Providing **Joint** Talk-Based Feedback To Instructors and Learners

EDUC 473/CS 393

CUEMATH

schoolhouse

Course logistics

- HW3 due this Friday
- Any questions about blog post? We'll share past reading commentaries w/ you today!
- Plan for today:
  - Dora's brief lecture
  - Q&A with Mariah
  - Paper discussion by Aakriti, Alex, Mayank, Kaiyu

CUEMATH

schoolhouse

Learning Analytics and Knowledge (LAK)

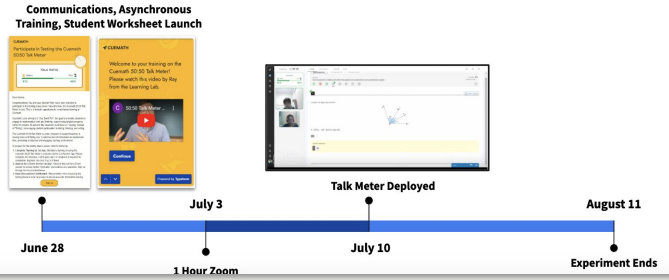
Does Feedback on Talk Time Increase Student Engagement? Evidence from a Randomized Controlled Trial on a Math Tutoring Platform

Dorottya Demszky
ddemszky@stanford.edu
Stanford University
United States

Rose E. Wang
Stanford University
United States
rewang@stanford.edu

Sean Geraghty
CueMath
India
sean.geraghty@cuemath.com

Carol Yu
CueMath
India
carol.yu@cuemath.com



Learning at Scale (L@S) Work-in-Progress paper

Enhancing Tutoring Effectiveness Through Automated Feedback: Preliminary Findings from a Pilot Randomized Controlled Trial on SAT Tutoring

Joy Yun*
joyyun@stanford.edu
Stanford University
Palo Alto, California, USA

Yann Hicke*
ylh8@cornell.edu
Cornell University
Ithaca, New York, USA

Mariah Olson
mariah@schoolhouse.world
Schoolhouse.world
Mountain View, California, USA

Dorottya Demszky
ddemszky@stanford.edu
Stanford University
Palo Alto, California, USA

ABSTRACT

To address educational inequities, high-quality SAT tutoring is crucial for students who need support. Many tutors, however, are novices and require coaching to be effective. To examine this issue, we conducted a pilot randomized controlled trial (RCT) to evaluate the effectiveness of personalized, automated feedback for novice tutors. This feedback aimed to enhance their tutoring skills and, consequently, improve student outcomes. In our RCT, we not only assessed the effectiveness of the feedback provided to novice tutors but also examined the impact of extending this feedback to both tutors and their students, compared to just the tutors alone. Furthermore, we explored how the use of social versus personal goal-oriented language in the feedback influences educational outcomes. Our preliminary findings from this pilot indicate that providing feedback to both tutors and learners led to a statistically significant improvement in average SAT practice test scores. Additionally, this approach significantly increased the tutor talk time ratio, an outcome that was somewhat unexpected and requires further in-

ACM Reference Format:
Joy Yun, Yann Hicke, Mariah Olson, and Dorottya Demszky. 2024. Enhancing Tutoring Effectiveness Through Automated Feedback: Preliminary Findings from a Pilot Randomized Controlled Trial on SAT Tutoring. In *Proceedings of the Eleventh ACM Conference on Learning @ Scale (L@S '24), July 18–20, 2024, Atlanta, GA, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3657604.3664681

1 INTRODUCTION

SAT scores play a crucial role in college admissions, influencing students' future career paths. However, high-quality SAT tutoring, typically costly and accessible mostly to affluent students, exacerbates educational inequities. In response, free volunteer SAT tutoring initiatives like Schoolhouse.world have emerged, serving over 74,300 learners with the help of more than 10,800 volunteer tutors. These programs offer a potential solution to reduce disparities in

CUEMATH

Learning Analytics and Knowledge (LAK)

