STEM Summit

Neil Heffernan

Who is Neil?

Dr. Neil Heffernan

- Professor of Computer Science
- Director, Learning Sciences & Technologies Graduate Program
- Over 40 Peer-Reviewed Data Mining Studies
- About 2 dozen Peer-Reviewed Randomize Controlled Experiments.

Who is Cristina?

Cristina Heffernan
Chief Teacher Trainer,
Co-Founder of ASSISTments
•7 years as Math Teacher
•8 Year as math coach and PD provider

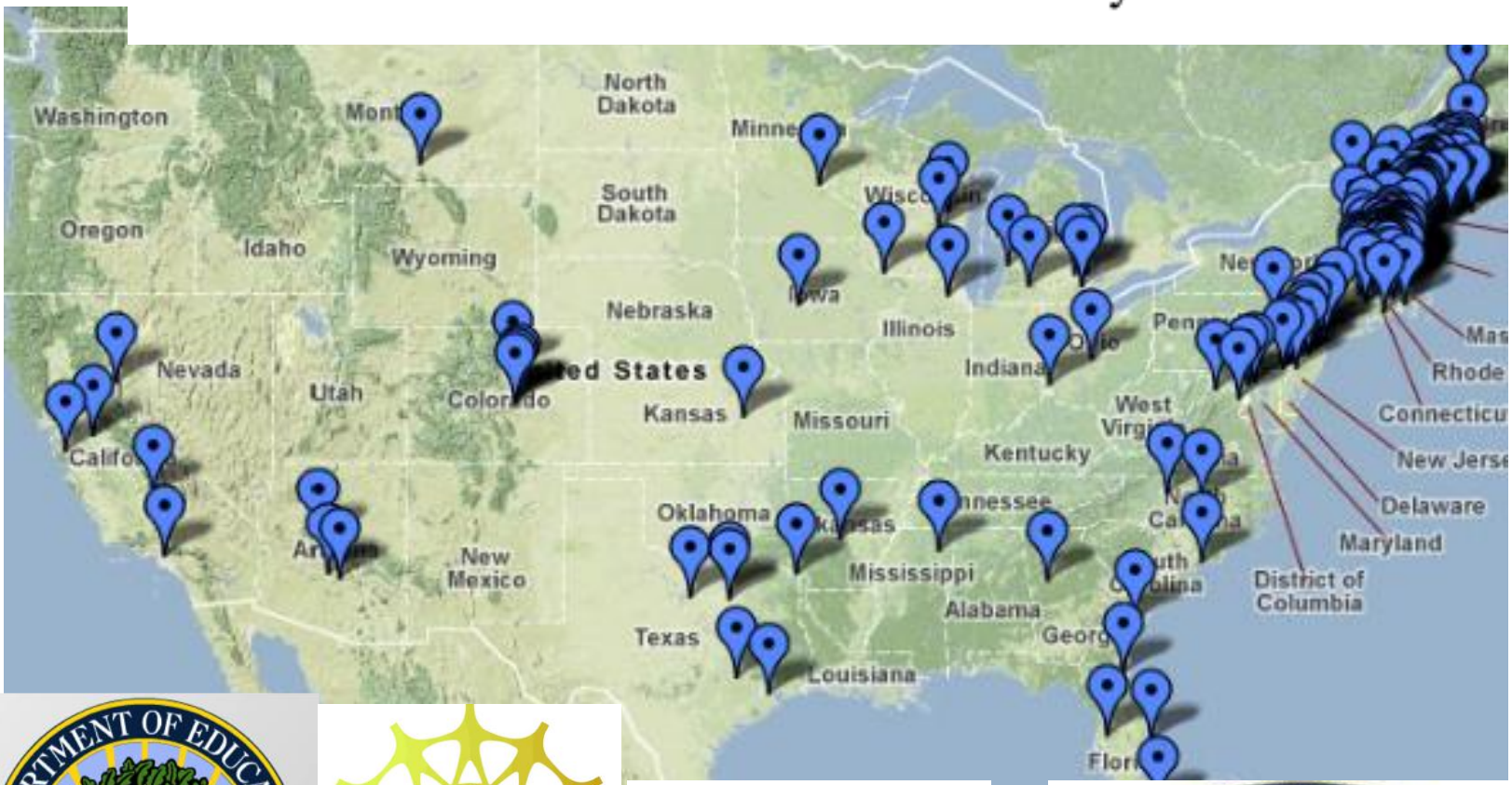
•Leads Professional Development
•Trains Trainers
•Monitors Content
•Runs Studies

