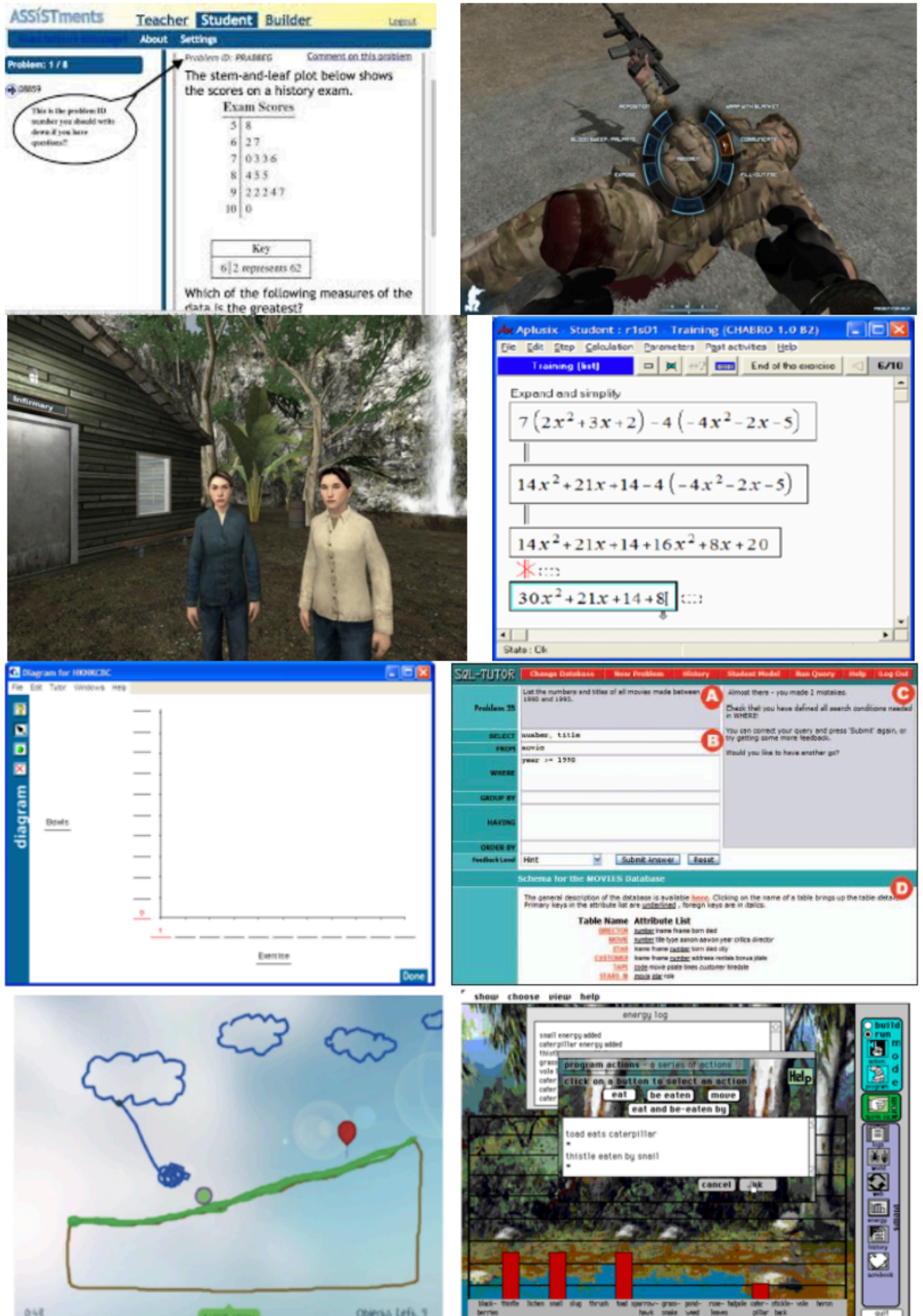
Implications to the Theory Conformance

Bosch & D'Mello, 2013	US	--	29	Undergrads	Unnamed	L	RJP on 100 fixed points	25min	Exc	3
Bosch, & D'Mello, 2017	Midwestern US	17-21	99	Undergrads	Unnamed	L	RJP on 100 fixed points	25min	Exc	5
D'Mello & Graesser, 2012	Southern US	--	28; 30	Undergrads	Auto-Tutor	L	RJP every 20s; fixed points	32min; 35min	Exc	4;5
D'Mello et al., 2007	Southern US	--	28	Undergrads	Auto-Tutor	L	RJP ev. 20s	32min	Inc	2
D'Mello et al., 2009	Southern US	--	41	Undergrads	Unnamed	L	RJP on fixed points	35min	Exc	1
D'Mello & Graesser, 2010	Southern US	--	28; 30	Undergrads	Auto-Tutor	L	RJP ev. 20s; fixed points	32min; 35min	Exc	3;3

- They are likely to have overstated their possible effects, possibly finding positive results where negative results would have been more accurate.
- Results need to be reanalyzed with the appropriate base rate values for L.

Re-Analysis and Synthesis

- Acquired 10 past affect datasets (2 with experiment and control conditions)
- 8 different virtual learning environments, 2 traditional classrooms without a virtual learning system
- 9 in classroom environment (using BROMP), 1 in lab setting (self-report)
- 5 in the US, 5 in the Philippines
- Varied student demographics (age, gender, urbanicity, SES), school type (public/private), subject matter



Baker, R.S., Ocumpaugh, J.L., Andres, J.M.A.L. (in press) BROMP Quantitative Field Observations: A Review. In R. Feldman (Ed.) Learning Science: *Theory, Research, and Practice*.

