Learning Math by Teaching Math: A Case Study in Intelligent Tutoring and Educational Data Mining

Ma. Mercedes T. Rodrigo

Ateneo Laboratory for the Learning Sciences
http://penoy.admu.edu.ph/~alls

A word of thanks


Matsuda


Koedinger


Keiser


Yarzebinski


Ong


Bringula


Dela Cruz


Basa

Let's talk about you

- What are your usual teaching strategies for learning math?
- What are the usual assessment strategies?
- How do your teachers process the assessment data?

Context

- How can we use computers for mathematics education

The usual suspects

- Drills
- Educational games

Less usual, less common suspects

- Intelligent tutoring systems

What are ITSs?

- Computer-based learning environments that use artificial intelligence to provide the learner with customized feedback and guidance

Outline

- Introduction to SimStudent
- Data collection methods
- SimStudent findings
  - Effectiveness of learning by teaching
  - Self-explanation
  - Prior learning and other factors

Outline

- Introduction to SimStudent
- Data collection methods
- SimStudent findings
  - Effectiveness of learning by teaching
  - Self-explanation
  - Prior learning and other factors

Learning by Teaching

- A well-known effective learning strategy
  - Across ages, domains, structures, etc…
- Cognitive and social theories have yet to be investigated
  - When and how students learn by teaching?

Solution: Teachable Agent

- Pedagogical agent that students can teach
- Does the TA work as a peer for students to learn by teaching?
- Do we need a genuine machine learning technology?
  - What is it the ML brings us that would otherwise impractical?

SimStudent: a teachable agent

- Learns cognitive skills from tutored-problem solving
- Programming by demonstration
  - Production model
  - Hybrid learning for what, when, and how to apply a skill
    - Version space
    - Inductive logic programming
    - Iterative-deepening search

Tutoring Stacy

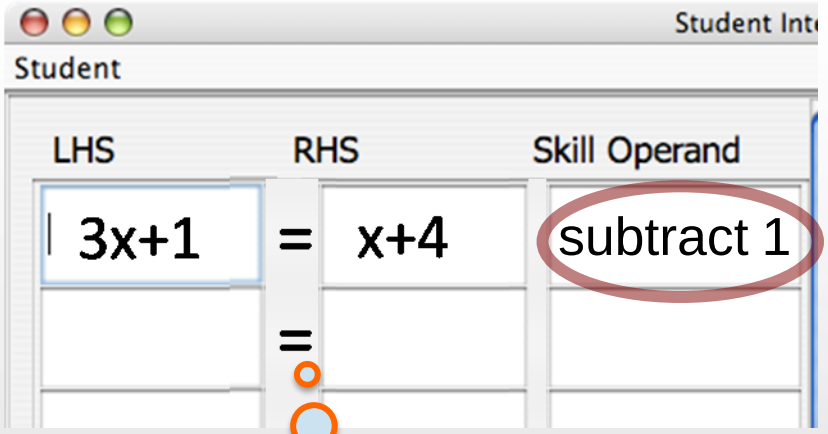
- Try tutoring Stacy to solve problems such as
  - 3x = 6
  - 2x+1 = 5
  - 5x+3 = 6x-2
- Try tutoring the WRONG rules and see what happens

