Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 25-1289

The Power of Personalized Attention: Comparing Pedagogical Approaches in Small Group and One-on-One Early Literacy Tutoring

Hsiaolin Hsieh
Stanford University

David Gormley
Stanford University

Carly D. Robinson
Stanford University

Susanna Loeb
Stanford University

Tutoring has played a significant role in pandemic-related learning recovery, supporting student learning and engagement. A recent randomized controlled trial estimated that one-on-one virtual early literacy tutoring was nearly twice as effective as two-on-one tutoring for improving student learning (Robinson et al., 2024). To better understand this gap, we analyze transcripts from 16,629 tutoring sessions in this RCT—which included over 3.7 million tutor utterances—using natural language processing and machine learning techniques. We explore how tutors allocate attention across content instruction, relationship building, and classroom management between one-on-one and two-on-one formats. While tutors dedicate similar time to content instruction and relationship building across both formats, students receiving one-on-one tutoring receive more attention and personalized support. To improve the effectiveness of two-on-one tutoring, it may be beneficial to equip tutors with strategies that engage multiple students simultaneously, thereby reducing downtime and minimizing the potential for disengagement.

VERSION: October 2025

Suggested citation: Hsieh, Hsiaolin, David Gormley, Carly D. Robinson, and Susanna Loeb. (2025). The Power of Personalized Attention: Comparing Pedagogical Approaches in Small Group and One-on-One Early Literacy Tutoring. (EdWorkingPaper: 25-1289). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/ykhg-zr64

The Power of Personalized Attention: Comparing Pedagogical Approaches in Small Group and One-on-One Early Literacy Tutoring

Hsiaolin Hsieh, David Gormley, Carly D. Robinson, Susanna Loeb

Abstract

Tutoring has played a significant role in pandemic-related learning recovery, supporting student learning and engagement. A recent randomized controlled trial estimated that one-on-one virtual early literacy tutoring was nearly twice as effective as two-on-one tutoring for improving student learning (Robinson et al., 2024). To better understand this gap, we analyze transcripts from 16,629 tutoring sessions in this RCT—which included over 3.7 million tutor utterances—using natural language processing and machine learning techniques. We explore how tutors allocate attention across content instruction, relationship building, and classroom management between one-on-one and two-on-one formats. While tutors dedicate similar time to content instruction and relationship building across both formats, students receiving one-on-one tutoring receive more attention and personalized support. To improve the effectiveness of two-on-one tutoring, it may be beneficial to equip tutors with strategies that engage multiple students simultaneously, thereby reducing downtime and minimizing the potential for disengagement.

Introduction

A recent RCT focusing on the efficacy of a virtual early literacy tutoring program for students from kindergarten to second grade compared one-on-one to two-on-one virtual tutoring. In keeping with the prior literature, the study found that one-on-one tutoring produced larger gains in students' literacy achievement, with the estimated effect size of one-on-one tutoring approximately twice that of the two-on-one tutoring, though not statistically different (Robinson et al., 2024). The virtual tutoring environment, with the video and audio that it produces, provides an unprecedented opportunity to study the details of tutoring sessions. By analyzing tutor pedagogy, we can better understand why one-on-one tutoring tends to be more effective than two-on-one tutoring, and, potentially, how to improve two-on-one approaches so that schools can leverage the value of tutoring at a lower cost.

Advancements in computational methods have made it possible to transcribe audio recordings and analyze those transcriptions at scale, offering insights into human interactions and behaviors. By employing techniques in natural language processing (NLP), machine learning (ML), and large language models (LLMs), researchers have created automated discourse models and analytical tools to identify teacher and tutor talk moves (e.g., Demszky et al., 2021). They also investigate how these interactions influence student engagement and learning outcomes (Abdelshiheed et al., 2024; Booth et al., 2024; Jacobs et al., 2022; Liu et al., 2024; O'Connor & Michaels, 2019).

This study leverages these computational methods and audio-recordings of tutoring from the RCT of one-on-one versus two-on-one tutoring to uncover how tutoring differs across these settings. It asks: *How does tutor pedagogy differ between two-on-one and one-on-one formats in ways that might explain their differential effectiveness?* We find that students in one-on-one sessions receive more time focused on personalized content and relationship building than those in two-on-one sessions and that tutors employ different cultural, linguistic, and relational approaches in the two contexts to enhance content learning and foster relationships.

Background

High-impact tutoring—tutoring with a consistent tutor, using data and high-quality instructional materials, over a substantial number of hours—has emerged as one of the most effective interventions for accelerating student learning (Nickow et al., 2024; Robinson & Loeb, 2021). Research consistently shows that incorporating tutoring interventions into the school day leads to substantial positive effects on student learning, with especially strong impacts on early literacy (Dietrichson et al., 2017; Nickow et al., 2024). These interventions can also strengthen student engagement and improve attendance (Lee et al., 2024), perhaps because they facilitate a consistent connection with a dedicated tutor who can cultivate a sense of belonging and accountability (Guryan et al., 2021). Recent research also provides evidence of positive effects for students affected by school closures and learning disruptions from the COVID-19 pandemic (Cortes et al., 2025; Lee et al., 2024).

Despite its promising outcomes, tutoring is available to less than half of the target student population (Center on Reinventing Public Education [CRPE], 2024). Many students who are eligible and can benefit from learning recovery interventions are not receiving the services (CRPE, 2023).

Recruiting tutors is a key challenge in scaling tutoring (Groom-Thomas et al., 2023). While effective tutors can come from a variety of backgrounds (Nickow et al., 2024), finding enough tutors in the local community to support all struggling students can be difficult, especially in rural communities and areas with tight labor markets (Groom-Thomas et al., 2023). One solution is virtual tutoring, where the students are in school but the tutor is not and the tutoring session takes place online. Virtual tutoring expands the labor market beyond geographic boundaries, which allows schools to provide tutoring even if there are not enough tutors in the local area. Virtual tutoring also addresses the cost of commuting, which can be substantial relative to the compensation. Many school districts invested in virtual tutoring in response to the pandemic, and recent research shows that the virtual approach can be effective for improving student learning (Carlana & La Ferrera, 2021; Gortazar et al., 2023; Robinson et al., 2024; Hashim et al., 2025).

Funding is another key challenge to tutoring uptake: effective tutoring programs require many tutors and experienced educators to oversee and coach tutors (Groom-Thomas et al., 2023). To manage the costs of high-impact tutoring, which often range from $1,200 to $2,500 per student, education agencies often increase the student-tutor ratio to two or more students per tutor in a session. While larger student-tutor ratios can reduce costs and may facilitate peer (student-student) learning, larger ratios are generally associated with lesser learning gains (Nickow et al., 2024). The differential effectiveness between one-on-one and two-on-one tutoring motivates our analysis of the differences between one-on-one and two-on-one tutoring sessions.

Tutoring improves student learning outcomes by addressing two major challenges of classroom teaching: the diverse academic levels and needs of students and the inability to provide students with the attention and human relationship they need for motivation and engagement (Guryan et al. 2023; Roorda et al., 2017, Scales et al., 2020). The key question—and the focus of this study—is how the dynamics of tutoring differ between one-on-one and two-on-one formats across the two critical ingredients of tutoring: personalization of academic content and relationship building. One-on-one tutoring might be more impactful than two-on-one tutoring because it allows for more personalized instruction, where the tutor focuses the sessions on the material that can most benefit the student. Alternatively or in addition, one-on-one tutoring may be more effective because the tutor develops a stronger relationship with the student, engaging them and motivating their learning. We explore both the quantity of time tutors spend on instruction and relationship building in one-on-one and two-on-one sessions and the quality of interactions during those times. By examining these differences in detail, we can better understand the mechanisms that might explain variations in student learning outcomes and inform best instructional practices for small-group tutoring.

Quantity: Quantity of instructional time affects learning (Gromada and Shewbridge, 2016; Holland et al., 2015; Patall et al., 2010; Yeşil Dağlı, 2019) and may differ between one-on-one and two-on-one sessions. Researchers have found that students spend 40 percent more time on-task in the one-on-one environment than in whole-class instruction (Bloom, 1984).

Tutor attention mediates the value of instructional time. Educator attention is a limited resource (Kahneman, 1973). In a one-on-one session, tutors may direct their undivided attention to their sole student, while in two-on-one sessions tutors may split their attention between their students (Zhang et al., 2025). Thus students in two-on-one sessions experience three distinct types of instructional time: instruction directed at both them and the other student, instruction directed at only them, and instruction directed only at the other student. Personalized instruction directed at one student in a two-on-one session may be approximately equivalent to that in a one-on-one session, though the presence of the other student may introduce social dynamics, such as collaboration or competition, that may influence learning (Zajonc, 1965, Johnson & Johnson, 2009). However, the value of instruction directed at the other student or at both students together is unclear.

Quantity of time spent on relationship building also may affect learning and differ between one-on-one and two-on-one sessions. A meta-analysis finds that positive student-teacher relationships are associated with higher engagement and increased student achievement (Roorda et al., 2017). High-impact tutoring emphasizes consistent tutor-student pairings because this consistency provides more time to build stronger tutor-student relationships (Robinson et al., 2021). Again, the scarcity of tutor attention means that two-on-one students are likely to get less direct relationship-building time with their tutor than one-on-one students.

Differences in quantity between one-on-one and two-on-one tutoring may also stem from differences in the time spent on disruptions and from logistical or structural components of tutoring, such as the number and length of sessions and the pace of instruction (i.e., the tutor speaking speed). On the one hand, students in one-on-one tutoring may benefit from more frequent sessions because scheduling is simpler with just two people involved. On the other hand, students in two-on-one tutoring might experience longer sessions and a faster tutor speaking pace, as the tutor works to address the needs of multiple students. Examining the number and length of sessions, tutor talk time, tutor-student pairing consistency, and tutor speaking speed could allow us to understand if students in two-on-one and one-on-one sessions receive different quantities of instruction.