My Background-

- Teacher of America Teacher
  - Middle school math and science
  - I implemented something in my classroom by hand
- Now can we help many researchers
- We give away all our data free, with student and teacher identifiers removed.
- We care a great deal about open science.
- Free

ASSISTments™

A Free Public Service of Worcester Polytechnic Institute



DEPARTMENT OF EDUCATION · UNITED STATES OF AMERICA

NSF

BILL & MELINDA GATES foundation

DEPARTMENT OF THE NAVY · ONR · USN · Science & Technology

ASSISTments is an online platform that: (video)

- Gives immediate feedback to teachers, students, school administrators, and parents
- Has flexible content for teachers to use pre-built problem sets, edit pre-built problem sets, or build their own problem sets
- Has useful features- show video

ASSISTments™

A Free Public Service of Worcester Polytechnic Institute

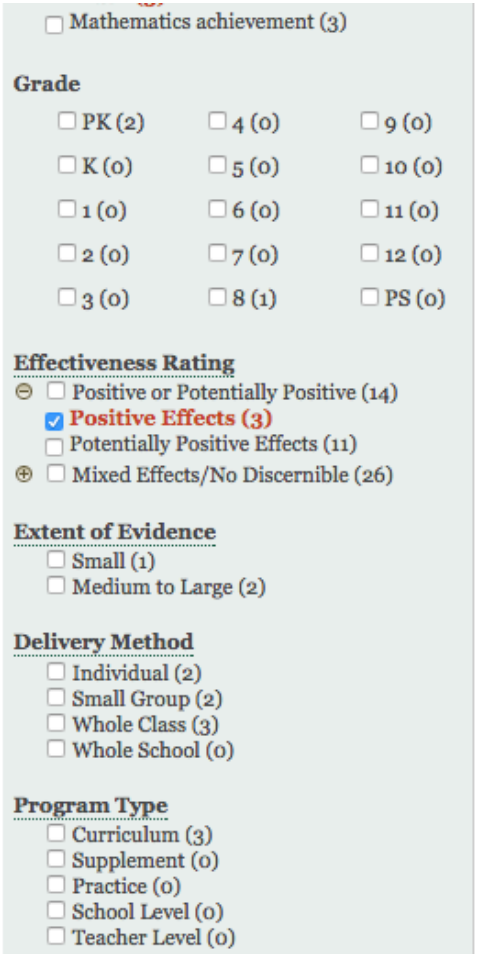
Jeremy Roschelle
@roschelle63

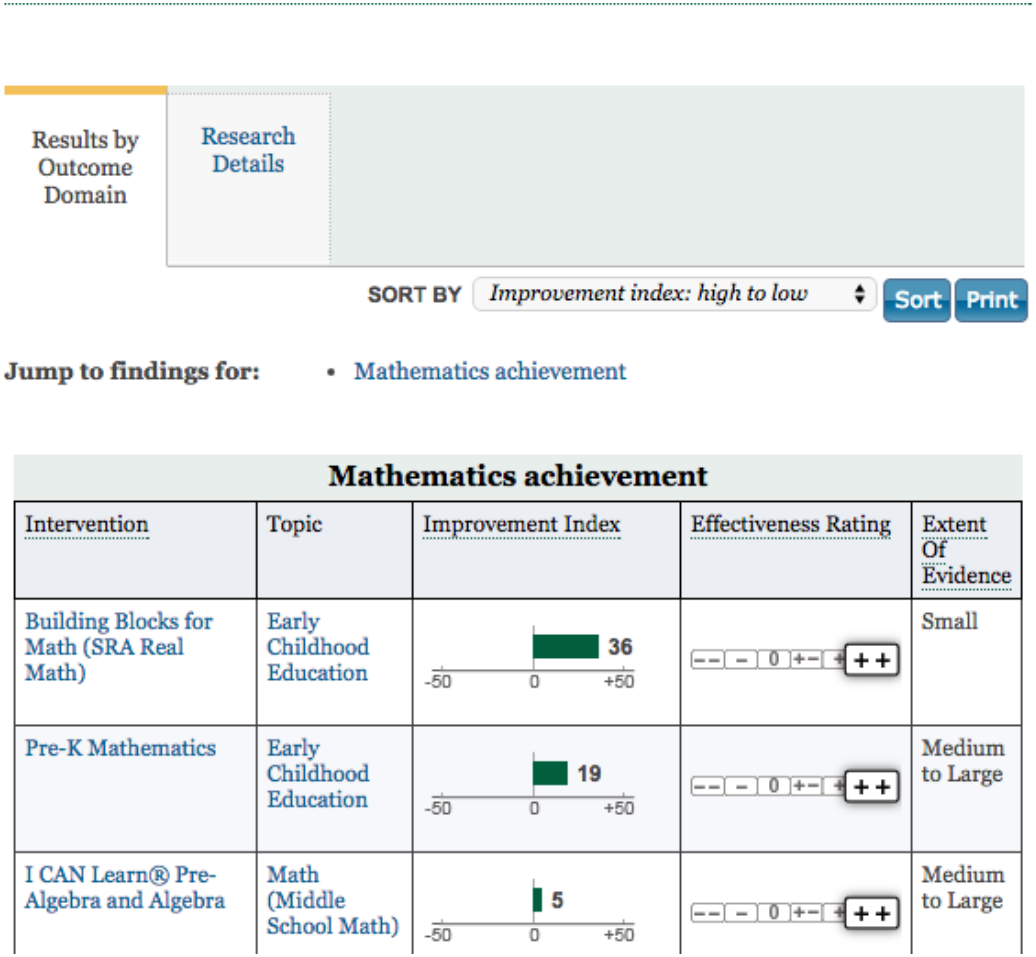
Follow