SimStudent models human learning

- It is possible for SimStudent to learn the wrong thing

Example: Learning to subtract a constant term

First example



I see 3x, 1, x, and 4 in the equation.
I wonder where the '1' came from…

Learning to subtract a constant number

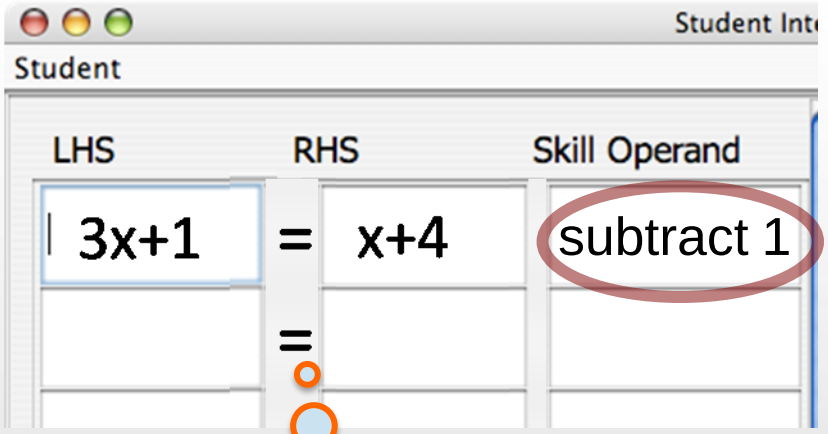
Subtract the difference between 4 and 3…

Subtract the coefficient of X…

Subtract the last term on the left-hand side…

Example: Learning to subtract a constant term

First example



I see 3x, 1, x, and 4 in the equation.
I wonder where the '1' came from…

Learning to subtract a constant number

Subtract the difference between 4 and 3…

Subtract the coefficient of X…

Subtract the last term on the left-hand side…

Outline

- Introduction to SimStudent
- Data collection methods
- SimStudent findings
  - Shallow learning
  - Self-explanation
  - Prior learning and other factors

Structure of the Study

Control Condition

Target unit:
Equation with variables on both sides

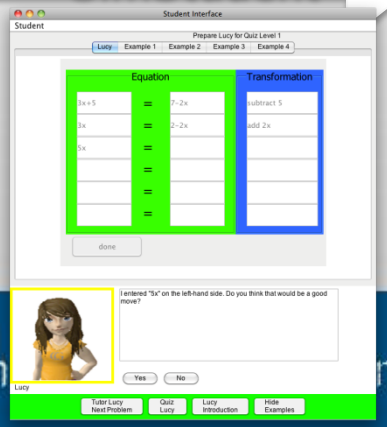
Day 1: CT Review

Day 2: Pre-test

Day 3-4(or 5)

Within classroom RCT

Day 3-4(or 5): SimStudent



Day 6: Post-test

2 weeks: Delayed-test

Three versions, counter balanced

Pre, post, and delayed post

- Isomorphic
- Counterbalanced
- Divided into procedural and conceptual parts
- Only the procedural results were used to measure gains

Process data

- Problems tutored
- Feedback provided
- Steps performed
- Examples reviewed
- Hints requested
- Quiz attempts

Data collection venues

- Pittsburgh, PA
- University of the East, Manila

Outline

- Introduction to SimStudent
- Data collection methods
- SimStudent findings
  - Effectiveness of learning by teaching
  - Self-explanation
  - Prior learning and other factors