Quality: Tutoring sessions likely vary in the quality, as well as the quantity, of interactions, whether instructional or relationship-building interactions. For example, one-on-one sessions may allow tutors to personalize instruction and relationship building more than they are able to

do in two-on-one sessions. We use personalization as a key indicator of the quality of targeted instruction and relationship building because both types of interactions likely hinge on their level of personalization.

According to Vygotsky, the quality of instruction is closely tied to how well it is personalized. He advanced the concept of the Zone of Proximal Development (ZPD) for which personalized instruction should target just above a student's current ability level (Vygotsky, 1978). In one-on-one tutoring, tutors can identify their student's ZPD and provide the necessary scaffolding to support learning within that range (Vygotsky, 1978; Wood, Bruner, & Ross, 1976). In two-on-one tutoring, tutors must simultaneously manage two distinct ZPDs. If the students have differing abilities, the tutor may encounter cognitive overload, where external demands (different student needs) and internal processing (individual scaffolding strategies) exceed their available attentional resources (Sweller, 1988). This overload can reduce a tutor's effectiveness in teaching (Feldon, 2007).

Similarly, one-on-one tutoring may be better suited for developing highly-personalized student-teacher relationships because the tutor can focus entirely on understanding and connecting with a single student. In contrast, two-on-one tutoring requires the tutor to cultivate two relationships at once, which may result in less in-depth knowledge of each individual student and weaker personalization (Ichii, 2022).

Research Questions: Ultimately, one-on-one sessions may deliver a higher quantity of instruction and relationship building by offering the tutor's undivided attention and a higher quality through more personalized content and interactions. This study is the first to compare the quantity and quality of instruction and relationship building in one-on-one versus two-on-one tutoring, using a large set of tutoring session transcripts from an RCT. Specifically, we address the following research questions:

- RQ1: How do the *structural features of tutoring* vary between one-on-one and two-on-one tutoring? In particular, how does tutoring differ in the number and length of sessions, tutor talk time, tutor-student pairing consistency, and tutor speaking speed?
- RQ2: How does *time allocation* vary between one-on-one and two-on-one tutoring? In particular, how much talk time do tutors spend on instruction, relationship building, and classroom-management, and how do tutors in two-on-one sessions allocate their attention between their students for these topics?
- RQ3: How does the *content* of tutors' instructional time and relationship-building time vary between one-on-one and two-on-one tutoring? In particular, do tutors in one-on-one sessions provide more individualized instruction and more individualized interactions during relationship-building?

Program and Study Context

OnYourMark Education (OYM) is a virtual early literacy tutoring program with curriculum grounded in the Science of Reading, with a focus on phonics, phonological awareness, and fluency. OYM tutoring is delivered in one-on-one or two-on-one sessions embedded into the school day. The program aims to promote positive tutor-student relationships through a small student-tutor ratio and by pairing students with a consistent tutor.

During the 2022-23 school year, Uplift Education, a charter management organization in Texas partnered with OYM to provide early literacy tutoring to 2,206 kindergarteners, first graders, and second graders in 12 of its elementary schools as part of a RCT to evaluate the effectiveness of the tutoring program. Students eligible to receive tutoring at these schools were randomized into one of three groups: control (business-as-usual without tutoring), individualized tutoring (1 student : 1 tutor), or tutoring in pairs (2 students : 1 tutor). Students assigned to receive tutoring met their tutor online for 20 minutes during the school day, four times per week. Tutoring rolled out in September and continued through May.

The study was conducted across 12 schools using a stratified randomization design. To prioritize equity and address school staff concerns about randomization, approximately ten high-need students per school were guaranteed tutoring slots and excluded from the analytic sample. After reserving these seats, school leaders identified additional eligible students based primarily on beginning-of-year Dynamic Indicators of Basic Early Literacy Skills (DIBELS) assessments, with final selections left to local discretion. Eligible students were grouped into pairs based on school, scheduling availability, and overlapping areas of literacy need, and 2,085 students were ultimately randomized within 34 school-by-grade strata. Student pairs were assigned as clusters to either treatment or control conditions, ensuring balance across grade levels and initial performance. Treatment students were further randomized to receive either one-on-one or two-on-one tutoring, and then assigned to tutors at random. This process resulted in 510 students in one-on-one tutoring, 570 students in two-on-one tutoring, and 1,005 students in the control condition (see Robinson et al., 2024 for more details on the study design). The present study focuses on a subset of the 1,080 treatment students, whose tutoring sessions were recorded and analyzed.

Data

This study uses the session metadata and video recordings from February to May 2023, covering approximately ten weeks of the OYM RCT. During this time, OYM recorded a total of 23,825 tutoring sessions. A typical OYM tutoring session is scheduled for 20 minutes; however, the actual duration varied based on when the tutor and student(s) logged onto the platform and began the lesson. Some sessions were cut short due to technical issues or attendance problems, while others were extended to meet instructional needs. Consequently, the dataset includes session recordings of varying lengths, ranging from 1 to 40 minutes. To focus on the most common tutor-student interactions, we excluded 15% of the shortest and longest sessions for each group

size from the analysis. As a result, the sample for analysis includes 16,629 session recordings: 10,474 from students assigned to one-on-one settings and 6,155 from students assigned to two-on-one settings. Due to logistical difficulties, 27.10% of two-on-one sessions were conducted with just one student and 1.02% of one-on-one sessions were conducted with two students. Of these sessions, 4,451 took place in kindergarten, 6,377 in first grade, and 5,646 in second grade, while the remaining 155 sessions involved students from different grades.

To explore our outcome of interest, tutor pedagogy as demonstrated by their language, we utilized automatic speech recognition (specifically WhisperX) to transcribe audio recordings. For this study, we focus exclusively on the transcripts of tutors because the students' transcripts are not reliable due to persistent background noise in the tutoring venues (e.g., classrooms, hallways, or cafeterias). The mean word error rate (WER) for tutor utterances, calculated from a random sample of 12 transcripts stratified by grade and group size, is 16.24% (SD=4.10), indicating 83.76% of the transcription is correct. This rate is within the acceptable performance range for automatic speech recognition, especially given the wide range of WERs of WhisperX (3-60% depending on the quality and difficulty of the datasets; Bain et al., 2023; Ferraro et al., 2023; Radford et al., 2022). Our WER was driven by substitutions (9.28%), followed by deletions (4.52% error rate), and insertions (2.44% error rate). Our sample consists of 3,754,469 tutor utterances in total. On average, tutors spoke 224 utterances (1,277 words) per session. We constructed our outcome measures from these transcripts, as detailed below in the description of the methods used to address each research question.

We link these transcriptions to student-level and tutor-level data from OYM and the school district's administrative data. The student-level data includes grade, date of birth, race/ethnicity, gender, whether the student received free or reduced price lunch or were otherwise indicated as economically disadvantaged based on the receipt of other public assistance, whether the student had an Individualized Education Plan or 504 Plan, whether the student was designated as an English learner, and the student's availability for tutoring within the school day. The student-level data also includes student beginning, middle and end of school year achievement scores from DIBELS. For tutor-level data, OYM provided administrative information including the name of the tutor, tutor demographics, and tutoring attendance.

Methods

The goal of this study is to explore how tutor pedagogy differs across one-on-one and two-on-one sessions in ways that may explain their varying effectiveness. We employ text analysis techniques to annotate and analyze the transcripts of tutoring sessions based on tutor recordings.

In this paper we report the results of bivariate comparisons. Because students and tutors were randomized into one-on-one or two-on-one tutoring, we do not need to control for potentially confounding factors when comparing the session types. We do run analyses controlling for all available pre-treatment measures (see Appendix), but the results are so similar that we do not report them in the main text.

RQ1: Examining the Structural Features of Tutoring

Our first research question asks: how do the structural features of tutoring vary between one-on-one and two-on-one tutoring? Structural features define the frequency and intensity of the tutoring. We investigate the structural features of the tutoring sessions, focusing on how one-on-one and two-on-one sessions differ in the following aspects:

- number of sessions per week
- session length
- percent of sessions with the same tutor
- tutor talk time
- tutor speaking rate (speed)

To address this research question, we calculate the frequency and duration of each tutoring session and tutor utterance using timestamps from the platform's metadata. We define a tutor utterance as a single instance of tutor speech during which tutors are speaking without interruption. We determine the number of attended sessions per week at the student level based on their assigned conditions. We measure tutor consistency by the percentage of sessions student(s) met with the same tutor. We define session length as the overlapping period when both the tutor and at least one of the students are present. We utilize NLP to count the number of words in each tutor utterance, aggregating these counts at the session level. We measure utterance duration using the timestamps. By combining word counts with duration, we estimate the tutors' speaking speed in words per minute at the session level. This metric provides insight into the pacing of tutor speech during the sessions.

RQ2: Examining Time Use During Tutoring Sessions Through Text Classification

Our second research question asks: how does time allocation and tutor attention differ between one-on-one and two-on-one tutoring for the following interaction types? We distinguish three types of time use:

- Content instruction
- Relationship building
- Classroom management

To address this question, we develop text classification models to label tutor utterances. We utilize text analysis techniques based on machine learning and pre-trained contextual embeddings (BERT and RoBERTa) and employ expert annotations to fine-tune a public transformer model on Hugging Face. We train three separate text classifiers to label utterances according to content instruction, relationship building, and classroom management. Note that we borrow the term "classroom management" from the whole class setting to refer to non-instructional tasks related to maintaining order by addressing behavioral or tech issues. We also develop an additional text classifier to determine how the tutor allocates their attention in two-on-one sessions between students. By combining the labeled utterances with their timestamps

from session log files, we analyze the tutor talk time and the percentage of tutor utterances for different types of tutor interactions in each tutoring session.