Can we improve math homework? Presented @SRI_Education @assistments positive findings at the White House Symposium on Digital Learning today


Panel: Internet-Scale Experimentation

What Works Clearing House





Mathematics achievement

Intervention	Topic	Improvement Index	Effectiveness Rating	Extent Of Evidence
Building Blocks for Math (SRA Real Math)	Early Childhood Education	36 (-50 0 +50)	++	Small
Pre-K Mathematics	Early Childhood Education	19 (-50 0 +50)	++	Medium to Large
I CAN Learn® Pre-Algebra and Algebra	Math (Middle School Math)	5 (-50 0 +50)	++	Medium to Large

ASSISTments Homework Efficacy Study

SRI International

Hypothesis

Online Math Homework will improve student learning

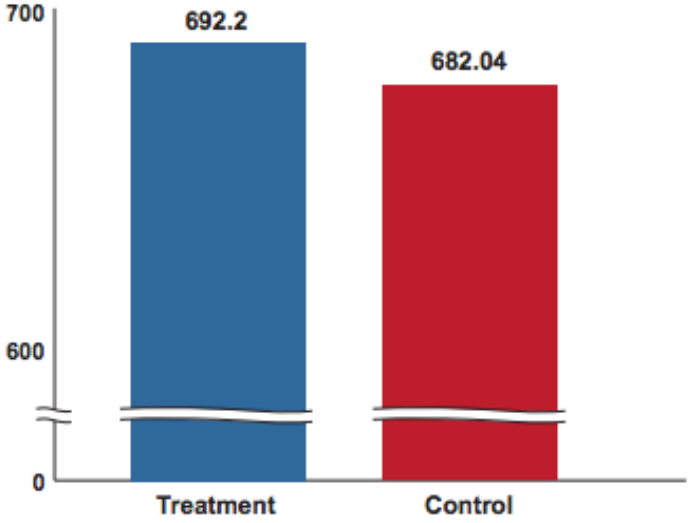
Procedure

- Random Assignment
  - Schools paired by similar size and prior math scores
  - Coin toss for each pair

- Schools stay in group for two years
  - Year 1: Teacher preparation, practice and coaching
  - Year 2: Teachers use with new cohort of students
- Teachers follow school homework policies
- Content comes textbooks, plus skill builders