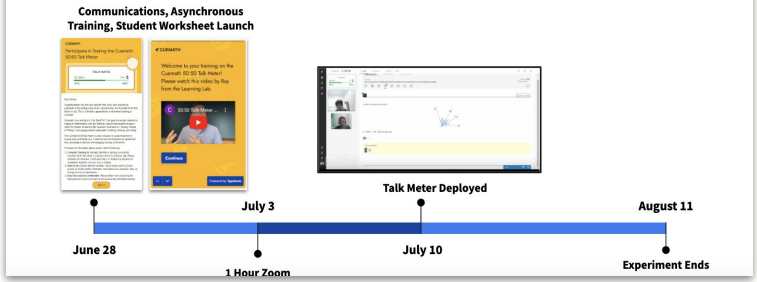
Does Feedback on Talk Time Increase Student Engagement? Evidence from a Randomized Controlled Trial on a Math Tutoring Platform

Dorottya Demszky
ddemszky@stanford.edu
Stanford University
United States

Rose E. Wang
Stanford University
United States
rewang@stanford.edu

Sean Geraghty
CueMath
India
sean.geraghty@cuemath.com

Carol Yu
CueMath
India
carol.yu@cuemath.com



Importance of student talk in learning

- Talking about math is central to learning math
- There is a lot of room for improving student talk time in mathematics teaching

72-88% teacher talk in many contexts!

Motivating students to be active participants of their learning

- Usually engaging students in discourse falls on the shoulders of the teacher
- Automated feedback to teachers on talk time was shown to improve student engagement
- Would giving feedback to **BOTH** teachers and students be even more effective at motivating students to engage?

Research Questions

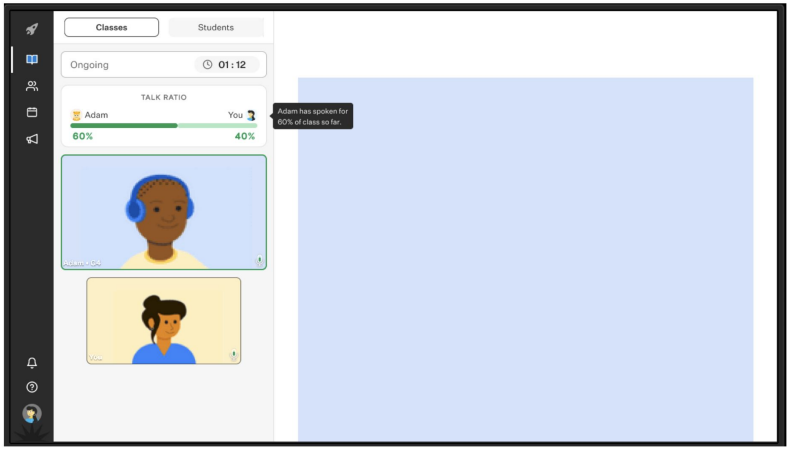
1. What is the **impact** of the TalkMeter on student-teacher interactions?
2. How did **tutors perceive** the Talk Meter and the impact it had on their instruction and student engagement?
3. How did **students perceive** the Talk Meter?

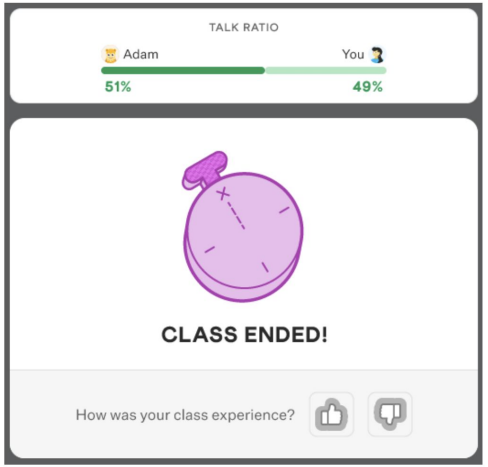
CueMath

- 1:1 math tutoring to ~37k students worldwide
- Most tutors are women from India with STEM backgrounds
- Students come from all over the world (esp. US, UK and India)
- Mostly grades 1-7, but also grades 8-12