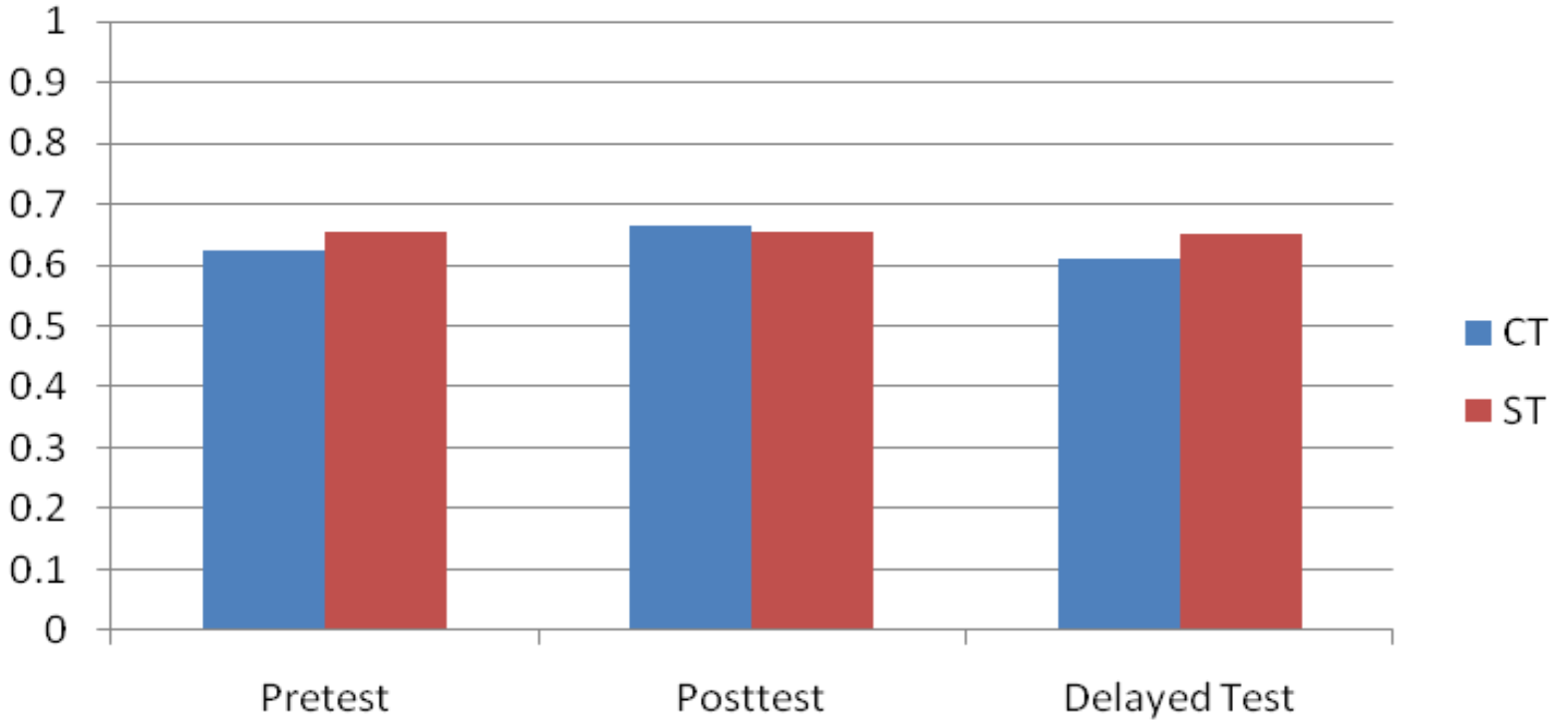
What do you think?

- Is learning by teaching effective?
- Will the students using SimStudent learn more than those in the control condition?