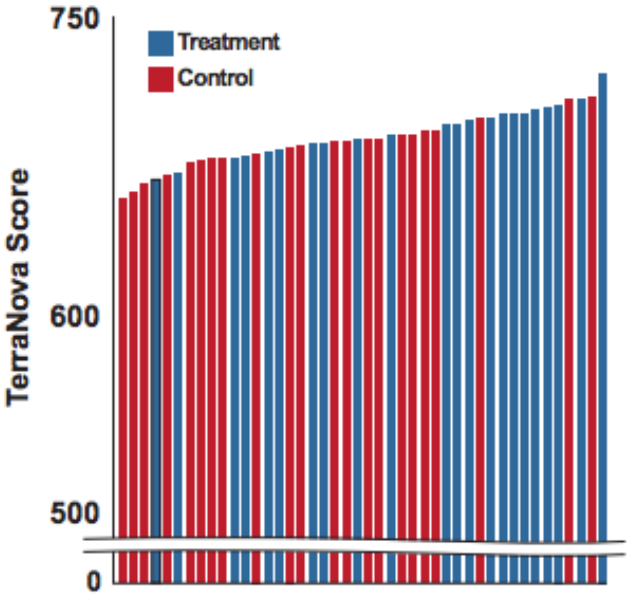
Impact Findings

Students in 7th grade whose teacher used ASSISTments learned more math as measured by a standardized test (TerraNova), controlling for their 6th grade math and reading scores on the prior year state achievement test.



Mean TerraNova Score

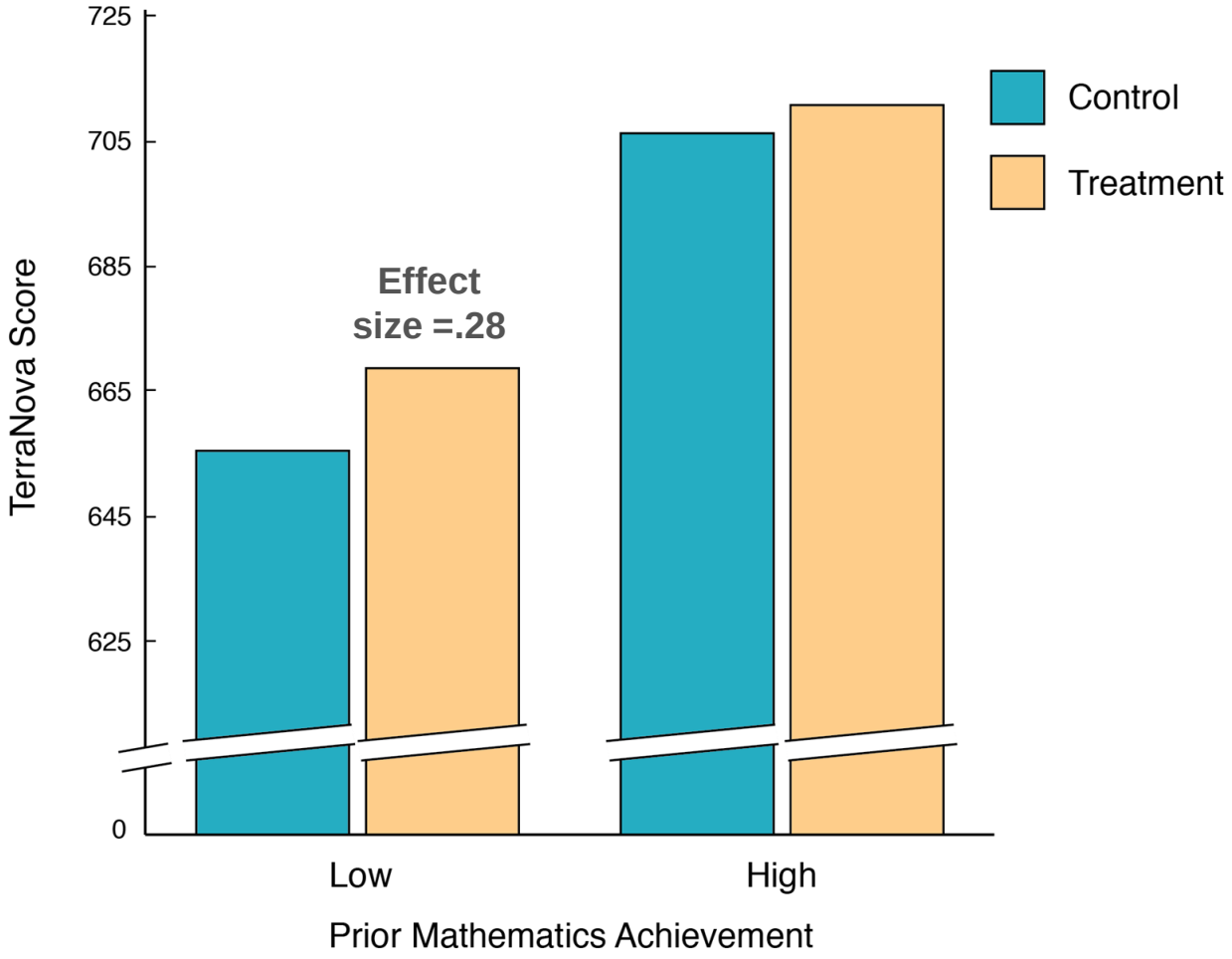
Many of the schools with the highest TerraNova scores were those who used ASSISTments (blue).