Talk Meter

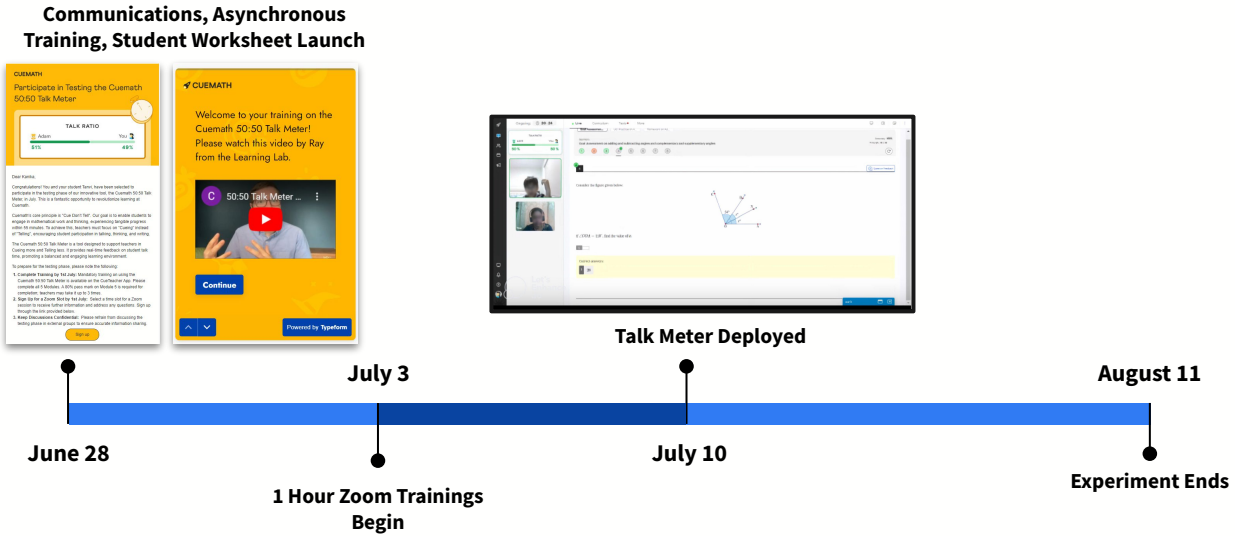
- Shown every 20 minutes
- Red (student talk <= 25%), yellow (between 25-50%) or green (>= 50%),





Randomized Controlled Trial

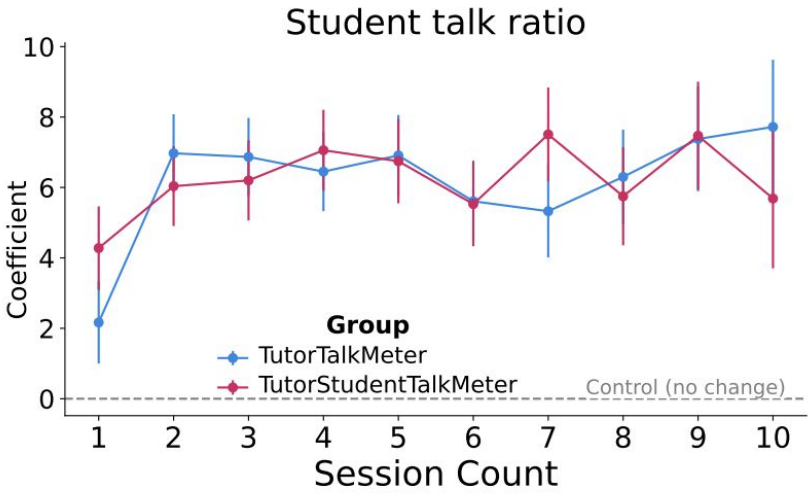
- 742 tutors and 1,266 students
- Three arms: (1) Control (2) Tutor Talk Meter (3) Tutor Student Talk Meter