Re-Analysis and Synthesis

- Redefined L statistic

Re-Analysis and Synthesis

- Redefined L statistic
- Standardized treatment of edge cases

The cases below illustrate situations where transition calculations may not be straight forward:

1. L is 0 for any transition going into a state that did not occur in a student's affect sequence. In that case, $P(next) = 0$ and $P(next \mid prev) = 0$, and thus, $L = 0$.
2. The L value is undefined for any transition out of a state that does not occur for a student, as we do not know what would have followed that state if it had occurred.
3. When a student remains in one affective state throughout an observation period, all transitions to states other than that state are 0, and all transitions to the single affective state seen have undefined L, as the denominator of the equation is 0 in that case.
4. When self-transitions are discarded from the data, an affect sequence consisting of a single state is reduced to a single state. In this case, since there would be no affective state in the next value, L is undefined for all states.

In all cases where L is undefined, those values are discarded from further analysis.

Karumbaiah, S. et al. (2018) The Implications of a Subtle Difference in the Calculation of Affect Dynamics. [ICCE18]

Re-Analysis and Synthesis

- Redefined L statistic
- Standardized treatment of edge cases
- Self-transitions excluded
  - reveals a larger number of affective patterns that might otherwise be suppressed by persistent affective states

Re-Analysis and Synthesis

- Redefined L statistic
- Standardized treatment of edge cases
- Self-transitions excluded
- Stouffer's Z to summarize significance levels from multiple affect datasets

$$\sum_{i=1}^{k} z(p_i)/\sqrt{k}$$

Across All Datasets

STOUFFER'S Z AND COMBINED P-VALUES FOR THE TWELVE NON-SELF-TRANSITIONS STUDIED IN THIS PAPER.

Transition	Stouffer's Z	Combined p
ENG_CON	**6.770**	**1.28e-11**
ENG_FRU	-10.878	1.46e-27
ENG_BOR	-12.296	9.40e-35
CON_ENG	-1.605	0.108
CON_FRU	-4.863	1.15e-06
CON_BOR	-7.763	8.25e-15
FRU_ENG	-3.906	9.35e-05
FRU_CON	-2.075	0.037
FRU_BOR	-0.007	0.99
BOR_ENG	-1.344	0.178
BOR_CON	-8.885	6.37e-19
BOR_FRU	-3.861	1.12e-04

Karumbaiah, S. et al. (2021) A Re-Analysis and Synthesis of Data on Affect Dynamics in Learning. [IEEE TAC21]

Between Two Countries

United States

Transition	Stouffer's Z	Combined p
ENG_CON	**18.337**	**4.14e-75**
ENG_FRU	**8.114**	**4.90e-16**
ENG_BOR	-0.511	0.609
CON_ENG	**2.028**	**0.042**
CON_FRU	-2.581	0.009
CON_BOR	-2.183	0.029
FRU_ENG	-1.768	0.077
FRU_CON	0.752	0.452
FRU_BOR	-0.905	0.365
BOR_ENG	**2.110**	**0.034**
BOR_CON	-5.150	2.60e-07
BOR_FRU	0.264	0.791

Philippines

Transition	Stouffer's Z	Combined p
ENG_CON	-6.634	3.26e-11
ENG_FRU	-21.100	7.88e-99
ENG_BOR	-15.668	2.46e-55
CON_ENG	-3.816	1.35e-04
CON_FRU	-4.144	3.40e-05
CON_BOR	-8.319	8.80e-17
FRU_ENG	-3.561	3.69e-04
FRU_CON	-2.973	0.0029
FRU_BOR	0.674	0.499
BOR_ENG	-3.543	3.94e-04
BOR_CON	-7.281	3.32e-13
BOR_FRU	-5.279	1.29e-07

Karumbaiah, S. et al. (2021) A Re-Analysis and Synthesis of Data on Affect Dynamics in Learning. [IEEE TAC21]

Conclusion

- Non-conformance to the theoretical model
  - Our synthesis supports only one (*engaged concentration -> confusion*) of the six hypothesized transitions
  - The widely-accepted theoretical model is either invalid or has a more limited scope than what it is currently being used for in the community
  - It is highly unlikely that there is a general multi-step pattern in affect dynamics
  - There may still be some contextually relevant patterns useful to understand student experience

Conclusion

- Non-conformance to the theoretical model
- Methodological implications for future affect dynamics research
  - If a study excludes self-transitions, the test for significance and the interpretation of the result must choose the appropriate chance level of L statistic and not zero.

Continued Methodological Improvements in Transition Analysis

Yang, J. C. et al. (Eds.) (2018). Proceedings of the 26th International Conference on Computers in Education. Philippines: Asia-Pacific Society for Computers in Education

The Implications of a Subtle Difference in the Calculation of Affect Dynamics

Shamya KARUMBAIAH[a], Juliana Ma. Alexandra L. ANDRES[a], Anthony F. BOTELHO[b], Ryan S. BAKER[a], Jaclyn OCUMPAUGH[a]

The Case of Self-Transitions in Affective Dynamics

Shamya Karumbaiah, Ryan S Bakerand Jaclyn Ocumpaugh

University of Pennsylvania, USA
shamya@upenn.edu, ryanshaunbaker@gmail.com,
jlocumpaugh@gmail.com