Content Instruction Classifier: The tutoring program uses a curriculum based on the Science of Reading for young learners. We annotate utterances related to reading skills—such as sounding out words, blending or segmenting syllables, and identifying rhymes—as content instruction. Utterances that do not pertain to content instruction mainly include greetings, small talk, playing games, and disruptions (e.g., troubleshooting technical issues or managing behavior). This binary text classifier (content instruction vs. non-content instruction) achieves an accuracy of 84 percent.

Relationship-Building Classifier: Building strong teacher-student relationships can be achieved through a variety of methods. We fine-tuned this classifier to identify tutor utterances that exemplified four categories of relationship-building practices: providing motivational praise for student learning, making personal connections with students, showing care and affection, and engaging in enjoyable activities together (Breiseth, 2020; National Student Support Accelerator [NSSA], n.d.), as well as a null category representing non-relationship building. Table 1 provides sample tutor utterances for each category. This multi-class relationship-building classifier achieved a micro-average F1 score of 0.87 (macro-F1 = 0.66), with better performance in classifying utterances that establish personal connections (F1-score = 0.80) or demonstrate care and affection (F1-score = 0.68).

Classroom-Management Classifier: This classifier consists of three categories: behavioral classroom management, technical classroom management, and non-classroom management. Table 1 provides sample tutor utterances for each category. The classroom management classifier achieved a micro-average F1 score of 0.95 (macro-F1 = 0.83), with an F1-score of 0.86 for identifying technical challenges and 0.67 for identifying behavioral issues.

Attention Classifier: One complication in comparing time allocation for two-on-one and one-on-one sessions arises because the tutor in a two-on-one session is choosing not only what to focus time on but also how to split their attention between their two students. In the two-on-one setting, students experience three forms of attention: direct attention, observation of their peers' direct attention, and shared attention targeted at both them and their peer. To account for this variation, we employed an Attention Classifier (see Zhang et al., 2025 for more information) to categorize utterances into four groups: direct instruction for "Student A," instruction directed toward a peer "Student B," instruction aimed at both students ("Both"), or instruction aimed at only one student, with ambiguity regarding the recipient ("Only One" or "NA"). By cross-tabulating these attention labels with the three other classifiers (content instruction, relationship building, and classroom management), we can investigate how tutors allocate their attention between students at various points during the session.

RQ3: Examining the Details of Personalized Instruction and Relationship Building

Our third research question asks: how does the content of tutors' instructional time and relationship-building time vary between one-on-one and two-on-one tutoring? To answer this question, we qualitatively examine how tutors facilitate content instruction and build relationships, as these are two key components of tutoring that contribute to student learning and engagement (Neitzel et al., 2022; Nickow et al., 2024). Specifically, we investigate whether tutors provide more personalized content instruction and relationship building in one-on-one tutoring sessions compared to two-on-one sessions.

For content instruction, we utilize the utterance labels from our second research question. We apply BERTopic, a topic modeling algorithm based on pre-trained BERT contextual embeddings, to cluster and summarize ten topics (i.e. themes) and compare the frequency of these topics across formats. For relationship building, we use our multi-class relationship building classifier from research question two to explore the frequency of relationship-building subcategories according to group size. For the topics and subcategories that show significant differences between the two group sizes, we conduct a deeper qualitative analysis of the transcripts, focusing on the language used by tutors.

Results

RQ1 Finding: Similar Structural Features of Tutoring

Overall, tutoring sessions in one-on-one and two-on-one formats exhibit similar distributions of structural features, as shown in Figure 1. We examined the mean differences between these formats by analyzing each feature separately in the following sections.

Number of Sessions Per Week

During the ten-week period from February to May 2023, students assigned to one-on-one tutoring attended an average of 25.64 sessions (SD=6.93), while those in two-on-one tutoring attended an average of 25.46 sessions (SD=6.93). The difference between the two groups is not statistically significant (p=.65). Students in both settings attended an average of 2.65 sessions per calendar week, with no statistically significant difference (p=.81).

Percentage of Sessions with the Same Tutor

Students consistently received tutoring from the same tutor, regardless of group size. In one-on-one tutoring, students met with the same tutor an average of 88.74 percent of the time (SD=16.07), while those in two-on-one tutoring had a similar average of 88.60 percent (SD=15.79). The difference between the two groups is not statistically significant (p =.88).

Session Length and Tutor Talk Time

The average length of a tutoring session was slightly longer in the two-on-one format (M=19.73 minutes, SD=1.38) than in the one-on-one format (M=18.84 minutes, SD=1.67). This difference

is statistically significant (p<.001). Tutors also spoke for a slightly longer duration in two-on-one sessions (M=9.04 minutes, SD=2.16) than in one-on-one sessions (M=8.27 minutes, SD=2.10), again showing a statistically significant difference (p<.001). These variations may be due to the logistical challenges of the two-on-one format, such as waiting for the second student to join before starting the lesson.

Tutor Speaking Speed

Tutors spoke slightly faster in two-on-one sessions than in one-on-one sessions. Tutors' average speaking speed was 149.73 words per minute (WPM; SD=25.62), on average, with two-on-one sessions at 151.10 WPM (SD=23.98) compared to one-on-one at 148.94 WPM (SD=26.50), p<.001. While statistically different, these average rates are qualitatively quite similar, given that average talking speeds are 140-160 words per minute. Another way to conceptualize this is that a two WPM difference is only about 40 extra words over the course of a 20 minute session, where an average of 1,277 total words are spoken.

RQ2 Finding: Varying Time Use and Substantial Difference in Tutor Attention

Time Use across Different Interaction Types

Figure 2 displays the distribution of tutor talk time (in minutes) by interaction type, including content instruction, relationship building, and classroom management. In a typical tutoring session, without disaggregating by group size, tutors spent approximately six minutes of talk time on content instruction and close to one minute each on relationship building and classroom management. Relative to one-on-one sessions, tutors in two-on-one sessions spent an average of 0.53 minutes more on content instruction (p<.001), 0.02 minutes more on relationship building (p< .01), and 0.31 minutes more on classroom management (p<.001). Tutors spent approximately 31.55 seconds more on content instruction during two-on-one sessions, while also devoting roughly 18.50 seconds more to classroom management. In terms of percentages, time spent in two-on-one sessions is roughly 9.81 percent greater for content instruction and 63.33 percent greater for classroom management compared to one-on-one sessions. These percentages are large, especially for classroom management, but the magnitudes for the amount of time spent on management are quite small for both groups.

Allocation of Tutor Attention in Two-on-One Sessions

Within the time use described above, tutors allocate their attention differently in the two types of tutoring. In contrast to one-on-one sessions, where tutors dedicate their full attention to a single student, tutors in two-on-one sessions divide their attention between two students. The single student in one-on-one tutoring receives the entire focus of the tutor, while the two students in two-on-one tutoring share that attention.

Figure 3 shows the percentages of utterances that tutors directed to each student by interaction type across group sizes. In one-on-one sessions, 65.20% of utterances were related to content instruction. In two-on-one sessions, tutors dedicated an average of 65.66% of utterances to content instruction, shared between two students. In these two-on-one sessions, an average of 23.13% of utterances were content instruction directed to both students, 13.26% to one student ("Student A"), 14.64% to the other student ("Student B"), and 14.63% ambiguous (either to Student A or B). Assuming equal distribution of ambiguous utterances, each student receives 21.27% individualized content instruction (13.95% is the average of utterances classified as directed at student A and B plus 7.32% from half of the ambiguous utterances). In contrast, a student in a one-on-one session receives direct individualized content instruction in 65.20% of utterances.

A student in a two-on-one session also receives content instruction in other non-individualized ways: these students receive content instruction directed at both them and their peer in 23.13% of utterances and watch their peer receive direct instruction in 21.27% of utterances. However, instruction directed at two students, rather than individualized for one, may be less effective than individualized instruction given that tutoring relies on personalization. The value of watching instruction directed to a peer is likely even smaller because the instruction may not even be relevant to the student who is watching.

Tutor attention is also split between the two students during two-on-one tutoring for interactions related to relationship building (Figure 3). Again assuming equal distribution of ambiguous utterances, students in two-on-one sessions receive nearly three times less relationship building directed only at them than those in one-on-one tutoring. To be clear, they also receive relationship building directed at them and their peer, but it may be less personalized. Even if we assume that this type of interaction is equivalent to relationship building directed at only them, they still receive only two thirds of the amount of relationship building as one-on-one students. Note that we are not considering relationship building utterances directed only at the other student because we believe it is unlikely that exposure to the tutor building a relationship with the other student is beneficial to the first student.

Simply put, students in two-on-one sessions receive substantially less personalized content instruction and relationship building. Not only are large portions of the session directed at both them and their peer, and thus are less personalized, but substantial parts of the session are directed only to their peer, which may not be relevant to them at all.

RQ3 Finding: Details of Personalized Instruction and Relationship Building

Through topic modeling and our multi-class relationship building classifier, we find that one-on-one and two-on-one tutors employ different cultural, linguistic, and relational approaches to enhance content learning and foster relationships.

Individualized Content Instruction

Figure 4 illustrates the clusters of topics derived from BERTopic modeling, summarizing tutors' utterances during content instruction. Over 75 percent of the content-instruction utterances consist of brief, spontaneous positive feedback (Topic 1), including phrases like "good job," "great," "well done," "excellent," "nice," "perfect," etc. Other utterances pertain to specific learning tasks, such as phonics, blending sounds, punctuation, comprehension, and tests.