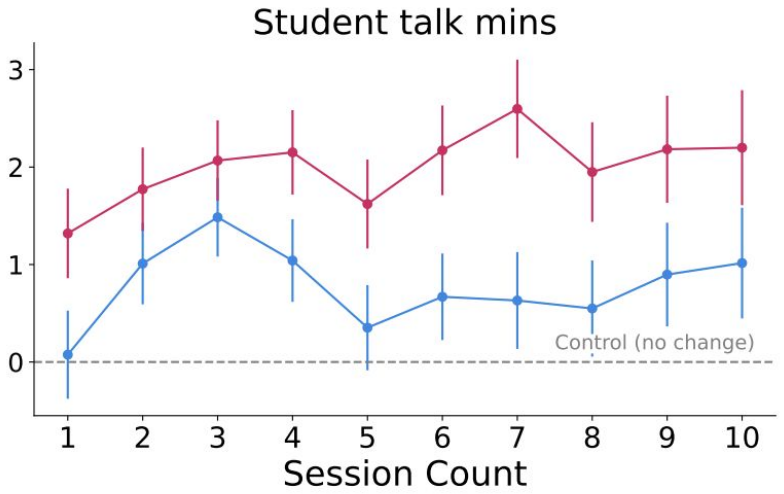


RQ1

Talk Meter increased student talk ratio

Student facing talk meter got students to talk more than the tutor facing talk meter alone!





RQ1

Talk Meter improved tutor questioning and students' use of math vocabulary

Student facing talk meter was more effective at improving quality of student math talk!

Table 3: Impact of the Talk Meter on language features.

	(1) Num. tutor focusing questions	(2) Num. tutor uptakes	(3) Num. tutor math terms	(4) Num. tutor unique math terms	(5) Num. student math terms	(6) Num. student unique math terms	(7) Num. student reasoning
G=TutorTM	0.987*	-0.225	-11.768**	-1.107**	5.567*	0.522	0.208
	(0.390)	(0.818)	(3.859)	(0.426)	(2.748)	(0.356)	(0.138)
G=TutorStudentTM	1.072**	1.383+	-9.670*	-0.873*	15.684**	1.566**	0.396**
	(0.404)	(0.813)	(3.980)	(0.394)	(3.253)	(0.349)	(0.145)
Control Mean	7.523	22.213	75.571	13.519	37.540	8.581	1.651
R2	0.275	0.398	0.188	0.239	0.220	0.244	0.194
Observations	2318	2318	2313	2313	2315	2315	2318

RQ2

Tutors found the Talk Meter helpful

- Increased awareness
- Student facing talk meter was effective at motivating students to talk, especially more introverted students
- It can feel unnatural to hold back from speaking
- Strategies:
  - open-ended questions
  - shortening explanations

> “It’s more impacted with [Student A] because [Student A] is one of my students who was really introvert. He hardly used to talk with me.... So once after this talk ratio, and still I’m struggling, but I think his participation has definitely increased.[. . .] [My other student S] is always excited. See, we have been keeping 50:50 ratio. And then sometimes he even said that see ma’am, I got the major ratio. I have been talking more. You’re not letting me to talk.

RQ3

Most students also liked the Talk Meter

- Gamification and competition
- Occasional intrusiveness

> When I see that it's red, I get a little bit sad and then I keep on talking, then I see it yellow, and then I keep on talking more. Then I see it green and then I'm super happy

> It's like a competition. So if you talk more, it's like, I think you're better at it.

> "It can get annoying because sometimes when I'm trying to look at a question, it just appears, and then sometimes I can't get rid of it.