Studying Affect Dynamics using Epistemic Networks

Shamya Karumbaiah and Ryan S Baker

University of Pennsylvania, USA

l.com

What's Next? Sequence Length and Impossible Loops in State Transition Measurement

Nigel Bosch
University of Illinois
Urbana–Champaign
pnb@illinois.edu

Luc Paquette
University of Illinois
Urbana–Champaign
lpaq@illinois.edu

Adjusting the L Statistic when Self-Transitions are Excluded in Affect Dynamics

Jeffrey Matayoshi
McGraw Hill ALEKS
jeffrey.matayoshi@aleks.com

Shamya Karumbaiah
University of Pennsylvania
shamya@upenn.edu

Affect dynamics, the investigation of how student affect transitions from one state to another, is a popular area of research in adaptive learning environments. Recently, the commonly used transition metric L has come under critical examination when applied to data that exclude self-transitions (i.e., transitions where

Using Marginal Models to Adjust for Statistical Bias in the Analysis of State Transitions

Jeffrey Matayoshi
McGraw Hill ALEKS
Irvine, California, USA
jeffrey.matayoshi@aleks.com

Shamya Karumbaiah
University of Pennsylvania
Philadelphia, Pennsylvania, USA
shamya@upenn.edu

TRACT

reas of educational research require the analysis of data that
n inherent sequential or temporal ordering. In certain cases,
hers are specifically interested in the transitions between
t states—or events—in these sequences, with the goal being

1 INTRODUCTION

As learning is a process that occurs over time, many areas of education and learning analytics research require the analysis of data that have a sequential or temporal ordering. Such analyses are important, as our understanding of the learning process can be

Conclusion

- Non-conformance to the theoretical model
- Methodological implications for future affect dynamics research
  - If a study excludes self-transitions, the test for significance and the interpretation of the result must choose the appropriate chance level of L statistic and not zero.
  - Continued methodological improvements to mitigate issues with low base rates and short sequences
  - <u>Open Questions</u>:
    - Field observations tend to sample at slower rates
    - Field observations are also coarser grained as compared to automated detection
    - How do these methodological choices impact the validity or applicability of the theoretical model?

Conclusion

- Non-conformance to the theoretical model
- Methodological implications for future affect dynamics research
- Need to focus on cultural factors in affect dynamics research
  - Given the differences in national culture, school culture, use of educational technology and forms of disengagement in the 2 countries - it is difficult at this point to understand *why* we see these differences
  - Best sense to re-consider affect dynamics as a generalizable phenomenon after it has been studied in a wider range of cultures

Why context matters in affect studies?

- Student demographics influence affect
  - **Culture** influences emotional expression and regulation (Tsai & Levenson, 1997; Uchida et al., 2009)
  - **Culture** influences frequency and emergence of affect (Kitayama et al., 2000)
  - **Age** influences emotional expressivity (Dunn & Brown, 1994; Gross et al., 1997) and inhibition (Cole, 1986)

Tsai, J., Levenson, R. (1997) Cultural influences on emotional responding: Chinese Am. & European Am. dating couples during interpersonal conflict. J. Cross-Cultural Psych.
Uchida, Y., et al. (2009). Emotions as within or between people? Cultural variation in lay theories of emotion expression and inference. Personality and Social Psychology Bulletin.
Kitayama, S., et al. (2000). Culture, emotion, and well-being: Good feelings in Japan and the United States. Cognition & Emotion.

Conclusion

- Non-conformance to the theoretical model
- Methodological implications for future affect dynamics research
- Need to focus on cultural factors in affect dynamics research
- Yet to study student affective experiences in collaborative setting

