Revising Tutor Responses in Math Education: Teacher Insights and NLP Strategies



Rose E. Wang



Carly Robinson



Dora Demszky

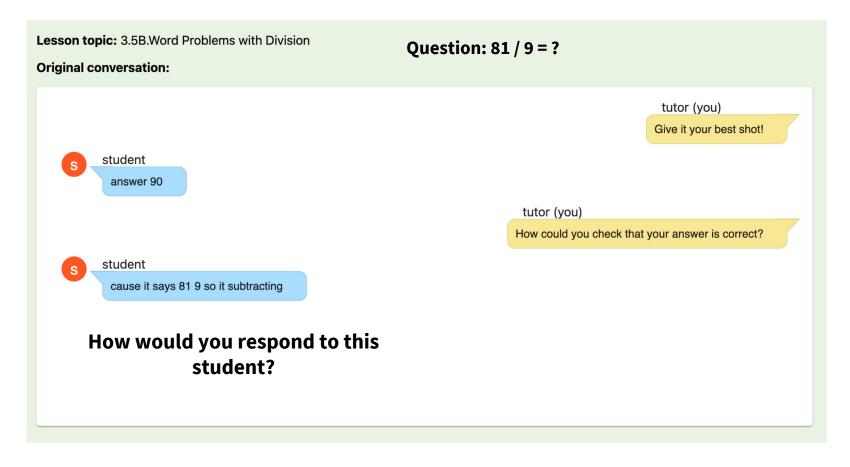
Stanford University

Text-based tutoring is a common approach to tutoring

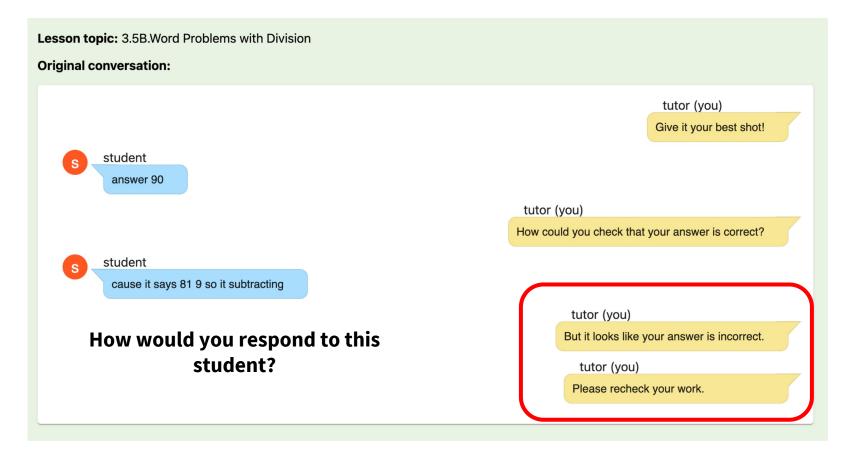
Advantages: Flexible (on-the-go demand and supply) and convenient

Let's take a look at some real examples of these textbased tutoring conversations

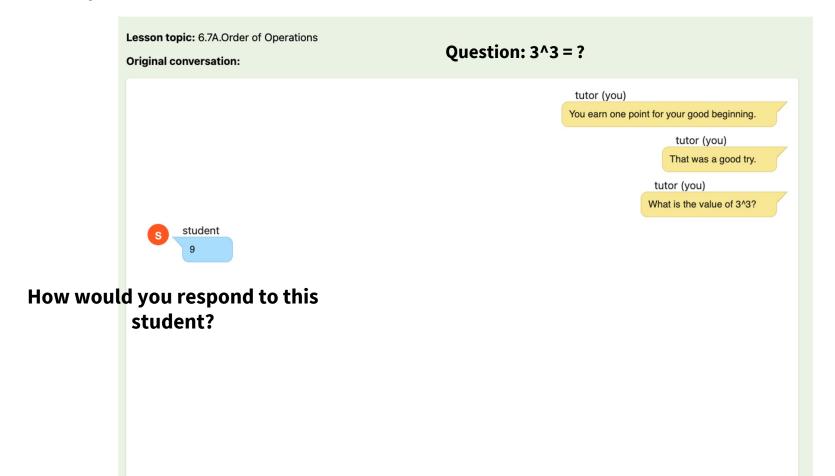
Example 1 (note: this is a real tutor-student conversation)