The percentage differences across tutoring formats for topics are small, except for Topic 3 (referring to a specific student) and Topic 7 (in Spanish). The difference in Topic 3 (one-on-one sessions having lower occurrences compared to two-on-one) aligned with the turn-taking and attention-splitting nature of two-on-one tutoring, as demonstrated by the attention classification analysis. Topic 7 highlights that tutors utilize their shared cultural and linguistic resources, especially Spanish, substantially more frequently during one-on-one sessions. The percentage difference for Topic 7 (0.13 percent of all utterances in one-on-one sessions versus 0.02 percent in two-on-one sessions) translates to 0.17 and 0.03 utterances per session in one-on-one and two-on-one sessions, respectively, $p<.001$. In the sample, 33.67% of students are multilingual learners and 69.18% are Hispanic.

Table 2 provides examples of tutor utterances in Spanish, showcasing how tutors leverage Spanish as a valuable resource for both content instruction and relationship building. This increased use of linguistic resources could happen for two reasons linked to personalization. First, a two person conversation is generally more individualized than a three person conversation (e.g. all three members of a three person conversation must speak Spanish or else the use of Spanish will exclude someone from the conversation). Second, one-on-one tutors can get to know their student better than two-on-one tutors can (one-on-one students receive a larger percentage of direct relationship building utterances than two-on-one students), and, as a result, one-on-one tutors may be more likely to learn important cultural and linguistic information about their student, such as whether they speak Spanish.

Figure 5 depicts the use of Spanish by tutors throughout the program. The x-axis represents the weeks of the tutoring sessions, while the y-axis summarizes the total number of sessions. The increased use of Spanish towards the end of the program coincided with tutors selecting the nursery rhyme "See You Later, Alligator," which includes the line "Mañana, Iguana." Different colors indicate different tutors (approximately 60 in total). The consistent appearance of the same colors suggests that certain tutors regularly employ Spanish during their sessions, while others do not. This consistent use indicates that those who were familiar with these linguistic resources applied them consistently, potentially fostering a welcoming environment for multilingual learners.

Personalized Relationship Building

Next we explore the frequency, in percentage of utterances, of relationship building subcategories (Figure 6). We find that tutors provide more motivational praise (mean difference (MD)=0.16, p<.001), make more personal connections (MD=0.23, p<.001), and spend a larger percentage of utterances on having fun with students in one-on-one sessions compared to two-on-one sessions (MD=0.15, p<.001). In contrast, tutors express similar levels of care or affection in both session types (MD=0.08; p=.10).

The differences in relationship building may stem from the social nature of one-on-one versus two-on-one tutoring sessions. Two person conversations allow for more individualized interactions between the tutor and the one student, while three person conversations require a focus on communal interests between the tutor and their two students.

Tic-tac-toe is a popular game in OYM tutoring. Tutors mentioned "tic-tac-toe" in 25.18 percent of one-on-one sessions and 21.29 percent of two-on-one sessions (p<.001). OYM encouraged tutors to play tic-tac-toe as a means for engaging students because the game is simple to learn and quick to run, fitting well within the 20-minute session duration. Tutors used tic-tac-toe both as an instructional tool and as a form of positive reinforcement (reward). For example, in instructional contexts, tutors might say, "We're playing tic-tac-toe, but you need to say the word correctly to place your mark," or "Now, it's time for a tic-tac-toe game, but we'll use sight words." As part of positive reinforcement, tutors might use language such as, "If we finish all of these sentences, do you want to play tic-tac-toe?" or "If we have enough time, we will definitely play tic-tac-toe." The game typically occurred at the end of the session but could also be played at the beginning as a make-up reward if time ran out in a previous session.

Tutors brought up "tic-tac-toe" more frequently during one-on-one sessions, and they also played the game differently in those sessions. Table 3 contains excerpts from tutor utterances related to the game in both two-on-one and one-on-one settings. Tic-tac-toe is designed for two players. In two-on-one tutoring sessions, students typically play against each other ("Okay, let's do a speed run on tic-tac-toe. You guys have to take turns, okay?"); if one student does not want to play, the other misses out ("I don't blame Student A for not wanting to play tic-tac-toe. He and I played a lot last semester."). In contrast, during one-on-one sessions, the student has the chance to play against the tutor, receiving undivided attention from an educator.

The dynamics of the interactions may depend on whether the engagement is between students or between an adult and a student. Student-student interactions might create more tension or competition, while adult-student interactions tend to encourage trust and engagement. The caring adult can guide the student with strategies for winning (e.g., "No, you don't want to go there, do you? You should go somewhere else.") and boost the student's confidence with compliments ("It's hard to beat you." "How did you do that? You beat me!").

Discussion

This study analyzes transcription data to examine the differences in tutor pedagogy between two-on-one and one-on-one tutoring sessions and how these differences may account for the gap in effectiveness between one-on-one and two-on-one tutoring, as reported by Robinson and colleagues (2024). Our transcripts come from a highly-structured early literacy tutoring program held virtually during the school day. Tutors and students were randomly assigned to the one-on-one or two-on-one setting, and, as a result, our comparisons are not confounded by selection bias and are unlikely to be confounded by missing variable bias.

We find that structural features, such as session frequency and tutor speaking speed, are similar between two-on-one and one-on-one settings. Our analysis also finds that the amount of time spent on content instruction and relationship building are similar across settings, though more time is spent on classroom management in two-on-one sessions. However, the attention classifier that we construct reveals that students in two-on-one sessions receive substantially less personalized instruction and relationship building. Tutors in two-on-one sessions spend large portions of the session addressing both students at once or directly addressing the other student in the session. In analysis of the details of personalization, we find that tutors in one-on-one sessions were more likely to speak Spanish with their students and have a larger portion of utterances dedicated to the relationship building subcategories of motivational praise, personal connections, and having fun together. While the magnitude of these differences is small, over the course of an academic year, these differences add up.

The results indicate that the differing learning outcomes may stem from variations and constraints in how tutors deliver content and build rapport. Specifically, students assigned to two-on-one tutoring receive far less individualized attention than students assigned to one-on-one sessions in both content instruction and relationship building. Students in two-on-one sessions also have to tolerate additional distractions from the extra time spent on classroom management, which could lead to disengagement.

Peer-learning is one of the potential strengths of small-group tutoring relative to one-on-one tutoring, yet, peer learning does not happen organically; it requires intentional facilitation of meaningful interactions, especially in virtual learning environments. In this study, students typically attend two-on-one tutoring sessions individually using their own computers and headphones. Student pairs occasionally are in the same room, but they are often in separate spaces. Regardless of their physical proximity, students often struggle to connect with peers through a computer screen without external prompting (Vrieling-Teunter et al., 2022). The early literacy curriculum further complicates matters, as it relies heavily on the deliberate practice of lower-cognitive skills, such as recall and reproduction. Mastery of these foundational concepts requires one-on-one attention from tutors, as opposed to cooperative discussions that are more effective for higher-order problem-solving in subjects like mathematics or reading comprehension.

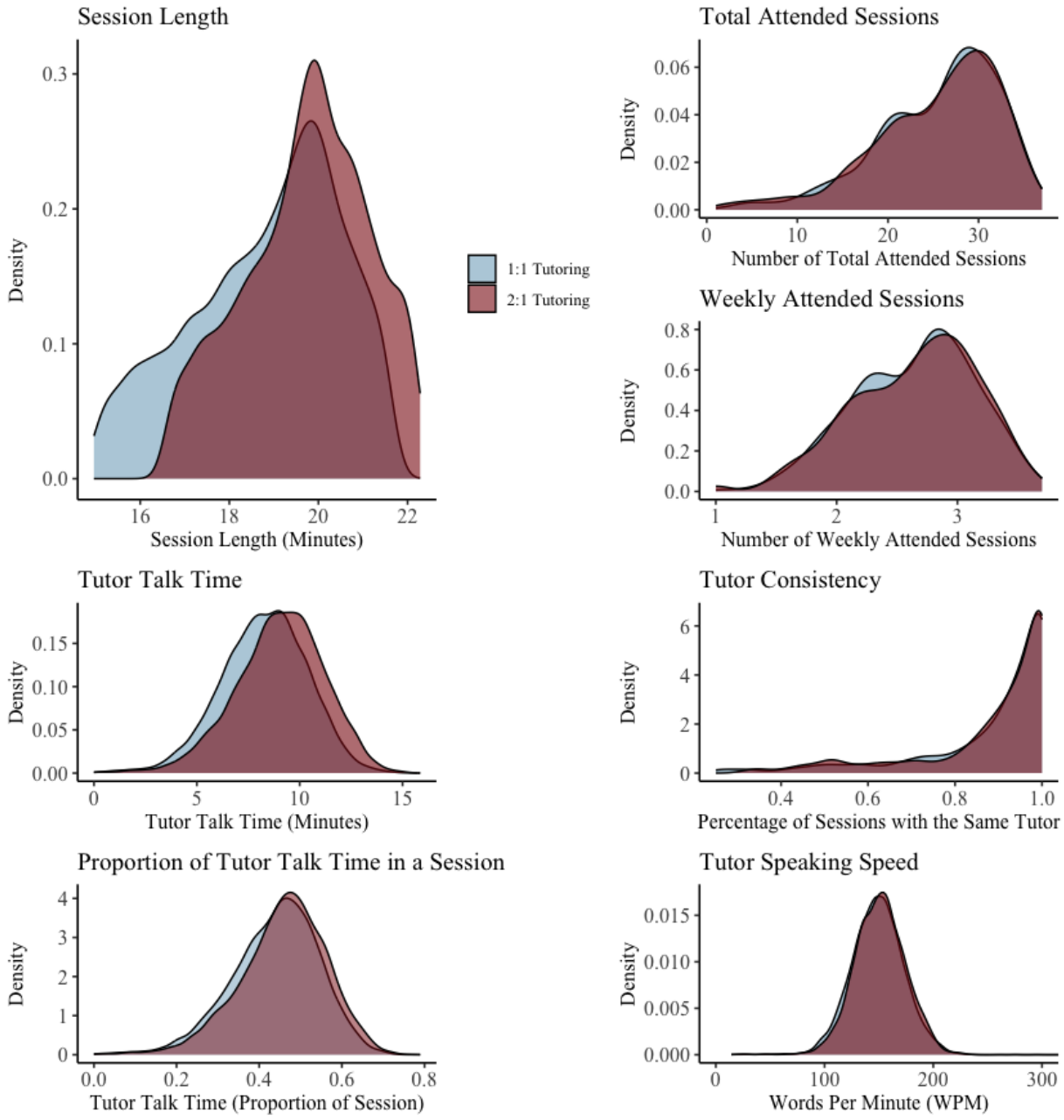
Given the existing funding shortages and associated cost implications, two-on-one or small-group tutoring will likely become more prevalent, underscoring the need for understanding how to make these settings as effective as they can be. This study shows that the approaches - such as playing tic-tac-toe for rapport-building - that work effectively in the one-on-one setting may not work as well in the two-on-one setting. Tutoring in the two-on-one setting may benefit from explicitly leveraging peer learning techniques and from using curricular materials that incorporate collaborative teamwork to reduce individual students' downtime. Using games that everyone in the session, including tutors and students, can participate in may also better foster teamwork and a sense of belonging. Some providers already enhance their tutoring sessions with well-designed, high-quality computer-assisted learning tools in small group contexts, which have demonstrated positive outcomes in both student learning and cost-effectiveness (Bhatt et al., 2024; Cortes et al., 2025; Slavin et al., 2011). Well-organized professional development for tutors may also improve two-on-one tutoring, especially in helping novice tutors develop their engagement skills. Recent advancements in generative artificial intelligence have created opportunities to provide coaching to tutors in real time at low cost (Agarwal et al., 2024; Demszky et al., 2023; Kim, 2024; L'Enfant, 2024; Wang et al., 2024).