My dissertation - Upstream Sources of Bias

Focus on Data Collection Method, System Design, and Theory

No matter what P looks like…

…and for any reasonable H…

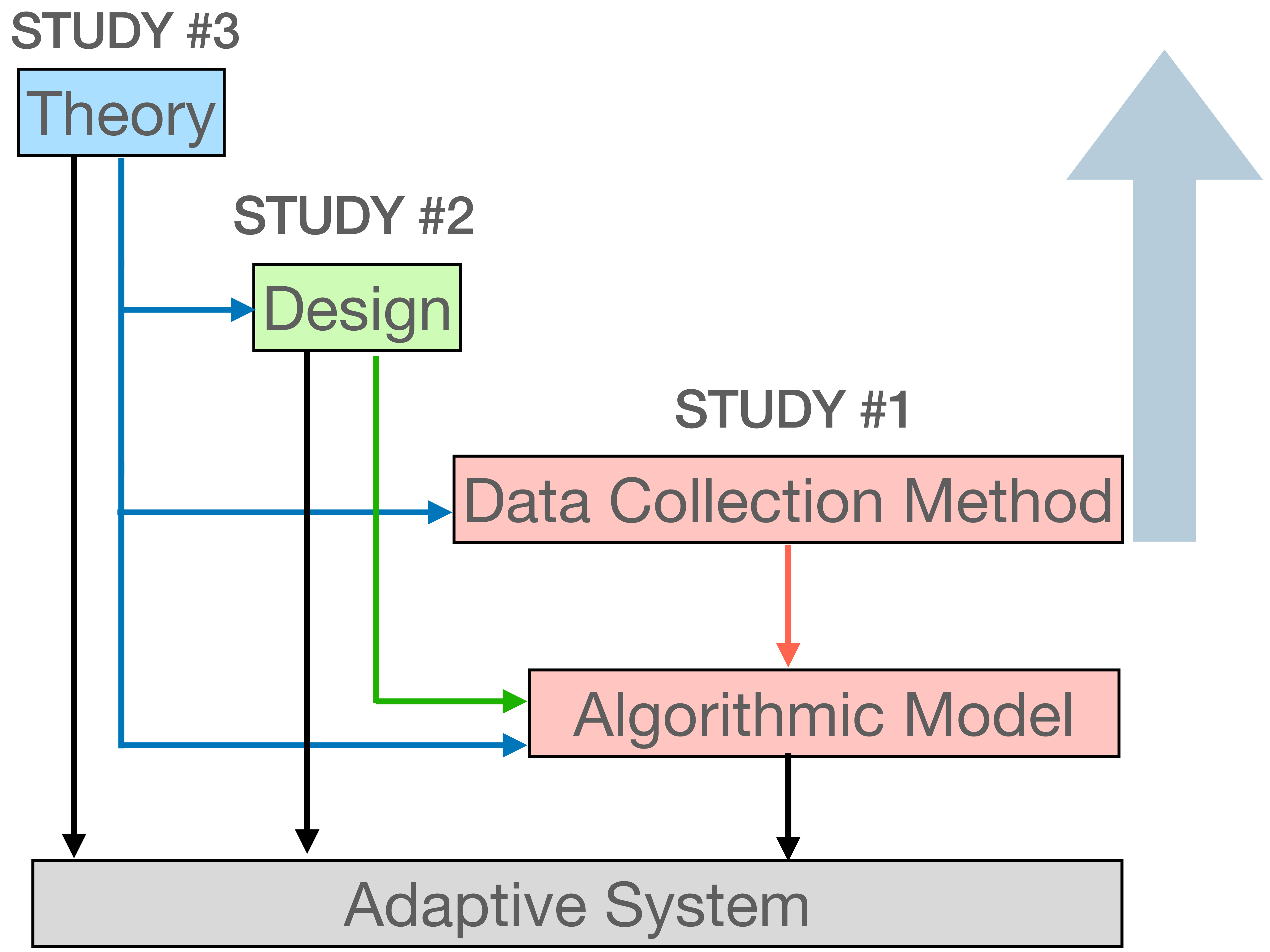
…if we have **enough data** S…

…then for **every h ∈ H**, we have

minimizing error on **data**

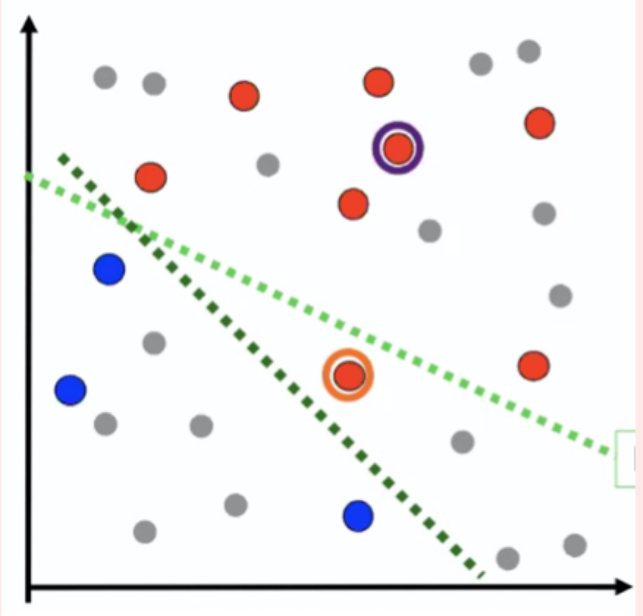
≈

minimizing **true/future** error



My dissertation - Upstream Sources of Bias

STUDY #1
ANNOTATED DATA COLLECTION



- Active machine learning to improve annotated data collection and cold start problem
- Varying effectiveness of methodological improvements

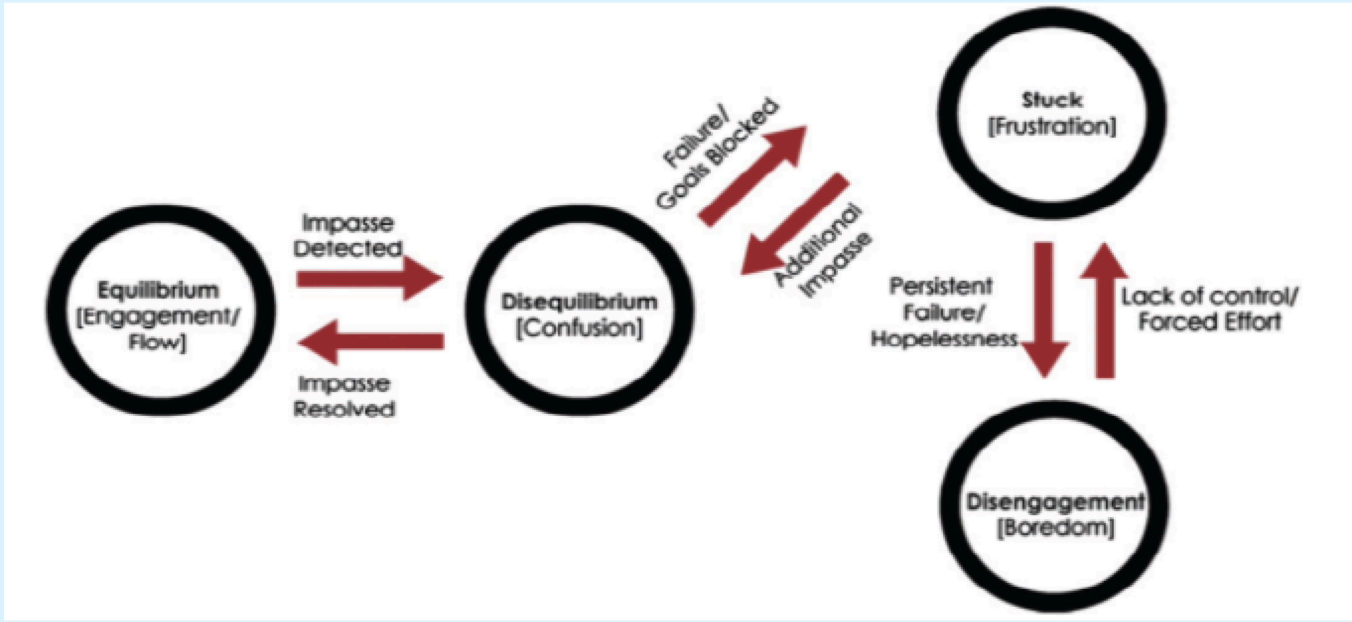
Karumbaiah, S. et al. (2021)
Using Past Data to Warm Start Active Machine Learning: Does Context Matter? [ACM LAK21]