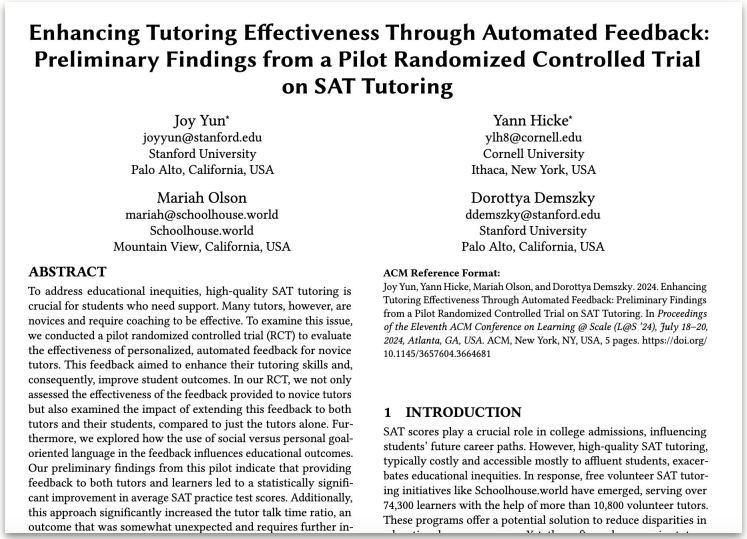
Conclusion & Future work

Talk Meter is a low cost way to increase student engagement

- Evaluating generalizability
- Improving user experience
- Expanding to other dimensions of talk

schoolhouse

L@S Work in Progress paper:



Enhancing Tutoring Effectiveness Through Automated Feedback: Preliminary Findings from a Pilot Randomized Controlled Trial on SAT Tutoring

Joy Yun*
joyyun@stanford.edu
Stanford University
Palo Alto, California, USA

Yann Hicke*
ylh8@cornell.edu
Cornell University
Ithaca, New York, USA

Mariah Olson
mariah@schoolhouse.world
Schoolhouse.world
Mountain View, California, USA

Dorottya Demszky
ddemszky@stanford.edu
Stanford University
Palo Alto, California, USA

ABSTRACT

To address educational inequities, high-quality SAT tutoring is crucial for students who need support. Many tutors, however, are novices and require coaching to be effective. To examine this issue, we conducted a pilot randomized controlled trial (RCT) to evaluate the effectiveness of personalized, automated feedback for novice tutors. This feedback aimed to enhance their tutoring skills and, consequently, improve student outcomes. In our RCT, we not only assessed the effectiveness of the feedback provided to novice tutors but also examined the impact of extending this feedback to both tutors and their students, compared to just the tutors alone. Furthermore, we explored how the use of social versus personal goal-oriented language in the feedback influences educational outcomes. Our preliminary findings from this pilot indicate that providing feedback to both tutors and learners led to a statistically significant improvement in average SAT practice test scores. Additionally, this approach significantly increased the tutor talk time ratio, an outcome that was somewhat unexpected and requires further in-

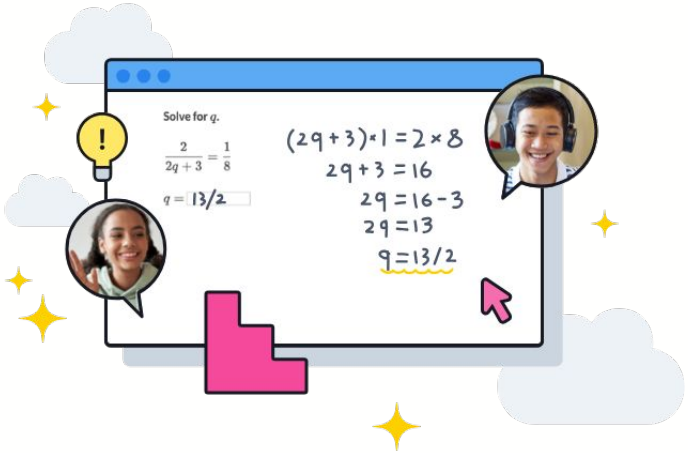
ACM Reference Format:
Joy Yun, Yann Hicke, Mariah Olson, and Dorottya Demszky. 2024. Enhancing Tutoring Effectiveness Through Automated Feedback: Preliminary Findings from a Pilot Randomized Controlled Trial on SAT Tutoring. In *Proceedings of the Eleventh ACM Conference on Learning @ Scale (L@S '24), July 18–20, 2024, Atlanta, GA, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3657604.3664681