This study contributes to the field by examining how technology and pedagogy intersect in tutoring. By utilizing automatic discourse analytics to assess tutor attention and tutor time allocation, this study provides insights into effective engagement strategies for small group tutoring. In the near future, integrating AI-powered feedback with these analytics could create a comprehensive framework for professional development, empowering tutors to refine their skills in real-time. This integration, in turn, may lead to improved educational outcomes for students. The findings of this study highlight the importance of data-driven approaches in identifying best practices, and suggest the potential for tailoring professional development efforts to meet the diverse needs of tutors. The research opens up opportunities for future studies that expand on these methods, leading to the development of more innovative teaching strategies to enhance student learning and engagement.

Figures

Figure 1

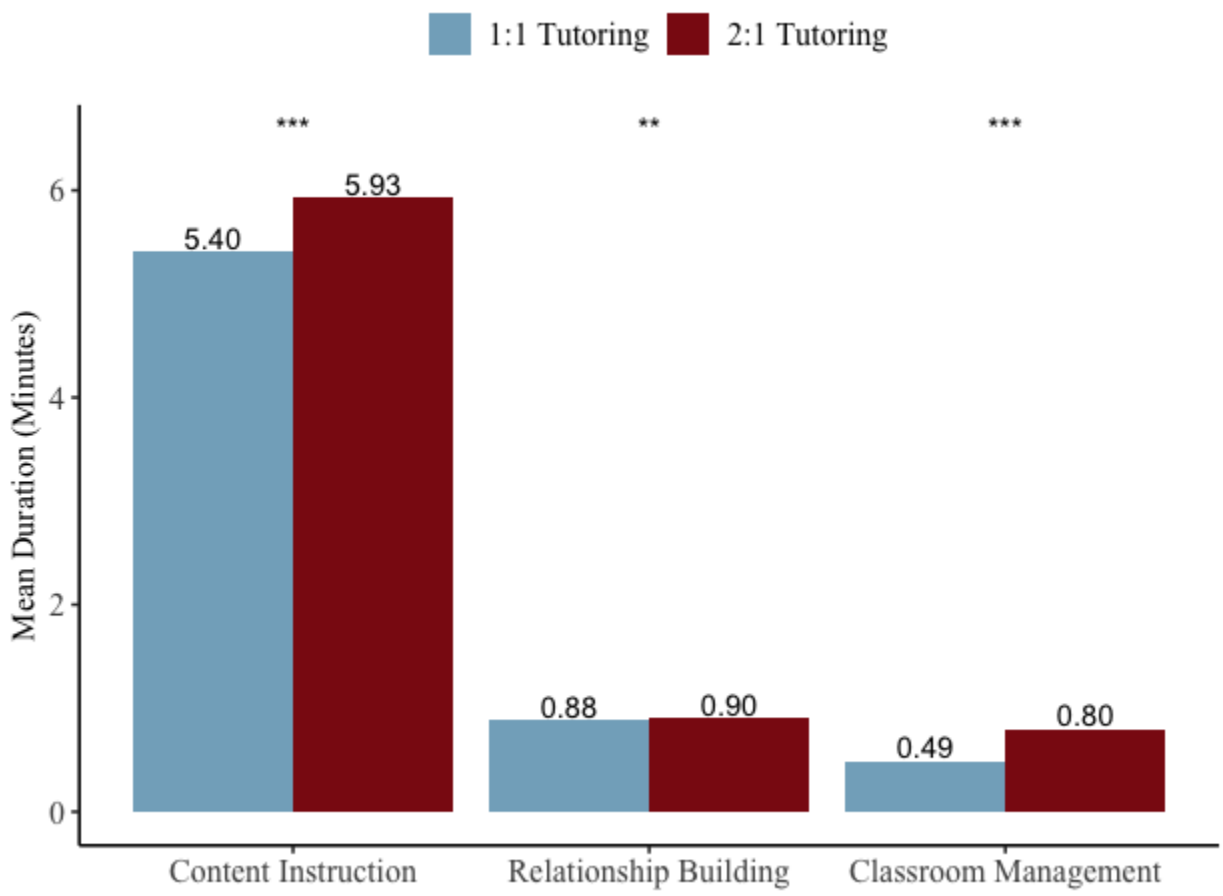
Distributions of Session Structural Features by Group Size



Note: This figure illustrates the distributions of structural session features for one-on-one sessions and two-on-one sessions. The blue areas represent one-on-one sessions and the red areas represent two-on-one sessions.

Figure 2

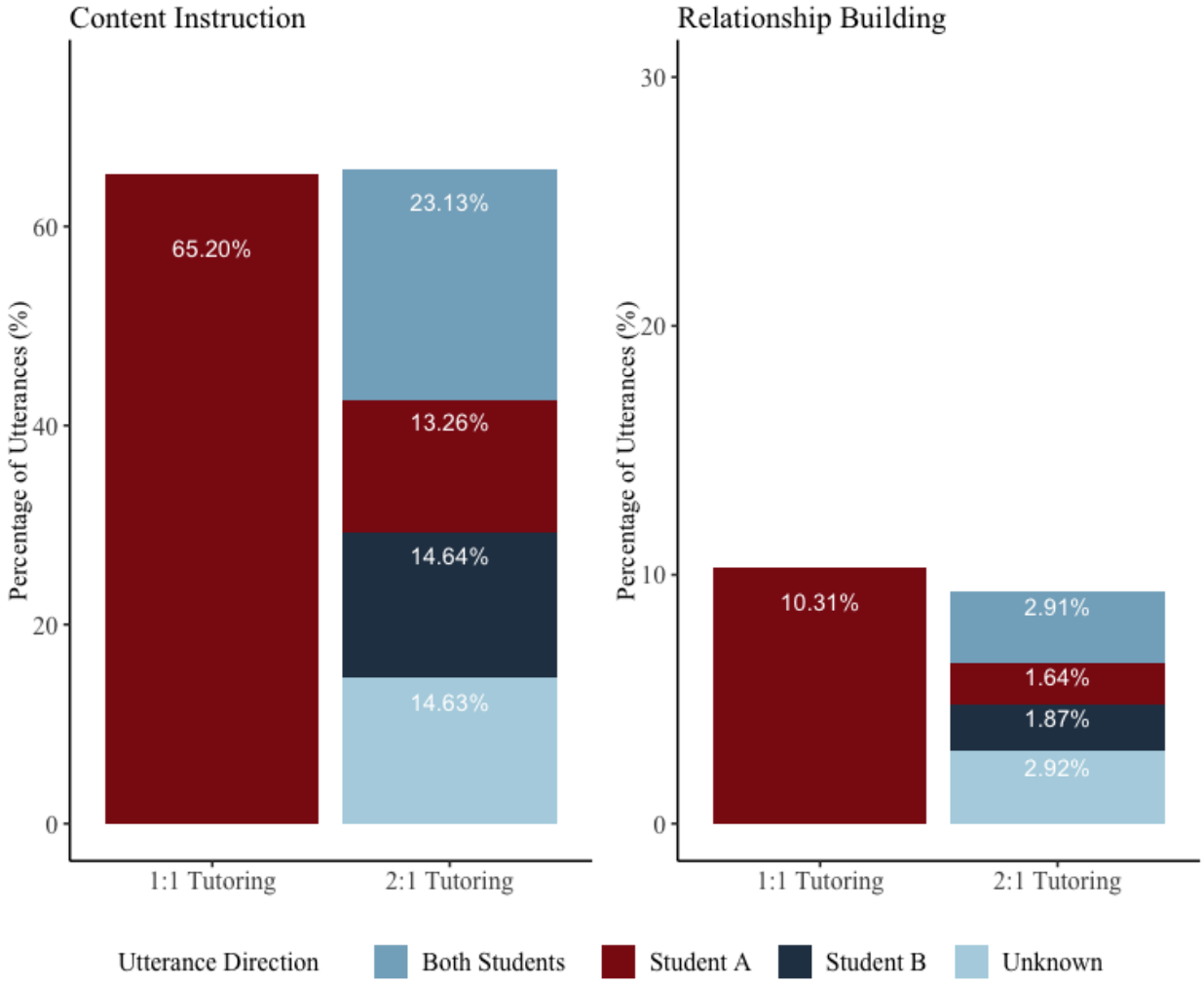
Time Use by Different Utterance Type and Group Size