STUDY #2
EDTECH DESIGN

- Differing implications of technology design on student outcomes
- Use of publicly-available, school-level demographics for bias research

Karumbaiah, S. et al. (2021)
Context Matters: Differing Implications of Motivation and Help-Seeking in Educational Technology. [IJAIED21]

STUDY #3
THEORETICAL MODEL



- Generalizability of a widely accepted theory
- Country-level differences in the non-conformance of empirical data

Karumbaiah, S. et al. (2021)
A Re-Analysis and Synthesis of Data on Affect Dynamics in Learning.
[IEEE TAC21]

The Search for Context

Bias in Adaptive Learning Systems

Upstream Sources of Bias

Contextualizing Theoretical Model of Affect

Contextualizing Origins of Bias

Continued Search For Context

Representation Bias

- Most research studies tend to be conducted in western countries with adaptive systems developed by designers in the west (Blanchard, 2012)
    - Little to no research on the generalizability when used in non-western contexts despite evidence on differences in student use (Ogan et al., 2012)

Ogan, A. et al. (2012). Collaboration in cognitive tutor use in Latin America: Field study and design recommendations.

Representation Bias

- Most research studies tend to be conducted in western countries with adaptive systems developed by designers in the west (Blanchard, 2012)
- Small-scale experiments also tend to recruit from a convenience sample due to practical constraints of research projects
  - For example, undergraduate, middle-class students in the United States (Kimble, 1987)
  - These participants are likely to exhibit significant differences in their behavior than other subpopulations (Henrich et al., 2010)

Henrich, J. et al. (2010). Most people are not WEIRD. Nature.

Kimble, G. A. (1987). The scientific value of undergraduate research participation. American Psychologist.

Representation Bias

- Most research studies tend to be conducted in western countries with adaptive systems developed by designers in the west (Blanchard, 2012)
- Small-scale experiments also tend to recruit from a convenience sample due to practical constraints of research projects
- Even when there is access to larger, more diverse datasets, it is often harder to collect student demographics data due to concerns over student privacy

Representation Bias

- Most research studies tend to be conducted in western countries with adaptive systems developed by designers in the west (Blanchard, 2012)
- Small-scale experiments also tend to recruit from a convenience sample due to practical constraints of research projects
- Even when there is access to larger, more diverse datasets, it is often harder to collect student demographics data due to concerns over student privacy
- Invalidates assumptions for student subpopulations not represented in the experiments informing upstream components

Measurement Bias

- Due to issues in data collection methods
- Reliability of measurements across different student subpopulations
  - For example, coder bias due to cross-cultural affect coding
  - Poor performance of automated facial recognition for female students and those with darker skin tones (Lohr, 2018)
  - Demographic differences in the reliability of self-reports (e.g., age, culture)

Lohr, S. (2018). Facial recognition is accurate, if you're a white guy. New York Times.

Historical Bias

- Forming theories and design choices by observing the world as it exists (including its biases)
  - For example, earlier cameras were designed to bring out high contrast and better resolution for white skin color (Roth, 2009)
- Historical biases embedded in upstream sources get perpetuated by downstream applications
  - For example, a predictive model that then automates the biased decision-making at a potentially larger scale

Roth, L. (2009). Looking at Shirley, the Ultimate Norm: Colour Balance, Image Technologies, and Cognitive Equity. Canadian Journal of Communication.

The Search for Context

Bias in Adaptive Learning Systems

Upstream Sources of Bias

Contextualizing Theoretical Model of Affect

Contextualizing Origins of Bias

Continued Search For Context

The Search for Context

Learning Sciences, Learning Analytics
2017-2021

Machine Learning, Learning Analytics
2015-2017

Computer Science, Software Engineering
2007-2015

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

"Although the **learning sciences** is continually evolving, what remains true of the tenets of this educational field is that learning happens through mediated processes that most often require collaboration with others whereby learning is inextricably linked to **context** and culture" - Dr. Yoon, EDUC 545

"**Learning analytics** is the measurement, collection, analysis and reporting of data about learners and their **contexts**, for purposes of understanding and optimizing learning and the environments in which it occurs." - SOLAR

Cited Articles

Karumbaiah, S., Baker, R.S., Ocumpaugh, J. (2021) Context Matters: Differing Implications of Motivation and Help-Seeking in Educational Technology. *International Journal of Artificial Intelligence in Education. [IJAIED21]*