Average TerraNova Scores of Participating Schools by Condition

Condition	# of Schools	# of Students	Mean TerraNova Score	Effect Size (g)	*p*-value
Treatment	22	1,572	692.20	0.27	0.004
Control	22	1,204	682.04		

Finding 3:Greater Effect for Low Prior Math



Makes sense: higher achievers need less homework help

Estimating the Effect of Web-Based Homework

Kim Kelly[1], Neil Heffernan[1], Cristina Heffernan[1],
Susan Goldman[2], James Pellegrino[2], Deena Soffer Goldstein[2]

Worcester Polytechnic Institute[1]
University of Illinois –Chicago[2]
[kkelly, nth]@wpi.edu

Abstract. Traditional studies of intelligent tutoring systems have focused on their use in the classroom. Few have explored the advantage of using ITS as a web-based homework (WBH) system, providing correctness-only feedback to students. A second underappreciated aspect of WBH is that teachers can use the data to more efficiently review homework. Universities across the world are employing these WBH systems but there are no known comparisons of this in

Homework

The Experiment

- **Experiment:** Correctness feedback and went over homework with the item report.
- **Control:** Entered in test mode, read answers in class the next day.

Kelly, K., Heffernan, N., Heffernan, C., Goldman, S., Pellegrino, G. & Soffer, D. (2013). Estimating the Effect of Web-Based Homework. In Lane, Yacef, Motow & Pavlik (Eds) The Artificial Intelligence in Education Conference. Springer-Verlag. pp. 824-827. A longer version is available here.

Control

#213215 Data driven	#213216 Data driven	#213217 Data driven
36%	42%	45%
0%	0%	0%
1/4^4,23% +feedback 1/1^8,23% +feedback	1/1^2,21% +feedback	1/1^7,27% +feedback

✓ 1/4^8	✓ 1/2^2	✓ 1/7^7
✗ 1/1^8	✗ 1/1^2	✗ 1/1^7
✓ 1/4^8	✓ 1/2^2	✓ 1/7^7
✗ 1/4	✗ 1/2	✗ 7
✓ 1/4^8	✗ 1/2^12	✗ 1/7^17
✗ 4/4^4	✓ 1/2^2	✓ 1/7^7

Experimental

#213215 Data driven	#213216 Data driven	#213217 Data driven
21%	55%	85%
0%	0%	0%
16^4,19% +feedback 1/4^4,15% +feedback		

✓ 1/4^8	✓ 1/2^2	✓ 1/7^7
✗ 1/4^4 1 times	✗ 1/2^13 1 times	✓ 1/7^7
✗ 16^4 1 times	✗ 1/1^2 1 times	✓ 1/7^7
✗ 4/4^4	✓ 1/2^2	✓ 1/7^7
✗ 1 1 times	✗ 2^2	✓ 1/7^7
✓ 1/4^8	✓ 1/2^2	✓ 1/7^7

Video of Classroom from Both Condtions

Homework Results

<u>The Results</u>

	Pre	Entering School	After Going Over
• Control	8%	58%	65%
• Experiment	7%	70%	82%

- Effect Size= .5 standard deviation

Put Connected Math answers into ASSISTments

- Rationale
  - Need to understand what questions are hard
  - What are the common wrong answers

ARRS

Skill Building with Automatic Re-Assessment and Re-Training(ARRS)