Note: This figure displays mean duration of time use, measured in tutor talk-time in minutes, during sessions (average length of 19.17 minutes) for one-on-one (blue) and two-on-one (red) sessions across three different time uses: content instruction, relationship building, and classroom management. *p<.05, **p<.01, ***p<.001

Figure 3

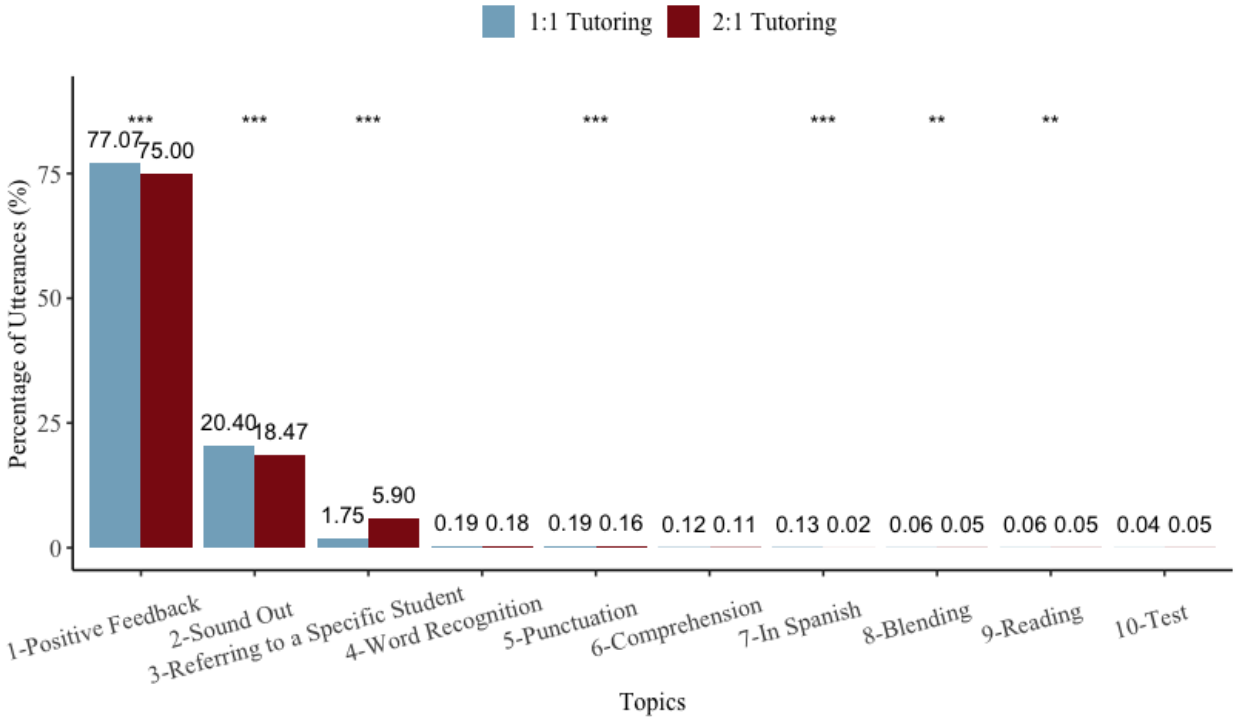
Average Percentage of Utterances by Attention Direction and Interaction Type



Note: This figure illustrates how tutors allocate their attention (in mean percentage of utterances) in one-on-one and two-on-one sessions during content instruction and relationship building.

Figure 4

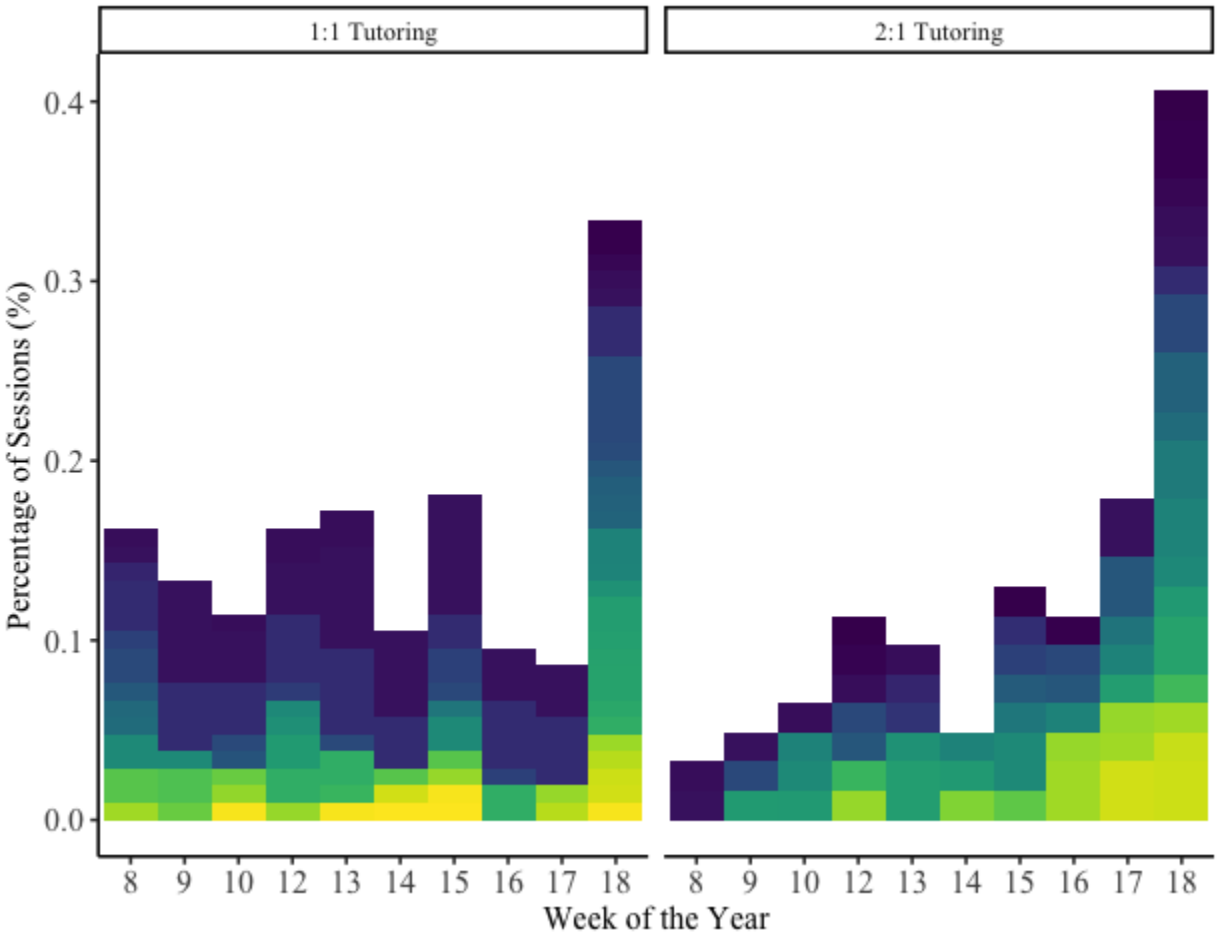
Clusters of Topics from Content-Instruction Utterances



Note: This figure demonstrates the percentage of utterances for ten content-related topics found through topic modeling across one-on-one (blue) and two-on-one (red) sessions. *p<.05, **p<.01, ***p<.001