And the results are…

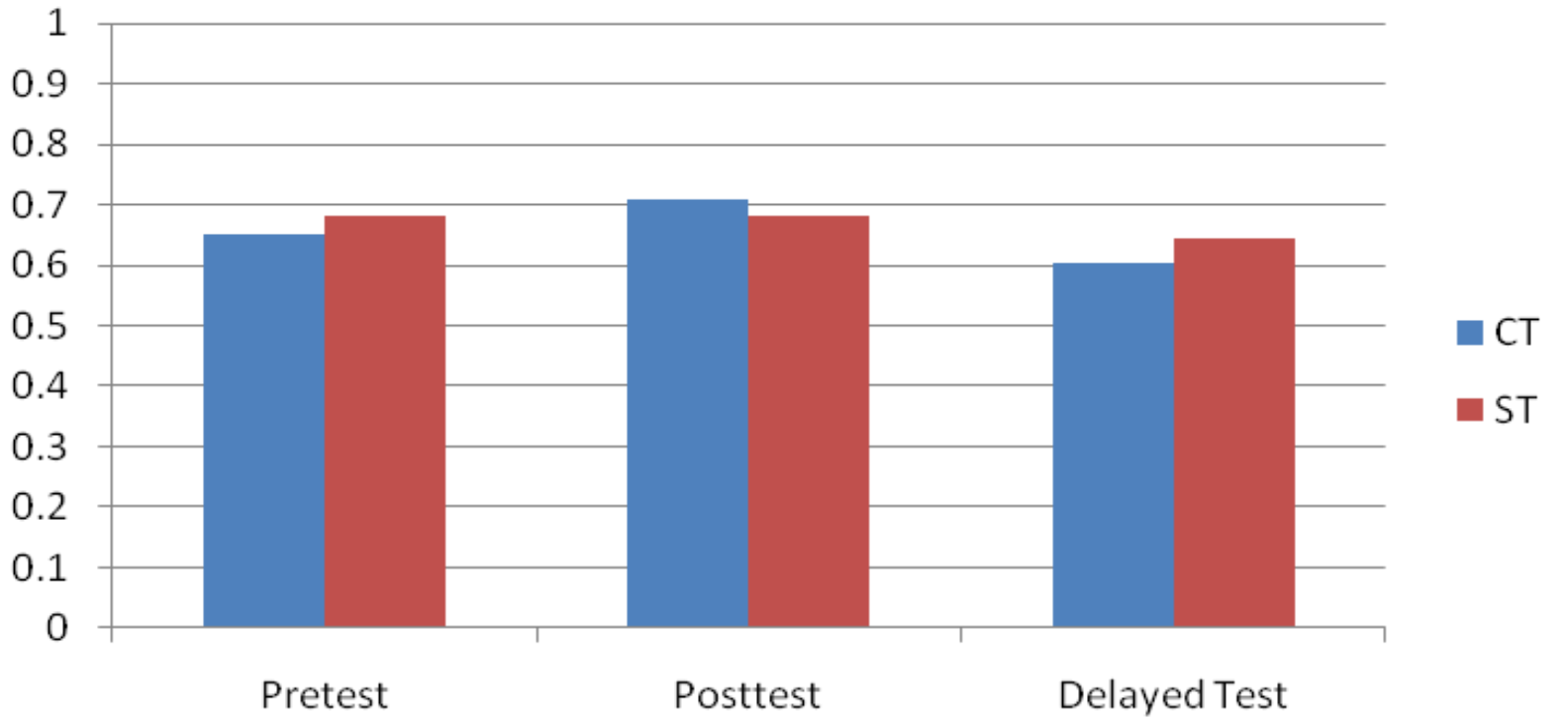
Overall Test Scores

N=74



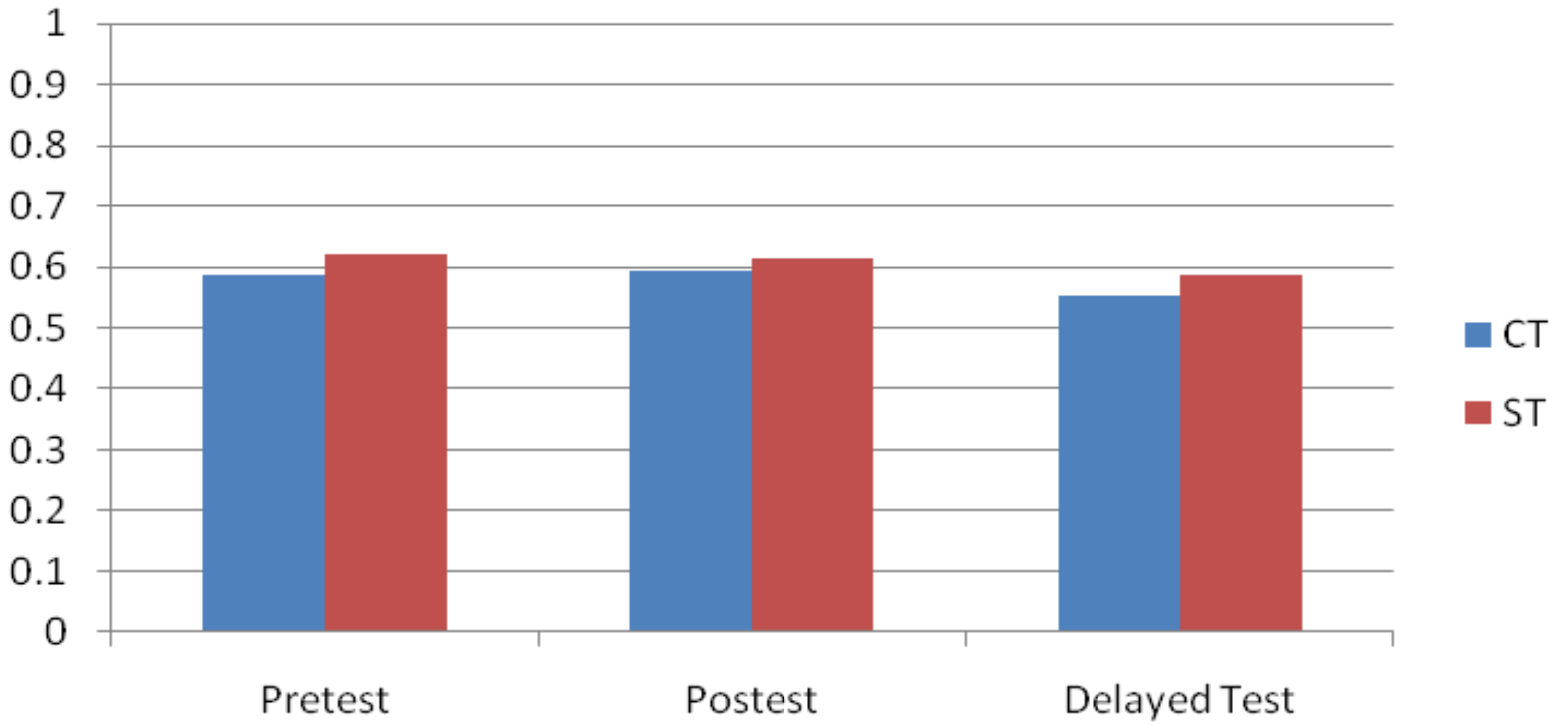
Procedural Items

N=74



Conceptual Items

N=74

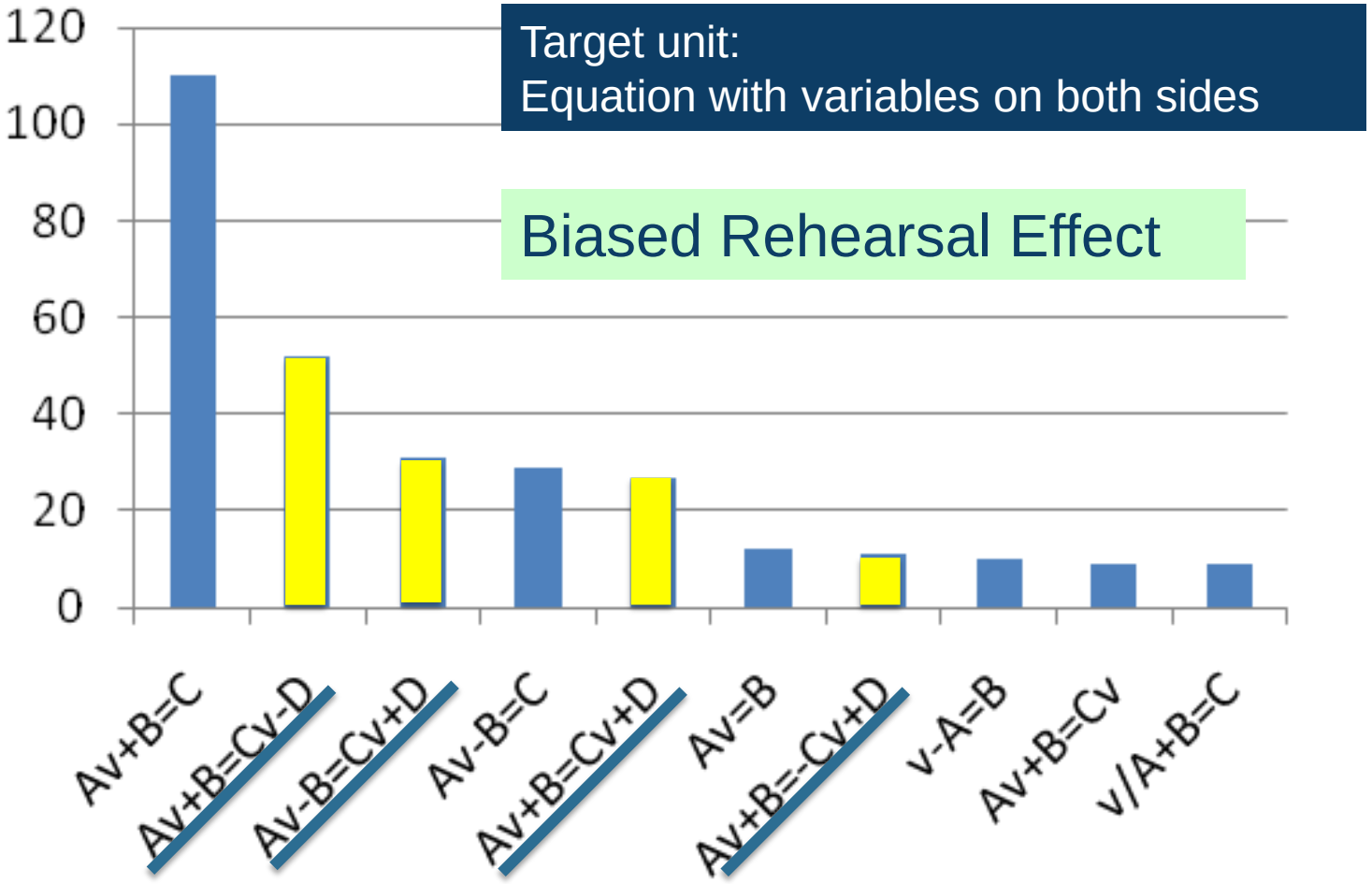


In brief

- No significant learning gains with SimStudent
- Students with weak prior knowledge learned more from the control condition

Why???

Problems Used for Tutoring



Outline

- Introduction to SimStudent
- Data collection methods
- SimStudent findings
  - Effectiveness of learning by teaching
  - Self-explanation
  - Prior learning and other factors

Self-Explanation Hypothesis

- Self-explanation facilitates learning from examples
- The same effect for such a meta-cognitive reflection might occur for tutor learning
- Prompting students to explain and justify their tutoring activities and decision might facilitate tutor learning