1 INTRODUCTION

SAT scores play a crucial role in college admissions, influencing students' future career paths. However, high-quality SAT tutoring, typically costly and accessible mostly to affluent students, exacerbates educational inequities. In response, free volunteer SAT tutoring initiatives like Schoolhouse.world have emerged, serving over 74,300 learners with the help of more than 10,800 volunteer tutors. These programs offer a potential solution to reduce disparities in

Schoolhouse

- Free online peer-to-peer tutoring
- Co-Founded by Sal Khan
- Community of 10k tutors and 70k learners in 100 countries
- SAT Bootcamps:
  - 4 weeks long
  - meeting 2x week on Zoom
  - 1:10 tutor:learner ratio



Challenges

- Lots of passive learning (90%+ tutor talk time)
- Novice tutors
- Tutors are volunteers & resources are limited so challenging to do extensive training

Research Questions

1. What is the **impact** of providing tutors with personalized feedback?
2. How does **joint feedback** to both tutors and learners impact the quality of the tutoring session?
3. How did **tutors and student perceive** the automated feedback?

Participants

~1.6k tutors, ~7k learners

Table 1: Demographics of the participant sample.

Tutors		Students	
Total number	1,609	Total number	6,572
Pronouns		***Pronouns***	
she/her	30%	she/her	55%
he/him	58%	he/him	32%
Other	12%	Other	13%
Grade level		***Grade level***	
Grade 9-10	29%	Grade 9-10	15%
Grade 11-12	67%	Grade 11-12	81%
Other	4%	Other	4%
Region		***Region***	
US	98%	US	86%
Non-US	2%	Non-US	14%

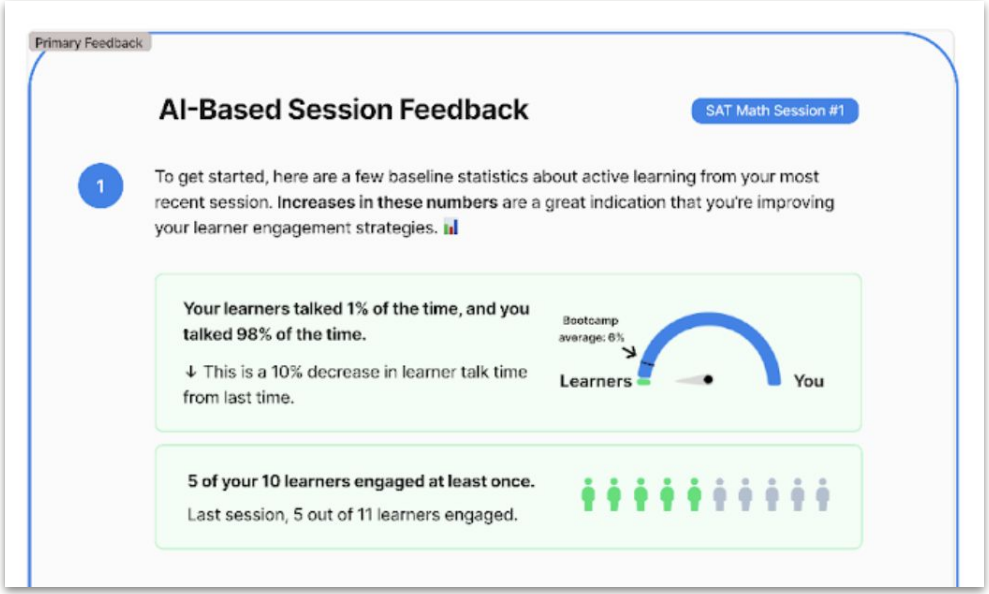
Randomized Controlled Trial

Group	Description
CONTROL	359 tutors receive no feedback
TUTOR FEEDBACK	350 tutors receive personalized feedback about their sessions
TUTOR FEEDBACK + LEARNER FEEDBACK PERSONAL	176 tutors receive personalized feedback while their learners receive socially oriented feedback about their sessions
TUTOR FEEDBACK + LEARNER FEEDBACK SOCIAL	176 tutors receive personalized feedback while their learners receive personal goal-oriented feedback about their sessions