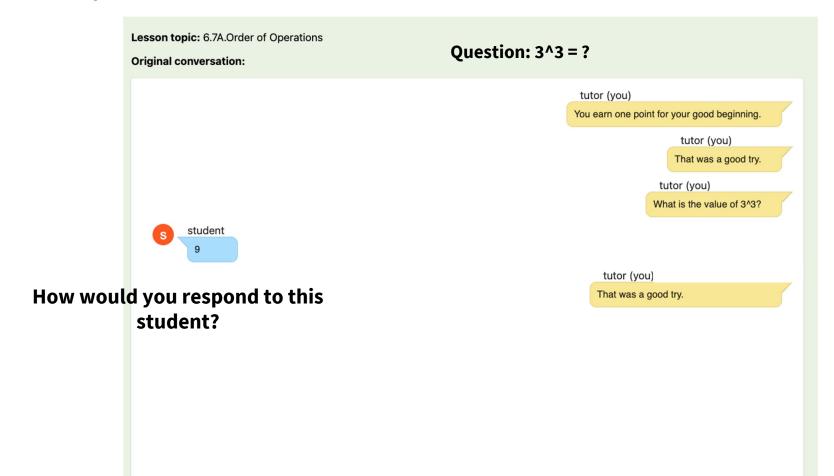
Example 1 (note: this is a real tutor-student conversation)



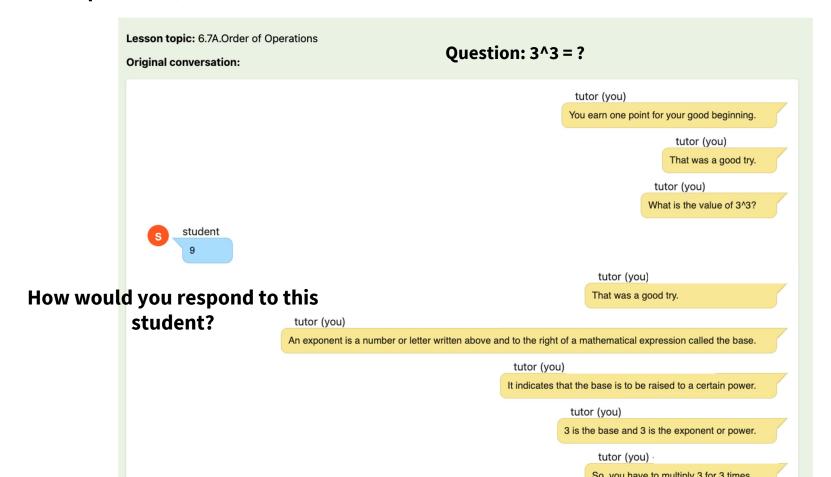
Example 2 (note: this is a real tutor-student conversation)



Example 2 (note: this is a real tutor-student conversation)



Example 2 (note: this is a real tutor-student conversation)



Issues with current responses to student mistakes

Tutors have domain knowledge about mathematics but

They provide unhelpful responses to students who struggle.

tutor (you) - REVISE THIS MESSAGE!

But it looks like your answer is incorrect.

tutor (you) - REVISE THIS MESSAGE!

Please recheck your work.

tutor (you) - REVISE THIS MESSAGE!

tutor (you) - REVISE THIS MESSAGE!

An exponent is a number or letter written above and to the right of a mathematical expression called the base.

 They do not respond to the type of error the student makes. tutor (you) - REVISE THIS MESSAGE!

It indicates that the base is to be raised to a certain power.

tutor (you) - REVISE THIS MESSAGE!

3 is the base and 3 is the exponent or power.

tutor (you) - REVISE THIS MESSAGE!