Karumbaiah, **S.,** Lan, A., Nagpal, S., Baker, R.S., Botelho, A., Heffernan, N. (2021) Using Past Data to Warm Start Active Machine Learning: Does Context Matter? *International Learning Analytics and Knowledge Conference. [LAK21]* [Nominated for Paper Award]

Karumbaiah, S., Baker, R.S., Ocumpaugh, J., Andres, J.M.A.L. (2021) A Re-Analysis and Synthesis of Data on Affect Dynamics in Learning. *IEEE Transactions on Affective Computing. [IEEE TAC21]*

Matayoshi, J., **Karumbaiah, S.** (2021) Investigating the Validity of Methods Used to Adjust for Multiple Comparisons in Educational Data Mining. *International Conference on Educational Data Mining. [EDM21]*

Karumbaiah, S., Brooks, J. (2021) How Colonial Continuities Underlie Algorithmic Injustices in Education. *IEEE Research in Equity and Sustained Participation in Engineering, Computing, and Technology. [IEEE RESPECT21]*

Matayoshi, J., **Karumbaiah, S.** (2021) Using Marginal Models to Adjust for Statistical Bias in the Analysis of State Transitions. *International Learning Analytics and Knowledge Conference. [LAK21]*

Karumbaiah, S., Baker, R.S. (2020) Studying Affect Dynamics using Epistemic Networks. *International Conference on Quantitative Ethnography*. *[ICQE20]* [Nominated for Paper Award]

Crossley, S.A., **Karumbaiah, S.**, Ocumpaugh, J., Labrum, M., Baker, R.S. (2020) Predicting Math Identity through Language and Click-stream Patterns in a Blended Learning Mathematics Program for Elementary Students. *Journal of Learning Analytics. [JLA20]*

Matayoshi, J., **Karumbaiah, S.** (2020) Adjusting the L Statistic when Self-Transitions are Excluded in Affective Dynamics. *Journal of Educational Data Mining. [JEDM20]*

Karumbaiah, S., Baker, R.S., Barany, A., Shute, V. (2019) Using Epistemic Networks with Automated Codes to Understand Why Players Quit Levels in a Learning Game. *International Conference on Quantitative Ethnography. [ICQE19]*

Karumbaiah, S., Baker, R.S., Ocumpaugh, J. (2019) The Case of Self-Transitions in Affective Dynamics. *International Conference on Artificial Intelligence in Education*. *[AIED19]*

Karumbaiah, S., Ocumpaugh, J., Baker, R.S. (2019) The Influence of School Demographics on the Relationship Between Students' Help-Seeking Behavior and Performance and Motivational Measures. *International Conference on Educational Data Mining. [EDM19]*

Crossley, S.A., **Karumbaiah, S.**, Ocumpaugh, J., Labrum, M., Baker, R.S. (2019) Predicting Math Success in an Online Tutoring System Using Language Data and Click-stream Variables: A longitudinal analysis. *Conference on Language, Data and Knowledge. [LDK19]*

Karumbaiah, S., Baker, R.S., Shute, V. (2018) Predicting Quitting in Students Playing a Learning Game. *Educational Data Mining. [EDM18]* [Nominated for Paper Award]

Nye, B. D., **Karumbaiah, S.,** Tokel, S. T., Core, M. G., Stratou, G., Auerbach, D., & Georgila, K. (2018) Engaging with the Scenario: Affect and Facial Patterns from a Scenario-Based Intelligent Tutoring System. *International Conference on Artificial Intelligence in Education. [AIED18]*

Karumbaiah, S., Andres, J.M.A.L., Botelho, A.F., Baker, R.S., Ocumpaugh, J. (2018) The Implications of a Subtle Difference in the Calculation of Affect Dynamics. *International Conference on Computers in Education*. *[ICCE18]* [Nominated for Paper Award]

Karumbaiah, S., Tao, Y., Baker, R.S., Ziyang, L. (under review) How does Students' Affect in Virtual Learning Relate to Their Outcomes? A Systematic Review. *[UR21]*

Karumbaiah, S., Syam, A., Baker, R.S., Shute, V. (under preparation) Understanding Student Behaviors in a Learning Game by Developing Qualitative Explanations of an Algorithmic Model. *[UP21]*

