Training, Assessing, and Matching Students in Technology-Mediated Peer Learning Environments

Soren Rosier*
Stanford University

Ojas Karnavat
University of California, Berkeley

Abstract

While tutoring has historically been the most effective learning intervention, it is costly to administer and challenging to integrate in classrooms. Peer tutoring is a relatively untapped solution due to barriers around time management, organization, pairing, and assessment. Following trials of an interactive training program called PeerTeach that rapidly improved middle schoolers' peer tutoring ability, our research team established a research practice partnership to design self-sustaining classrooms with peer tutoring as a central focus. This paper reports on our five-month mixed methods design-based research study with an Indian middle school where we found convincing evidence for the viability of technology-mediated peer tutoring as an antidote to key implementation challenges. Through this study, we gained valuable insights on critical roles technology can play in training effective peer helpers, optimally matching tutoring pairs, overcoming implementation pitfalls, and promoting motivational elements students most value.

Keywords: peer tutoring, collaborative learning, technology-mediated learning, design-based research

*corresponding author

"The happiness you get from teaching is just not comparable to anything else."
- 7th grade student, Pune, India

Introduction

One-to-one tutoring consistently produces the largest effect sizes among educational interventions (Bloom, 1984; Dietrichson et al., 2017). Meta-analyses confirm that tutoring surpasses nearly all other interventions in terms of effectiveness, with effect sizes typically ranging from 0.4 to 0.8 standard deviations (Nickow et al., 2020; Fryer, 2016). Beyond academic gains, tutoring promotes numerous affective benefits including improved subject interest, confidence, motivation, self-esteem, and communication skills (Lepper & Woolverton, 2002; Topping, 2017).

Despite its demonstrated efficacy, traditional tutoring remains financially inaccessible to most families and logistically challenging to implement at scale. The average cost of private tutoring in the US is $437 per month, with the top 5% of spenders paying approximately $2,000 monthly (Kim et al., 2025). Even "high-dosage" school-based tutoring programs require substantial investments of $2,500 to $3,800 annually per student (Ander et al., 2016). These programs face significant implementation challenges, including scheduling difficulties, spatial constraints, and inconsistent participation (Robinson & Loeb, 2021).

Peer tutoring represents a specific instructional arrangement within the broader spectrum of collaborative learning approaches. As Topping (2005) defines it, peer tutoring involves students actively helping each other to learn, with one student taking a direct instructional role as tutor while another serves as tutee. This arrangement differs from other collaborative approaches in its intentional role asymmetry and instructional focus (Falchikov, 2001). Within collaborative learning—where students work together toward academic goals (Dillenbourg, 1999)—peer tutoring occupies a distinct position characterized by its structured roles, explicit teaching functions, and typically dyadic format.

While peer tutoring demonstrates substantial learning benefits for both tutors and tutees (Bowman-Perrott et al., 2013; Leung, 2015), it remains underutilized in classroom practice due to persistent implementation barriers (Grubbs, 2009; Topping et al., 2017). Though technology has been leveraged to support various aspects of peer learning, a gap remains in understanding how comprehensive technology-mediated systems can address the full range of implementation barriers specific to peer tutoring.

Particularly understudied is how technology might simultaneously address the primary obstacles that prevent wider adoption of peer tutoring: teacher organizational burden, effective tutor training, optimal pairing processes, and assessment integration (Buchs et al., 2017; Jolliffe, 2015). The purpose of this design-based research study was to investigate how technology-mediated systems can be designed to facilitate effective peer tutoring implementation in middle school classrooms. Through a research-practice partnership with an Indian middle school, we developed and iteratively refined a comprehensive technological infrastructure to address key barriers that typically impede peer tutoring adoption. This study contributes to the literature by demonstrating how technology can specifically support tutor training, peer matching, routine establishment, and assessment—elements critical to creating sustainable peer tutoring communities.