SimStudent asks for explanations

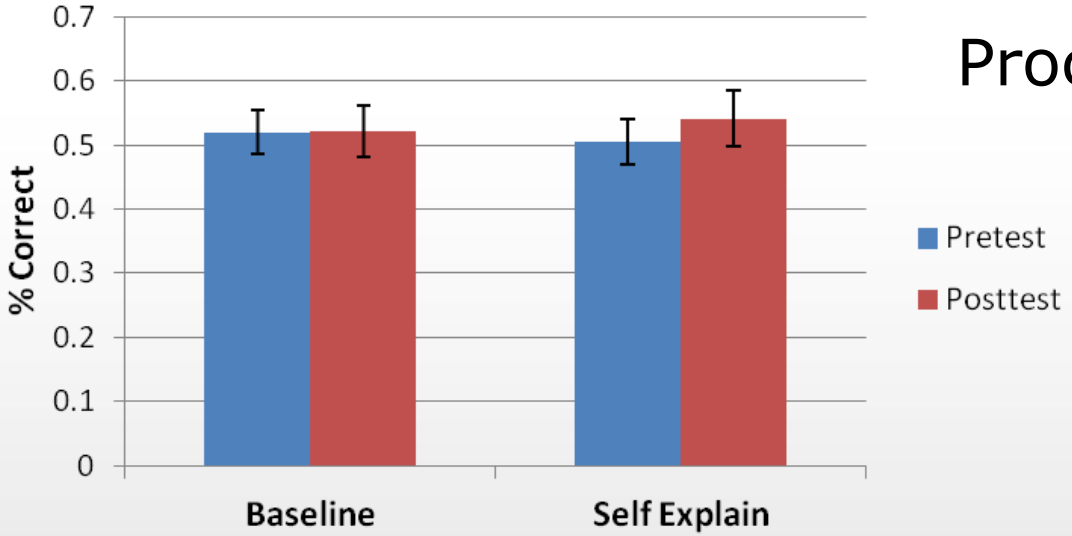
- A new problem posed by a student
  - "why should I do this problem?"
- Corrective feedback on an error
  - "But I put divide 3 for 3x=9. Why doesn't divide 2 work now?"
- A step demonstrated by the student as a hint
  - "Why did you do such?"

What do you think?

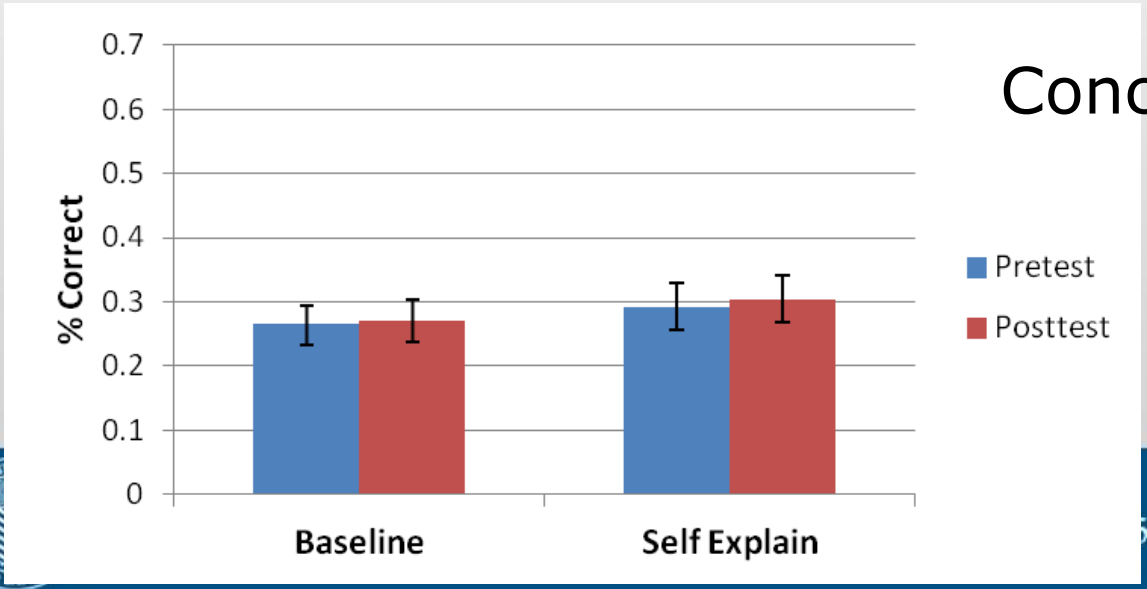
- Is self-explanation effective?
- Will students who self-explain learn more than those who do not?

And the results are...