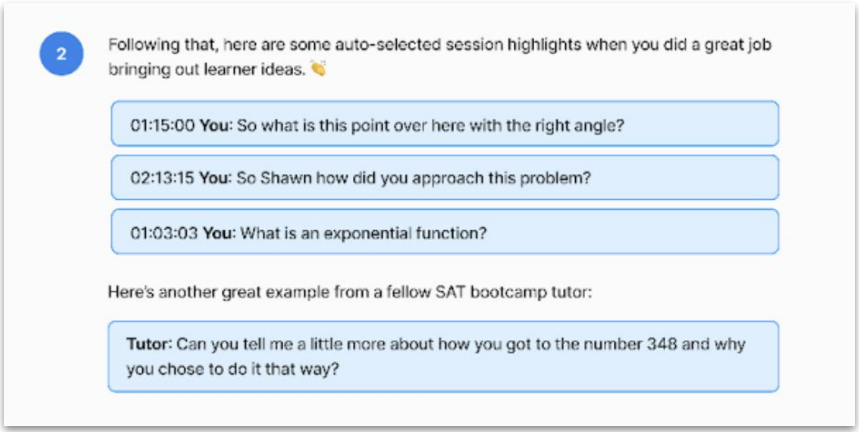
Tutor Feedback

Curriculum of Talk Moves:
Eliciting ideas → Building on Ideas → Pressing for Reasoning