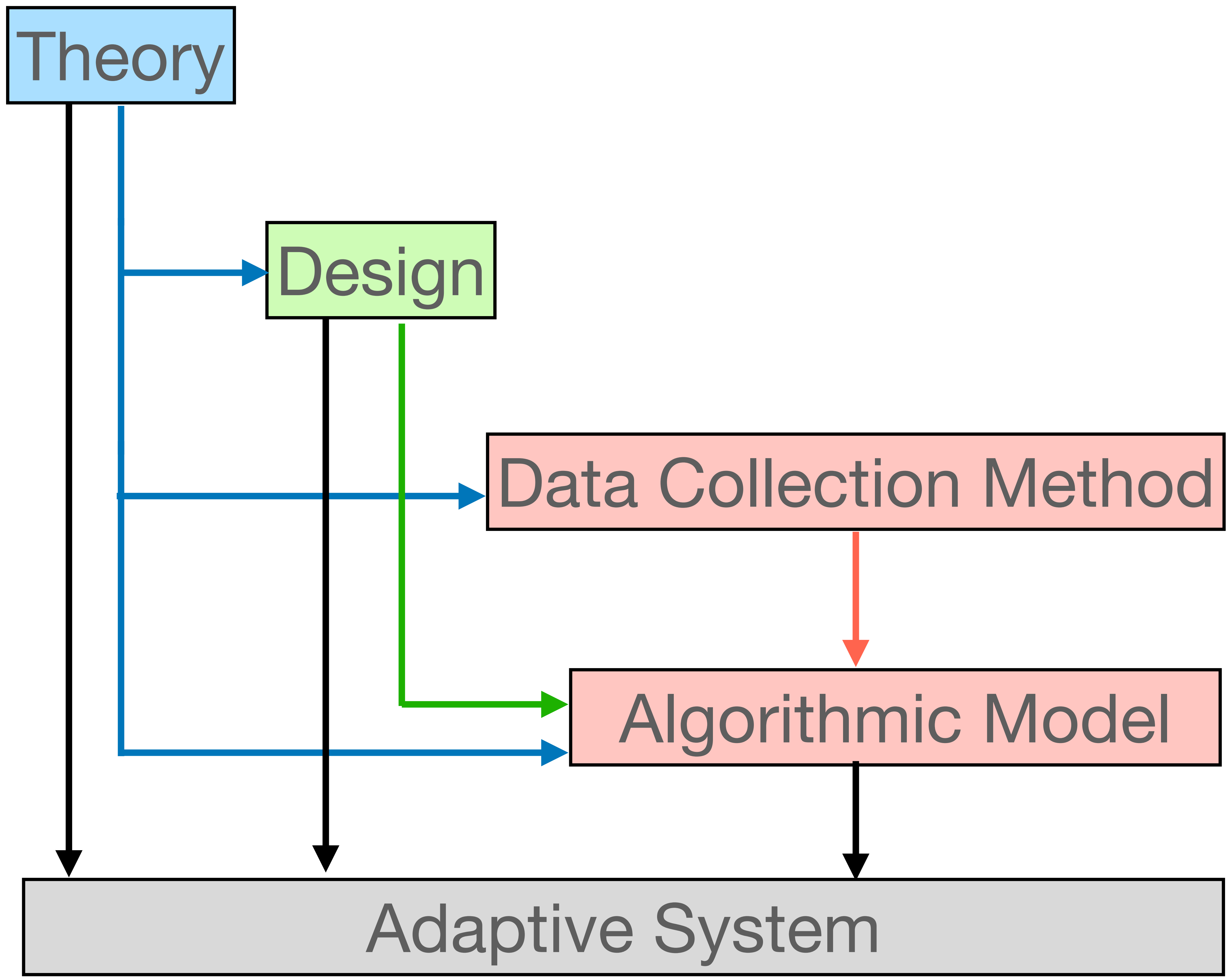
Understanding, Identifying and Mitigating Bias

Pushing Learning Analytics Forward Around Questions of Bias

- Future lines of inquiry
    1. Contextualizing sources, origins, and harms of bias: Theory Building
    2. Technical approaches to mitigating bias: Methodological Improvements
    3. Human-centered approaches to addressing bias: Social Change

#1 Contextualizing Sources, Origins & Harms of Bias

Theory Building for Research, Design, & Practice



- “what kinds of system behaviors are harmful, in what ways, to whom, and why?”

Blodgett, S. L., Barocas, S., III, H. D., & Wallach, H. (2020). Language (Technology) is Power: A Critical Survey of “Bias” in NLP. ACL.
Suresh, H., & Guttag, J. V. (2020). A framework for understanding unintended consequences of machine learning.

#1 Contextualizing Sources, Origins & Harms of Bias

Theory Building for Research, Design, & Practice

- **Sources**: data, model, theory
- **Origins**: measurement, representation, historical
- **Impacted Populations**: female students, rural school students, indigenous students
- **Harms**: missed learning opportunity, reduced interest in subject learned

- "what kinds of system behaviors are harmful, in what ways, to whom, and why?"
- Theoretical grounding (ed affective computing)
- Audits of real-world systems
- Shared language & nuanced understanding of bias

Blodgett, S. L., Barocas, S., III, H. D., & Wallach, H. (2020). Language (Technology) is Power: A Critical Survey of "Bias" in NLP. ACL.
Suresh, H., & Guttag, J. V. (2020). A framework for understanding unintended consequences of machine learning.

#2 Technical Approaches to Mitigating Bias

Methodological Improvements for Context Aware LA

No matter what P looks like…

…and for any reasonable H…

…if we have **enough data** S…

…then for **every h ∈ H**, we have

minimizing error on **data**

≈

minimizing **true/future** error

- Goal: Situate fair modeling in real-world educational contexts
- How to identify different student subgroups?
- Intersectionality matters
- Establish trade-offs and impossibilities