Figure 5

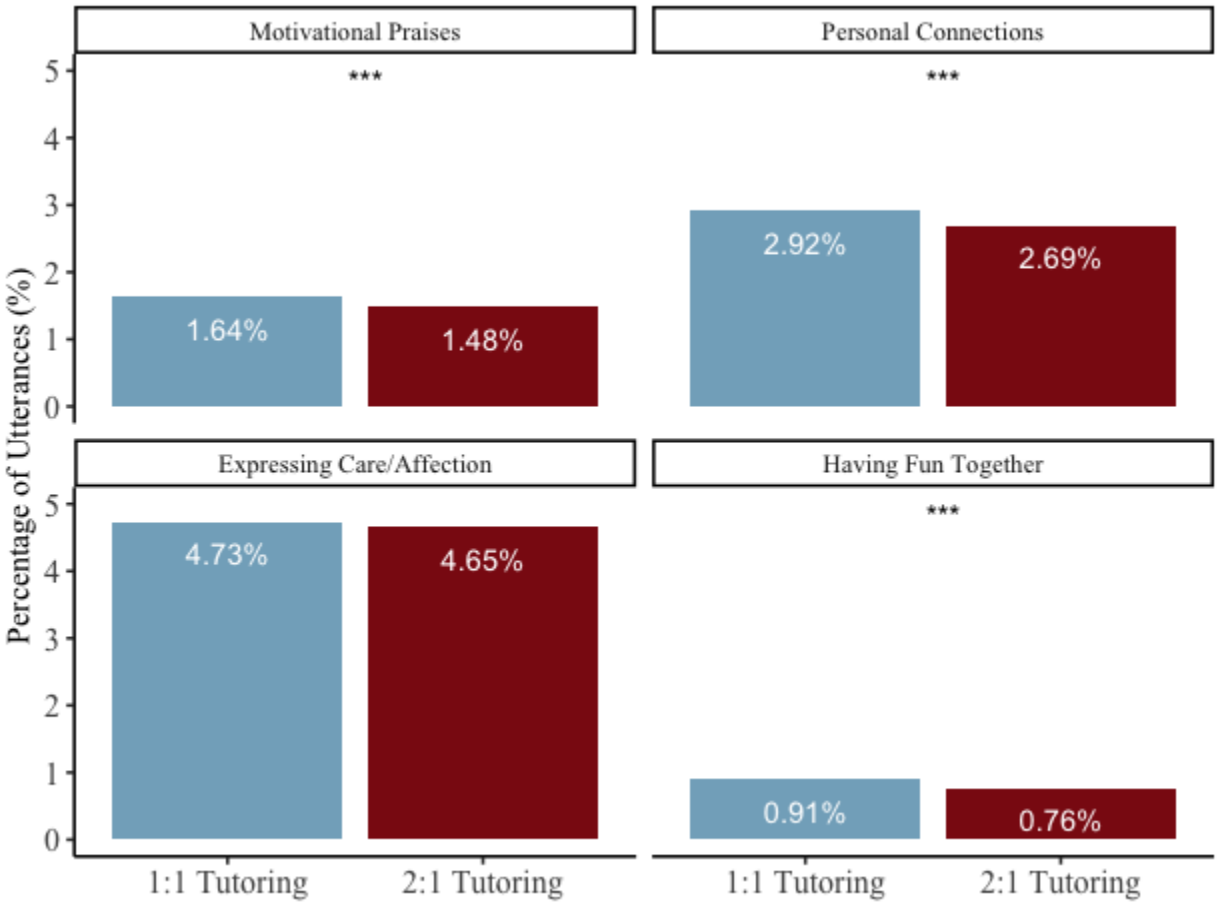
Histograms of Tutor Spanish Utterances across the Program



Note. This figure illustrates the use of Spanish by tutors throughout the program by group size. The x-axis gives the week of the tutoring sessions during the calendar year, starting at the end of February and ending in May. Each bar shows the percent of sessions in which tutors used Spanish, with each color representing a different tutor (approximately 60 in total). The consistent appearance of the same colors indicates that some tutors regularly use Spanish during their sessions.

Figure 6

Mean Percentage of Utterances by Relationship Building Subcategories and Group Size



Note: This figure illustrates the mean percent of utterances for four different relationship-building subcategories across one-on-one (blue) and two-on-one (red) sessions. *p<.05, **p<.01, ***p<.001

Tables

Table 1. Sample Utterances for Text Classifiers

Label	Sub-label	Definition	Sample Utterance
Content Instruction	YES	Instructions for learning to read	● Let's look at our irregular words. ● All right, let's sound it out." "How about this one? ● Good job!
	NO/Null	Non-content learning related	● How are you doing today? ● Ah, do you have a tree house? ● I'll let you draw because we finished everything. ● I need you to mute yourself, [Student B]. ● Could you hear me? I want you to listen carefully.
Relationship Building	Motivational praise	Positive reinforcement to create an engaging learning environment	● You are both fantastic spellers! ● It's okay if you don't know.
	Making personal connections	Sharing personal stories outside the lessons, answering icebreakers	● My dog wants to say hello to you guys. ● Do you play Minecraft?" ● ¿Te gustan los gatos? (Do you like cats?)
	Showing care or affection	Expressing interest in student well being	● I haven't seen you for a while. ● ¿Te divertiste? (Did you have fun?) ● What's the matter, baby? ● Oh, baby girl, that's sweet.
	Having fun together	Engaging in recreational activities like games	● All right, we're gonna play tic-tac-toe, all right? ● It's hard to beat you. ● We're gonna play Let's Spot It.
	Null	None of the above	
Classroom Management	Behavioral	Directing students to follow turn-taking or concentrate on the task	● You can't write it, but I'll let you write now because I made it so you could not write because you were not listening. ● Because I will make it so that you can't draw. ● Put your mouse away, because I don't want... I'll let you know when it's time to use that, okay?
	Technical	Troubleshooting technical glitches including unfamiliarity with or malfunctioning of the microphone or camera	● Uh-oh, no, we can't hear her. ● You still don't see anything? ● Are you able to hear me?

Null	None of the above

Table 2. Sample Tutor Utterances in Spanish

Utterance Type	**Examples**
Asset-based Instructional Practices	● *¿Entendiste cómo? (Did you understand how?)* ● *¿Cómo se dice lluvia en inglés? (How do you say rain in English?)* ● *En inglés, ¿cómo se dice huevo? (In English, how do you say egg?)*
Fostering a personal connection with students	● *Yes, I have Abuela Maria and Grandma Lily May.* ● *¿Te divertiste? (Did you have fun?)* ● *¿Tu perro es grande o es pequeño? (Is your dog big or your dog small?)*
Building a low-stress, high trust learning environment	● *¿Estás bien? (Are you ok?)* ● *Vamos a seguir practicando eso para que nos resulte más fácil, más fácil, más fácil, más fácil. (We are going to continue practicing that so that it becomes easier for us, easier, easier, easier.)* ● *Nos vemos mañana. (See you tomorrow)*

Table 3. Sample Tutor Utterances of Tic-Tac-Toe

Utterance Type	**Examples** (each bullet point is a unique session)
In two-on-one sessions	● *Okay, let's do a speed run on tic-tac-toe, okay? You guys have to take turns after each other, okay?* ● *I don't blame Student A for not wanting to play tic-tac-toe. He and I played a lot of tic-tac-toe last semester.*
In a one-on-one session	● *[Student], would you like to play tic-tac-toe or draw?* *...* *You think you're gonna win?* *...* *So where do you want to go first?* *It's hard to beat you.* *...* *No, you don't want to go there, do you?* *You should go somewhere else.* *You got three in a row.* *How did you do that?* *You beat me.*

References

Abdelshiheed, M., K. Jacobs, J., & K. D'Mello, S. (2024, July). Aligning tutor discourse supporting rigorous thinking with tutee content mastery for predicting math achievement. In International Conference on Artificial Intelligence in Education (pp. 150-164). Cham: Springer Nature Switzerland.

Agarwal, N., Sangwan, M., & Agarwal, E. (2024). Revolutionizing Pedagogy: Exploring AI Coaching in Enhancing Teacher Professionalism. In Integrating Generative AI in Education to Achieve Sustainable Development Goals (pp. 264-281). IGI Global.

Bain, M., Huh, J., Han, T., & Zisserman, A. (2023). Whisperx: Time-accurate speech transcription of long-form audio. arXiv preprint arXiv:2303.00747.

Bhatt, M. P., Guryan, J., Khan, S. A., LaForest-Tucker, M., & Mishra, B. (2024). Can technology facilitate scale? Evidence from a randomized evaluation of high dosage tutoring (No. w32510). National Bureau of Economic Research.

Bloom, Benjamin S. (1984). "The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring." Educational Researcher 13, no. 6: 4–16.

Booth, B. M., Jacobs, J., Bush, J. B., Milne, B., Fischaber, T., & DMello, S. K. (2024, March). Human-tutor coaching technology (htct): Automated discourse analytics in a coached tutoring model. In Proceedings of the 14th Learning Analytics and Knowledge Conference (pp. 725-735).

Breiseth, L. (2020). 10 Strategies for Building Relationships with ELLs. Colorín Colorado. Retrieved April 28, 2025, from https://www.colorincolorado.org/article/building-relationships-ells

Carlana, M., & La Ferrara, E.. (2021). Apart but Connected: Online Tutoring and Student Outcomes during the COVID-19 Pandemic. (EdWorkingPaper: 21 -350). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/0azm-cf65

Center on Reinventing Public Education [CRPE]. (2023). The State of the American Student: Fall 2023, We are Failing Older Students: Bold Ideas to Change Course. Available at https://crpe.org/wp-content/uploads/The-State-of-the-American-Student-2023.pdf

Center on Reinventing Public Education [CRPE]. (2024). The State of the American Student: Fall 2024, Solve for the Most Complex Needs: A Path Forward as Pandemic Effects Reverberate. Available at https://crpe.org/wp-content/uploads/CRPE_SOS2024_FINAL.pdf

Cortes, K. E., Kortecamp, K., Loeb, S., & Robinson, C. D. (2025). A scalable approach to high-impact tutoring for young readers. Learning and Instruction, 95, 102021.

Demszky, D., Liu, J., Mancenido, Z., Cohen, J., Hill, H., Jurafsky, D., & Hashimoto, T. (2021). Measuring conversational uptake: A case study on student-teacher interactions. arXiv preprint arXiv:2106.03873.

Demszky, D., Liu, J., Hill, H. C., Sanghi, S., & Chung, A. (2023). Improving Teachers' Questioning Quality through Automated Feedback: A Mixed-Methods Randomized Controlled Trial in Brick-and-Mortar Classrooms. EdWorkingPaper No. 23-875. Annenberg Institute for School Reform at Brown University.