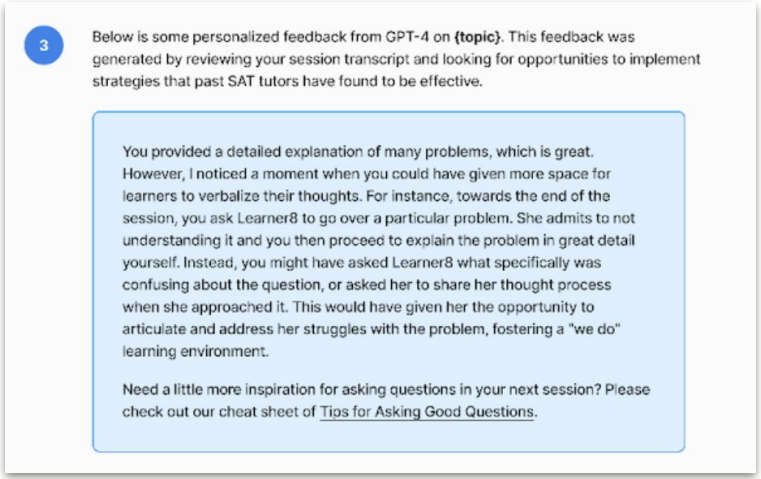
Part 1:
Stats



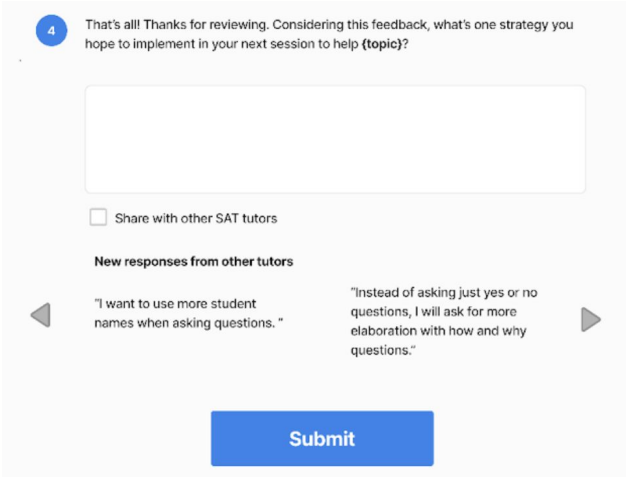
Part 2: Highlights



Part 3: GPT-based suggestions



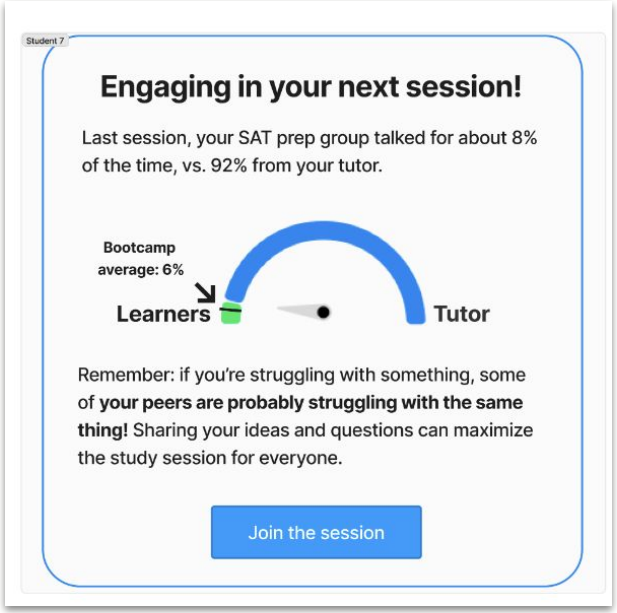
Part 4: Reflection

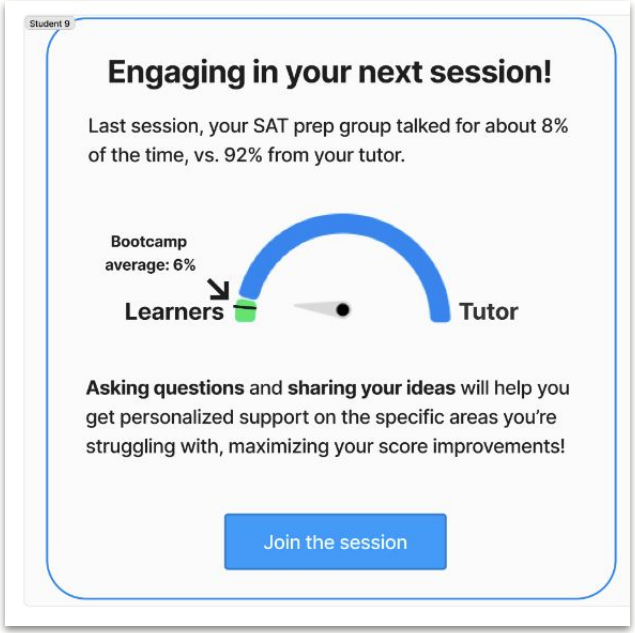


Learner Feedback

Right before they joined their session

Social vs Goal-Oriented Messaging





Findings (preliminary)

Tutor feedback improves tutors' rate of **eliciting, uptake of student ideas**

Tutor feedback improves **student reasoning**

Learner feedback, above and beyond tutor feedback improves **student NPS, students' session rating**

Tutors received the feedback very positively! (spent on average 4-5 mins)

Challenges / Negatives

- Talk time does not account for chat management
- Repetitiveness of feedback
- Technical issues / learner absences

Positives

- Liked section 1 (stats)
- Liked section 3 (GPT suggestions), although it was sometimes too long

Improvement & Usage

- Helping tutors convert passive learners into active ones!
- Make feedback more interactive

Main Affordances of Research-Industry Partnerships

- Closing the feedback loop (from users and providers back to R&D)
- Access to participants
- Real-world evaluation
- Sustainable research impact
- Scale

Mariah Olson

Co-Founder and Foundations lead, Schoolhouse.world