24 - Scientific Notation - THE SKILL BUILDING SET (Problem Set 11893)
Class progress: 18 not started, 6 in progress, 78 complete
Release date: November 08, 2011 06:00 AM
Due date: November 11, 2011 10:00 PM

23 - Greatest Common Factor (Problem Set 11892)
Class progress: 12 not started, 10 in progress, 80 complete
Release date: October 25, 2011 06:00 AM
Due date: October 28, 2011 10:00 PM

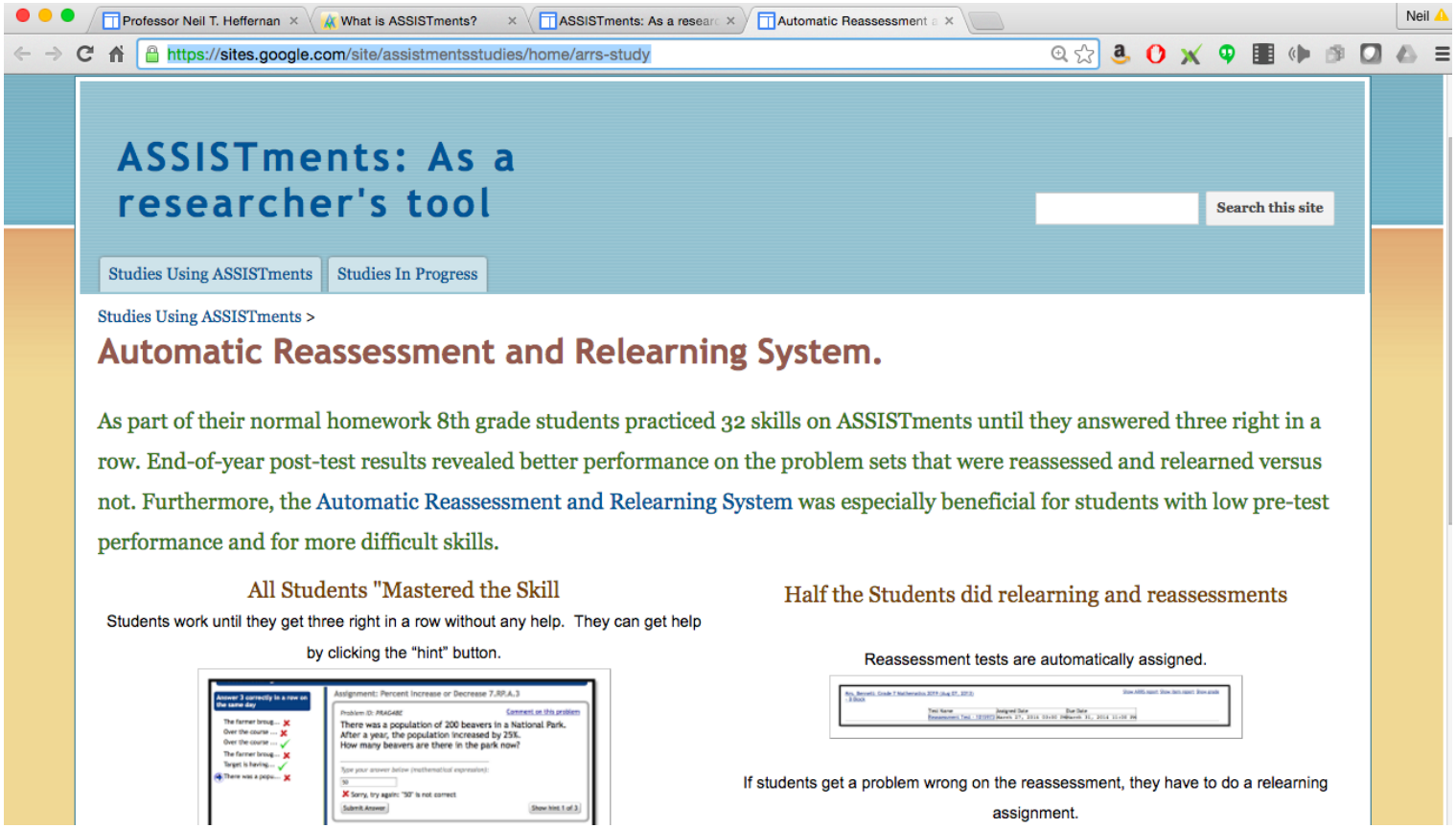
22 - Prime Factorization (Problem Set 7175)
Class progress: 13 not started, 0 in progress, 89 complete
Release date: October 25, 2011 06:00 AM
Due date: October 28, 2011 10:00 PM