While previous research has examined isolated components of peer tutoring—such as tutor training (Roscoe & Chi, 2007), pairing strategies (Webb & Mastergeorge, 2003), or assessment approaches (Slavin, 2015)—this study makes a unique contribution by investigating how a comprehensive, integrated technological system can simultaneously address multiple implementation barriers. Unlike prior work that has primarily focused on either the cognitive aspects of peer tutoring (Chi et al., 2001) or its social dynamics (Johnson & Johnson, 2009), our research integrates both dimensions through a technology-mediated platform that combines evidence-based training modules, data-driven matching algorithms, and automated assessment systems. Furthermore, existing studies have typically evaluated peer tutoring interventions in controlled settings with extensive researcher support (Topping et al., 2017). Instead, our design-based research approach examines implementation in authentic classroom contexts with an emphasis on sustainability. By investigating student experiences across the full implementation cycle using both quantitative and qualitative methods, this study offers novel insights into the interplay between technological design features, implementation barriers, and the elements of peer tutoring students find most valuable. These insights directly inform the development of scalable, technology-mediated solutions to transform peer tutoring from a promising but underutilized approach to a sustainable classroom practice.

Research Questions

Primary Research Question: How should technology-mediated peer tutoring environments be designed to overcome key implementation barriers and create sustainable co-learning communities in middle school math classrooms?

Secondary Research Questions:

1. What technological design features most effectively address barriers to peer tutoring implementation related to tutor training, student pairing, and assessment?
2. How do these technology-mediated design solutions influence the development of sustainable peer tutoring practices?
3. What aspects of peer tutoring do students value most, and how can technology enhance these valued elements?

Theoretical Framework

Peer tutoring represents a specific instructional arrangement within the broader spectrum of collaborative learning approaches. This study is guided by two primary theoretical foundations that inform both the design of the intervention and our interpretation of results: Vygotsky's sociocultural learning theory and Lewin's field theory.

Vygotskian Foundations of Peer Tutoring

Vygotsky's (1978) sociocultural learning theory provides the primary theoretical foundation for peer tutoring. Central to this approach is the concept of the Zone of Proximal Development (ZPD)—the space between what a learner can accomplish independently and what they can achieve with assistance. Peer tutoring operationalizes the ZPD by creating structured interactions where a more knowledgeable peer provides targeted support that helps the learner bridge this gap.

Unlike traditional classroom instruction, peer tutoring creates what Rogoff (1990) terms "guided participation," where learners actively construct understanding through collaborative dialogue and joint problem-solving. The asymmetrical helping relationship facilitates the internalization of social processes—where external, interpersonal activities are transformed into internal cognitive functions. This transformation occurs as tutees appropriate the problem-solving strategies modeled by their tutors, while tutors deepen their own understanding through the verbalization and reorganization of knowledge required for effective explanation (Roscoe & Chi, 2007).

Our technological design specifically supports this Vygotskian framework by training peer tutors in eliciting questions, revoicing, and probing—strategies that create the dialogic space necessary for knowledge construction within the ZPD. Furthermore, our matching algorithm operationalizes Vygotsky's theory by pairing students based on complementary knowledge states.

Lewin's Field Theory and Peer Interaction Dynamics

While Vygotsky's theory addresses the cognitive mechanisms of peer learning, Lewin's (1951) field theory provides a complementary framework for understanding the social and motivational dynamics of peer tutoring interactions. Field theory conceptualizes behavior as a function of the person and their environment ($B = f[P, E]$), emphasizing that learning is influenced by both individual characteristics and the social context.

Lewin's concepts of positive interdependence and psychological safety inform our approach to creating sustainable peer tutoring communities. Positive interdependence—the perception that one's success is linked to others'—is essential for motivating engagement in collaborative learning (Johnson & Johnson, 2009). Our technological design promotes this interdependence by making learning gains visible and celebrating collective achievement, helping students recognize the mutual benefits of the tutoring relationship.

Similarly, Lewin's emphasis on creating psychological safety within group settings guides our attention to compatibility and comfort in peer pairings. As Edmondson (1999) demonstrated, psychological safety—the shared belief that the environment is safe for interpersonal risk-taking—is essential for productive learning interactions. Our matching algorithm's incorporation of compatibility data operationalizes this theoretical insight, creating conditions where students feel comfortable enough to engage in the vulnerable acts of asking questions, admitting confusion, and taking intellectual risks.