Dietrichson, J., Bøg, M., Filges, T., & Jørgensen, A.-M. K. (2017). Academic interventions for elementary and middle school students with low socioeconomic status: A systematic review and meta-analysis. Review of Educational Research, 87(2), 243– 282. https://doi.org/10.3102/0034654316687036

Feldon, D. F. (2007). Cognitive Load and Classroom Teaching: The Double-Edged Sword of Automaticity. Educational Psychologist, 42(3), 123–137. https://doi.org/10.1080/00461520701416173

Ferraro, A., Galli, A., La Gatta, V., & Postiglione, M. (2023). Benchmarking open source and paid services for speech to text: an analysis of quality and input variety. Front. Big Data 6 (2023).

Gortazar L., Hupkau C., & Roldan A. (2023). Online tutoring works: Experimental evidence from a program with vulnerable children. EdWorkingPapers.com. Annenberg Institute at Brown University. https://www.edworkingpapers.com/ai23-743

Gromada, A. and C. Shewbridge (2016), "Student Learning Time: A Literature Review", OECD Education Working Papers, No. 127, OECD Publishing, Paris, https://doi.org/10.1787/5jm409kqqkjh-en.

Groom-Thomas, L., Leung, C., Loeb, S., Pollard, C., Waymack,N., & White, S. (2023) Challenges and Solutions: Scaling Tutoring Programs. Inter-American Development Bank.

Guryan, J., Christenson, S., Cureton, A., Lai, I., Ludwig, J., Schwarz, C., Shirey, E., & Turner, M. C. (2021). The effect of mentoring on school attendance and academic outcomes: A randomized evaluation of the Check & Connect Program. Journal of Policy Analysis and Management, 40(3), 841–882. https://doi.org/10.1002/pam.22264

Guryan, J., Ludwig, J., Bhatt, M. P., Cook, P. J., Davis, J. M., Dodge, K., ... & Stoddard, G. (2023). Not too late: Improving academic outcomes among adolescents. American Economic Review, 113(3), 738-765.

Hashim, Shirin, Katharine Pace Miles, and Erin Croke. (2025). Experimental Evidence on the Impact of Tutoring Format and Tutors: Findings from an Early Literacy Tutoring Program. (EdWorkingPaper: 25 -1176). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/zgy5-sr80

Holland, P., Alfaro, P., & Evans, D. (2015). Extending the School Day in Latin America and the Caribbean (SSRN Scholarly Paper No. 2620017). Social Science Research Network. https://papers.ssrn.com/abstract=2620017

Ichii, N. (2022). Students' Perceptions Toward Dyads and Triads in the English Classroom. ISSN: 2189-1036 – The IAFOR International Conference on Education – Hawaii 2022 Official Conference Proceedings, 245–255.

Jacobs, J., Scornavacco, K., Harty, C., Suresh, A., Lai, V., & Sumner, T. (2022). Promoting rich discussions in mathematics classrooms: Using personalized, automated feedback to

support reflection and instructional change. Teaching and Teacher Education, 112, Article 103631. https://doi.org/10.1016/j. tate.2022.103631.

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. Educational researcher, 38(5), 365-379.

Kahneman, D. (1973). Attention and effort (Vol. 1063, pp. 218-226). Englewood Cliffs, NJ: prentice-Hall.

Kim, J. (2024). "Leading teachers' perspective on teacher-AI collaboration in education". Education and Information Technologies, 29(7), 8693-8724.

L'Enfant, J. (2024). AI as a reflective coach in graduate ESL practicum: activity theory insights into student-teacher development. European Journal of Open, Distance & E-Learning, 26(1).

Lee, M. G., Loeb, S., & Robinson, C. D. (2024). Effects of High-Impact Tutoring on Student Attendance: Evidence from the OSSE HIT Initiative in the District of Columbia. EdWorkingPaper. Annenberg Institute at Brown University.

Liu, X., Zhang, J., Barany, A., Pankiewicz, M., & Baker, R. S. (2024, November). Assessing the Potential and Limits of Large Language Models in Qualitative Coding. In International Conference on Quantitative Ethnography (pp. 89-103). Cham: Springer Nature Switzerland.

National Student Support Accelerator [NSSA]. (n.d.) Toolkit for Tutoring Programs: Relationship Building. National Student Support Accelerator. Retrieved April 28, 2025, from https://nssa.stanford.edu/tutoring/instruction/relationship-building

Neitzel, A. J., Lake, C., Pellegrini, M., & Slavin, R. E. (2022). A synthesis of quantitative research on programs for struggling readers in elementary schools. Reading Research Quarterly, 57(1), 149-179.

Nickow, A., Oreopoulos, P., & Quan, V. (2024). The promise of tutoring for PreK–12 learning: A systematic review and meta-analysis of the experimental evidence. American Educational Research Journal, 61(1), 74-107.

O'Connor, C., Michaels, S. (2019) Supporting teachers in taking up productive talk moves: The long road to professional learning at scale. International Journal of Educational Research 97, 166–175

Patall, E. A., Cooper, H., & Allen, A. B. (2010). Extending the School Day or School Year: A Systematic Review of Research (1985–2009). Review of Educational Research, 80(3), 401–436. https://doi.org/10.3102/0034654310377086

Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2023, July). Robust speech recognition via large-scale weak supervision. In International conference on machine learning (pp. 28492-28518). PMLR.

Robinson, C., Kraft, M., Loeb, S., & Schueler, B. (2021). DESIGN PRINCIPLES FOR ACCELERATING STUDENT LEARNING WITH HIGH-IMPACT TUTORING.

Robinson, C. D., Pollard, C., Novicoff, S., White, S., & Loeb, S. (2024). The effects of virtual tutoring on young readers: Results from a randomized controlled trial. EdWorkingPapers. com.

Robinson, Carly D., and Susanna Loeb. (2021). High-Impact Tutoring: State of the Research and Priorities for Future Learning. (EdWorkingPaper: 21 -384). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/qf76-rj21

Roorda, D. L., Jak, S., Zee, M., Oort, F. J., & Koomen, H. M. Y. (2017). Affective teacher–student relationships and students' engagement and achievement: A meta-analytic update and test of the mediating role of engagement. School Psychology Review, 46(3), 239–261. https://doi.org/10.17105/SPR-2017-0035.V46-3

Scales, P. C., Pekel, K., Sethi, J., Chamberlain, R., & Van Boekel, M. (2020). Academic Year Changes in Student-Teacher Developmental Relationships and Their Linkage to Middle and High School Students' Motivation: A Mixed Methods Study. The Journal of Early Adolescence, 40(4), 499–536. https://doi.org/10.1177/0272431619858414

Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective programs for struggling readers: A best-evidence synthesis. Educational Research Review, 6(1), 1–26.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285. https://doi.org/10.1016/0364-0213(88)90023-7

Vrieling-Teunter, E., Henderikx, M., Nadolski, R., & Kreijns, K. (2022). Facilitating Peer Interaction Regulation in Online Settings: The Role of Social Presence, Social Space and Sociability. Frontiers in Psychology, 13. https://doi.org/10.3389/fpsyg.2022.793798

Vygotsky, L. S. (1978). Mind in Society: Development of Higher Psychological Processes. Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

Wang, R. E., Ribeiro, A. T., Robinson, C. D., Loeb, S., & Demszky, D. (2024). Tutor CoPilot: A human-AI approach for scaling real-time expertise. arXiv preprint arXiv:2410.03017.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry, 17(2), 89–100. https://doi. org/10.1111/j.1469-7610.1976.tb00381.x

Yeşil Dağlı, Ü. (2019). Effect of increased instructional time on student achievement. Educational Review, 71(4), 501–517. https://doi.org/10.1080/00131911.2018.1441808

Zajonc, R. B. (1965). Social Facilitation. Science, 149(3681), 269–274. https://doi.org/10.1126/science.149.3681.269

Zhang, Q., Wang, R. E., Ribeiro, A. T., Demszky, D., & Loeb, S. (2025). Educator Attention: How computational tools can systematically identify the distribution of a key resource for students. EdWorkingPapers.com. https://doi.org/10.26300/hvq6-d385

Appendix

Correlations between Tutor Utterances and Group Size

	Session Length		**Tutor Talk**		**Content Talk**		**Relationship Building Talk**		**Classroom Management Talk**	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
Two-on-One	**0.85**	<0.001	**0.80**	<0.001	**0.54**	<0.001	0.04	0.091	**0.31**	<0.001
EL	**0.16**	0.003	-0.12	0.308	0.05	0.563	**-0.08**	0.001	0.01	0.613
Female	0.00	0.921	-0.00	0.988	-0.11	0.185	0.02	0.406	-0.01	0.570
Hispanic	0.02	0.752	0.01	0.949	0.04	0.641	-0.00	0.879	0.01	0.666
SPED	0.05	0.616	-0.16	0.521	-0.06	0.739	**-0.08**	0.015	0.08	0.089
MOY Test Score	**0.00**	0.032	-0.00	0.403	**-0.01**	<0.001	0.00	0.080	0.00	0.971
(Intercept)	17.90	<0.001	9.02	<0.001	8.74	<0.001	0.60	0.001	0.47	0.001
Observations	16535		16535		16535		16535		16535	
R^2 / R^2 adjusted	0.072 / 0.072		0.031 / 0.031		0.033 / 0.033		0.008 / 0.008		0.089 / 0.089	

Note. This table displays the correlations between tutor utterances (measures of talk time in minutes) and tutoring format (one-on-one vs two-on-one) controlling for student baseline literacy proficiency and characteristics. Results show that, on average, tutoring sessions were longer (51 seconds) in two-on-one tutoring, where tutors also talked a little more (48 seconds). When disaggregating tutor utterances by interaction type, in two-on-one tutoring tutors dedicated slightly longer time to content instruction and classroom management (32.4 seconds and 18.6 seconds, respectively) and spent similar time on relationship building.