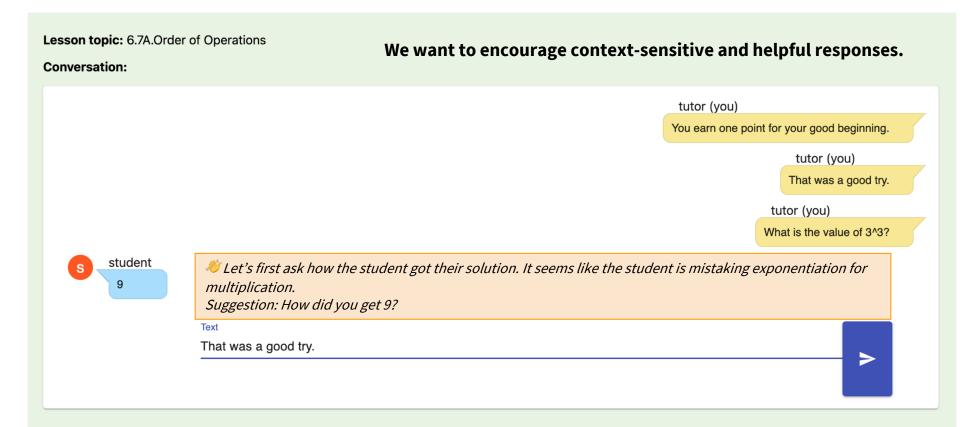
So, you have to multiply 3 for 3 times.

tutor (you) - REVISE THIS MESSAGE!

Good work so far.

That was a good try.

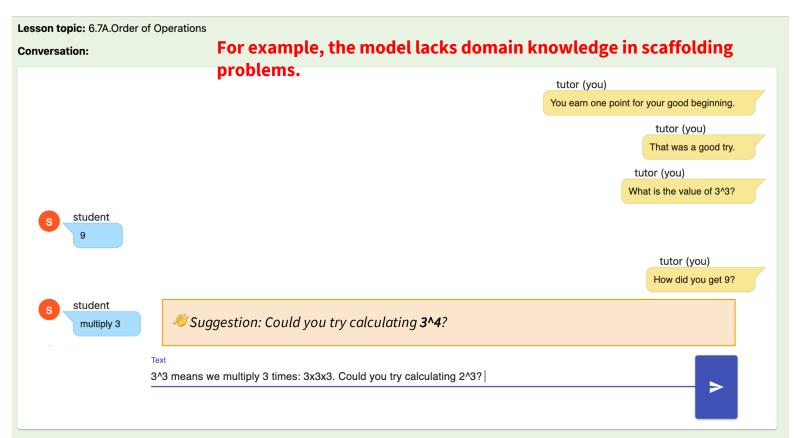
Our solution: Leveraging teacher insights and NLP-augmented methods to automate revisions



Our solution: Leveraging teacher insights and NLP-augmented methods to automate revisions

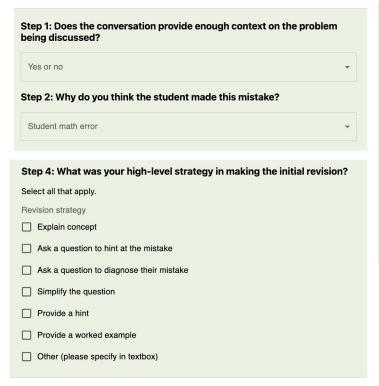


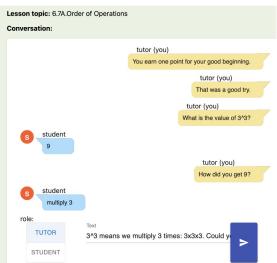
Our solution: Leveraging teacher insights and NLPaugmented methods to automate revisions



Ongoing work and next steps — feedback welcomed!

Collaborating with math teachers to collect a large revision dataset





Evaluate large language models like ChatGPT and GPT4 on this task

Train our own models with this data and investigate the importance of revising wrt. student error type