21 - Exponents (Problem Set 11890)
Class progress: 10 not started, 0 in progress, 92 complete
Release date: October 18, 2011 06:00 AM
Due date: October 21, 2011 10:00 PM

20 - Divisibilty (Problem Set 11889)
Class progress: 11 not started, 3 in progress, 88 complete
Release date: October 18, 2011 06:00 AM
Due date: October 21, 2011 10:00 PM

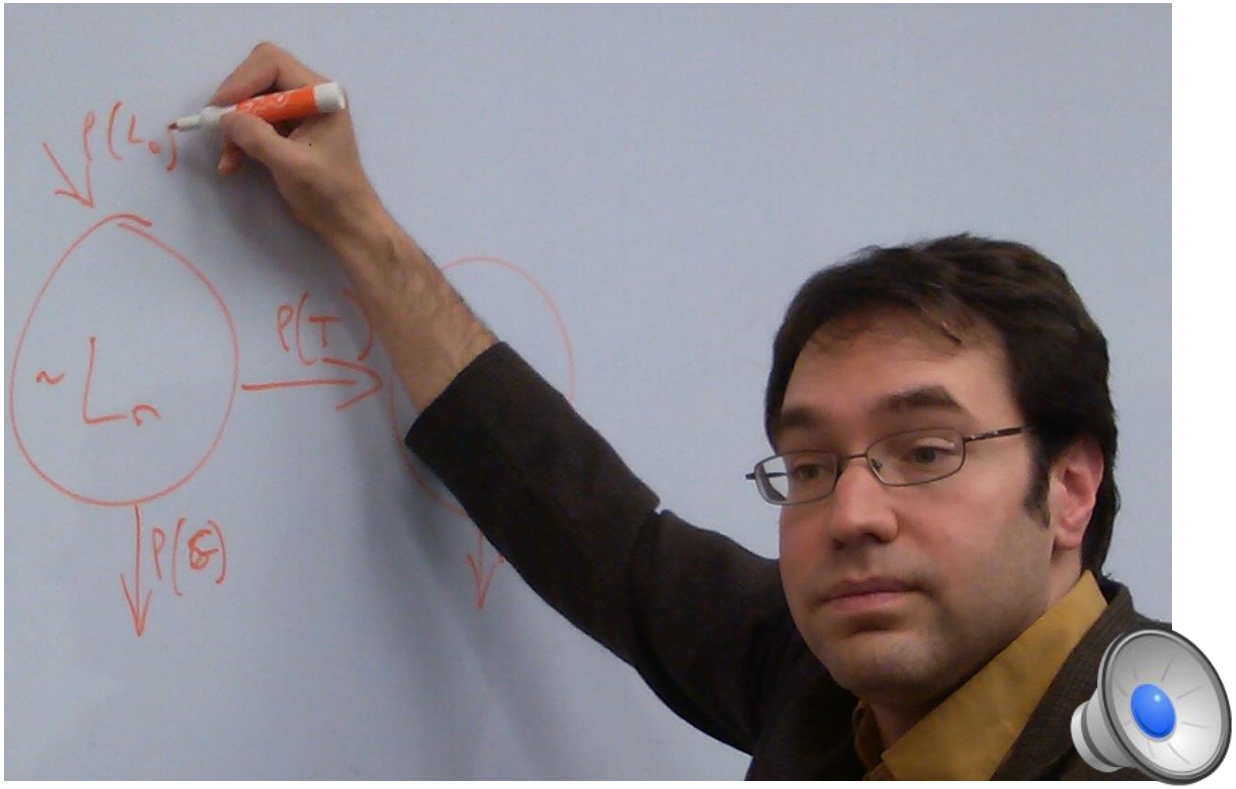
19 - Dividing Decimals (Problem Set 31278)
Class progress: 12 not started, 2 in progress, 88 complete
Release date: October 18, 2011 06:00 AM
Due date: October 21, 2011 10:00 PM

Spacing: Skill Builders and ARRS



Soffer, D., Das, V., Pellegrino, G., Goldman, S., Heffernan, N., Heffernan, C.,& Dietz, K. (2014) Improving Long-term Retention of Mathematical Knowledge through Automatic Reassessment and Relearning. American Educational Research Association (AERA 2014) Conference. Division C - Learning and Instruction / Section 1c: Mathematics. Link Nominated for the best poster of the session.