Results: Test Scores



Procedural Skill Test



Conceptual Knowledge Te

Self-Explanation Effect for Tutor Learning

- Regression analysis
  - Dependent variable
    - Normalized gain for the procedural skill test
  - Independent variables
    - Type of explanation (new problem, error, hint)
    - Format of explanation (dropdown, free input, mix)
    - Degree of elaboration (shallow, elaborated)
    - Frequency of denial (SE skipped / total question asked)
    - Domain specificity (# of math term used)
- No notable effect of self-explanation for tutor learning

In brief

- Overall, the version of APLUS and SimStudent used for the Self-Explanation study showed weak effect on procedural skill acquisition
  - The latest version showed a significant main effect of test (pre vs. post)
- SimStudent did learn skills from students – Quiz passed
  - This was actually shallow learning!
- SE students achieved the same test scores with fewer problems.
- No particular self-explanation effect confirmed
  - System may need to provide constructive feedback for "shallow" explanations

Outline

- Introduction to SimStudent
- Data collection methods
- SimStudent findings
  - Effectiveness of learning by teaching
  - Self-explanation
  - Prior learning and other factors

Prior knowledge

- Pre-test scores are highly predictive of post-test scores.
- Why do some students with low prior knowledge do better than others?
- What other factors affect tutor learning?

US vs PH study

- Pre, post, and delayed post
- SimStudent SE vs baseline

Overall test results

- PH
  - No statistically significant difference between pre and post
  - Statistically significant difference between pre and delayed post
- No condition difference in post-test scores in either country
  - SEs did not affect tutor learning

Country differences

- US students had higher prior learning
- Did better on the post-test
- Learned more

Within the PH dataset

- Some students in the PH dataset did better than others.
- Their SimStudent passed 2 out of 4 quiz sections (termed "passing Sx students")
- Why?
- What did they do that was different?

Copying quiz items

- Did passing S1 and S2 students use more quiz items for tutoring?

Copying quiz items

- Did passing S1 and S2 students use more quiz items for tutoring?

- No.
  - No notable difference in number of quiz items tutored.

Accuracy of tutoring

- Did students passing S1 and S2 students tutor their SimStudents more accurately?

Accuracy of tutoring

- Did students passing S1 and S2 students tutor their SimStudents more accurately?

- Yes!
  - Accuracy of tutoring is a key for success!
- Why were some students more accurate?

Prior knowledge

- Was there a correlation between prior knowledge and accuracy of tutoring?

Prior knowledge

- Was there a correlation between prior knowledge and accuracy of tutoring?

- Yes!
  - US students had higher pre-test scores and tutored more accurately

Prior knowledge

- In the PH
  - There was no diff in prior knowledge between passing and failing S1 students,
  - There was a difference in accuracy of tutoring
  - Weak trend in on the average normalized gain from pre- to post- in favor of passing students
- Passing S1 students learned more

Learning strategy and resource usage

- Did passing S1 students tutor more problems?

Learning strategy and resource usage

- Did passing S1 students tutor more problems?

- No.

Learning strategy and resource usage

- Did passing S1 students refer to more examples?

Learning strategy and resource usage

- Did passing S1 students refer to more examples?

- On average, yes, but this was not statistically confirmed.

Learning strategy and resource usage

- Did passing S1 students copy more example problems for tutoring?

Learning strategy and resource usage

- Did passing S1 students copy more example problems for tutoring?

- Yes!
  - PH students were switching probably copying the worked out examples line-by-line

What did we learn?

- Learning by teaching is tricky.
- Self-explanations can help…a bit.
- Prior knowledge matters…a lot.
- Accuracy of tutoring matters.
- Copying worked-out examples can help.

How did we learn all this?

- SimStudent and similar environments allow us access to fine-grained process data
- Enables deep analysis of student behaviors
- Provides empirical evidence to support or refute theory.

More thanks

- Marc Lester Armenta
- Regina Ira Antonette M. Geli
- Gabriel Jose G. Vitug

And thank *you*!

- Any questions?