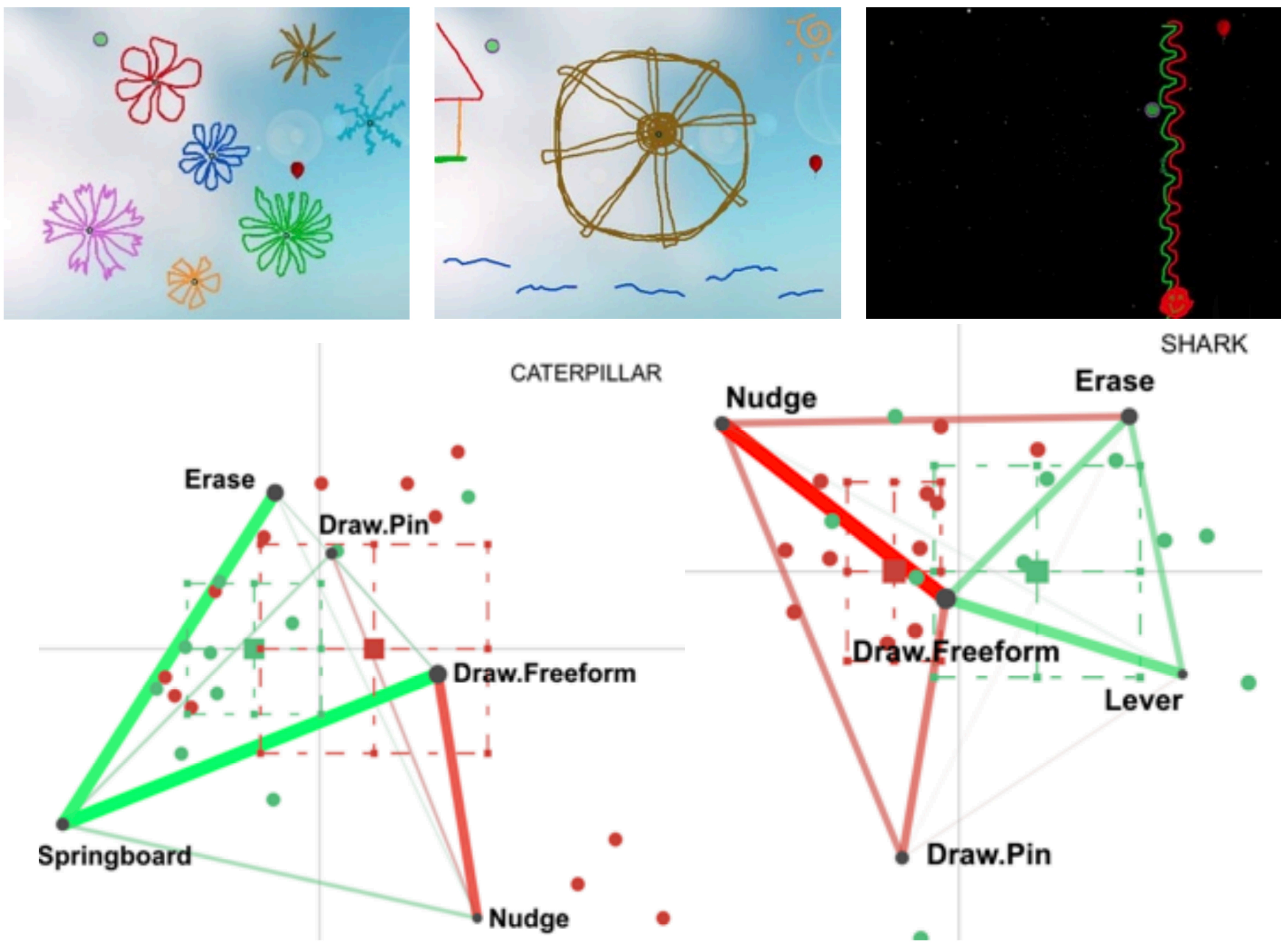
Baker, R. S., & Hawn, A. (2021). Algorithmic Bias in Education.

Kleinberg, J., Mullainathan, S., & Raghavan, M. (2016). Inherent trade-offs in the fair determination of risk scores.

Karumbaiah, S., et al. (2019) The Influence of School Demographics on the Relationship Between Students' Help-Seeking Behavior and Performance… [EDM19]

#2 Technical Approaches to Mitigating Bias

Methodological Improvements for Context Aware LA



- Goal: Generate qualitative explanations for biases in automated decisions
- Quantitative Ethnography
- Explainable AI techniques
- The shift from predicting to understanding

Karumbaiah, S. et al. (2019) Using Epistemic Networks with Automated Codes to Understand Why Players Quit Levels in a Learning Game. [ICQE19]

Karumbaiah, S., Syam, A., et al. (under preparation) Understanding Student Behaviors in a Learning Game by Developing Qualitative Explanations of an Algorithmic Model.

#3 Human-Centered Approaches to Addressing Bias

LA as a Tool to Drive Positive Social Change

An Imperative & An Opportunity

- Goal: Active voice of teachers and learners in equitable human-AI adaptivity
- Human-centered LA
- Fairness elicitation
- Collective audits to identify and detect bias

50 Years of Test (Un)fairness: Lessons for Machine Learning

Ben Hutchinson and Margaret Mitchell

We conclude by reflecting on what further lessons the history of test fairness may have for the future of ML fairness. Careful attention should be paid to legal and public concerns about fairness. The experiences of the test fairness field suggest that in the coming years, courts may start ruling on the fairness of ML models. If technical definitions of fairness stray too far from the public's perceptions of fairness, then the political will to use scientific contributions in advance of public policy may be difficult to obtain.

Karumbaiah, S., & Brooks, J. (2021) How Colonial Continuities Underlie Algorithmic Injustices in Education. [IEEE RESPECT21]

Karumbaiah, S., Shim, J., Yoon, S., et al. (under submission) Case Studies on Designing an Online Environment for In-service Teachers' Learning of Computational Thinking.

Understanding, Identifying and Mitigating Bias

Pushing Learning Analytics Forward Around Questions of Bias

- Future lines of inquiry
  1. Technical approaches to mitigating bias: Methodological Improvements
  2. Human-centered approaches to addressing bias: Social Change
  3. Contextualizing sources, origins, and harms of bias: Theory Building

Thank you!

@s_karumbaiah

shamya@cmu.edu

shamya.github.io