ITEST project :Ryan Baker: Affect Detection



Methodology: BROMP Protocol

- Android App: Time synchronized to the software that students were using
- Inter-rater reliability was high (Kappa was high)
- Then constructs features to try to help predict the judgments (treated as ground truth)
- 3000 observations
- Why sensor-less detectors?

Ryan Baker: Affect Dectors

- [x] Enable Affect Detector Key
  - ' C ' indicates 'Confused' affect
  - ' B ' indicates 'Bored' affect
  - ' F ' indicates 'Frustrated' affect
  - ' E ' indicates 'Engaged Concentration' affect
  - ' - ' indicates 'Not' state

We have several good publications

- These measure can be used to predict
  - End of Year Test state scores
  - Whether the student goes to college
  - Whether the student majors in STEM
  - How their gaming behavior have long term correlations
- Surprising Findings
  - Students that do the best on state exams show the most frustration
  - Frustration might not be bad

1. First paper reporting we can measure affect
    - San Pedro, M., Baker, R., *Gowda, S.,* & Heffernan, N. (2013). Towards an Understanding of Affect and Knowledge from Student Interaction with an Intelligent Tutoring System. In Lane, Yacef, Motow & Pavlik (Eds) *The Artificial Intelligence in Education Conference.* Springer-Verlag. pp. 41-50.
2. Then we showed we can better predict state test scores
    - Pardos, Z.A., Baker, R.S.J.d., San Pedro, M.O.C.Z., Gowda, S.M., Gowda, S.M. (2014) Affective States and State Tests: Investigating How Affect and Engagement during the School Year Predict End-of-Year Learning Outcomes. Journal of Learning Analytics, 1(1), 107–128.
        1. *First appeared as Pardos, Z.A., Baker, R.S.J.d., San Pedro, M.O.C.Z., Gowda, S.M., Gowda, S.M. (2013) Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes. Proceedings of the 3rd International Conference on Learning Analytics and Knowledge, 117-124.*
3. Then we showed we can predict who enrolls in colleage years later
    - *San Pedro, M.,* Baker, R., Bowers, A. & Heffernan, N. (2013) Predicting College Enrollment from Student Interaction with an Intelligent Tutoring System in Middle School. In S. D'Mello, R. Calvo, & A. Olney (Eds.) Proceedings of the 6th International Conference on Educational Data Mining
4. Then we showed we can predict college major they will pursue
    - San Pedro, M., Ocumpaugh, J., Baker, R., & Heffernan, N. (2014) Predicting STEM and Non-STEM College Major Enrollment from Middle School Interaction with Mathematics Educational Software. In John Stamper et al. (Eds) Proceedings of the 7th International Conference on Educational Data Mining. pp. 276-279. A longer version is here
5. We followed that up with looking at how gaming the system was important in making the decesion.
    - San Pedro, M.O., Baker, R., Heffernan, N., Ocumpaugh, J. (2015) Exploring College Major Choice and Middle School Student Behavior, Affect and Learning: What Happens to Students Who Game the System? Proceedings of the 5th International Learning Analytics and Knowledge Conference. pp 36-40.
6. Along the way Baker et al lead the team to making sure the decetors generalize across urban, suburban and rural areas
    - Ocumpaugh, J., Baker, R., Gowda, S., Heffernan, N., Heffernan, C. (2014) Population validity for Educational Data Mining models: A case study in affect detection. *British Journal of Educational Technology, 45* (3), 487-501. OI: 10.1111/bjet.12156
7. Finally, Heffernan has refit the decectors with new features
    - Wang, Y., Heffernan, N, & Heffernan, C. (2015) Towards better affect detectors: effect of missing skills, class features and [illegible]ong answers. Proceedings of the Fifth International Conference on Learning Analytics And Knowledge. pp 31-35. See data here and [illegible]ere