Understanding Implementation Challenges: A Theoretical Perspective

Understanding the barriers to peer tutoring implementation requires examining both the cognitive demands of effective tutoring and the social structures needed to support productive interaction. From a Vygotskian perspective, the limited implementation of peer tutoring partly stems from students' lack of internalized models for effective scaffolding behaviors. Without explicit training, students typically default to didactic explanation patterns that fail to navigate the ZPD effectively (Webb & Mastergeorge, 2003).

From a Lewinian perspective, implementation challenges also arise from insufficient attention to the field forces that shape collaborative behavior. The organizational burden teachers face when implementing peer tutoring reflects the substantial effort required to create the environmental conditions necessary for productive peer learning. Similarly, students' reluctance to engage fully in asymmetrical helping relationships often stems from insufficient positive interdependence and psychological safety within the learning environment.

Our technological design addresses these theoretically grounded challenges by: (1) providing explicit models and practice opportunities for effective scaffolding behaviors, supporting the internalization of productive tutoring strategies; (2) automating the logistical aspects of creating optimal learning conditions, reducing teacher burden; and (3) intentionally building positive interdependence and psychological safety through compatibility-based matching and visible recognition of collective achievement.

Practical Barriers to Implementation in Classrooms

Peer tutoring faces several critical implementation barriers that prevent widespread adoption despite its demonstrated effectiveness.

First, teachers perceive the implementation of structured peer interactions as excessively burdensome given their existing responsibilities (Buchs et al., 2017; Jolliffe, 2015). The organizational demands include creating appropriate materials, monitoring interactions, and carefully matching students based on knowledge states and interpersonal dynamics (Topping, 2015). As research on collaborative learning implementation has demonstrated, the perceived workload often outweighs anticipated benefits, even when teachers value the approach conceptually (Buchs et al., 2017).

Second, creating positive interdependence—a sense of shared goals and mutual benefit—presents significant challenges in asymmetrical helping relationships like peer tutoring (Johnson & Johnson, 2009). Without carefully engineered social structures, students may perceive limited personal benefit from participating, particularly in subjects traditionally viewed as individualistic (Slavin, 2015).

Third, students typically lack the pedagogical skills needed for effective peer instruction. Research on helping behaviors in collaborative contexts consistently shows that untrained students adopt ineffective approaches characterized by didactic explanation and limited responsiveness to peer understanding (Webb & Mastergeorge, 2003; Roscoe & Chi, 2007).

These interaction patterns are particularly problematic in peer tutoring, where the explicit teaching function demands specific pedagogical moves for effective knowledge construction (Chi et al., 2001). Without intervention, peer tutors engage in suboptimal behaviors that may actually diminish learning compared to no assistance at all (Webb & Mastergeorge, 2003).

Fourth, many teachers lack both the training and tools needed to develop effective collaborative skills in their students (Topping et al., 2017; Cohen, 1997). This challenge, while documented across collaborative learning approaches, takes on particular significance in peer tutoring contexts, where the explicit instructional role demands more specialized preparation than general collaborative work. As Cohen (1997) emphasized, productive peer interactions require "specific development of skills for discourse, either in advance of cooperative learning or through direct assistance when groups are in operation" (p. 7).

These barriers collectively account for the limited adoption of peer tutoring despite its demonstrated effectiveness. As Greenwood et al. (2002) observed in their implementation research on classroom-wide peer tutoring, even successful programs often fail to achieve sustainability once external support is withdrawn. This sustainability challenge highlights the need for technological infrastructures that can reduce implementation barriers and support long-term adoption of effective peer tutoring practices.

Methodology

Design-Based Research Approach

This study employed design-based research methodology (Design-Based Research Collective, 2003) to iteratively develop and refine technology-mediated peer tutoring environments in authentic classroom settings. This approach facilitates "the design of technology-enhanced learning environments and the development of emerging theories of learning" (Wang & Hannafin, 2005, p. 5) through systematic processes that integrate research and practice.

Using this methodology, we collaboratively designed technology-mediated solutions with practitioners, implemented them in classroom contexts, collected data on their effectiveness, and refined both the intervention and our theoretical understanding based on emerging findings. This iterative approach allowed us to develop practical tools that addressed immediate implementation challenges while generating knowledge about effective peer tutoring.

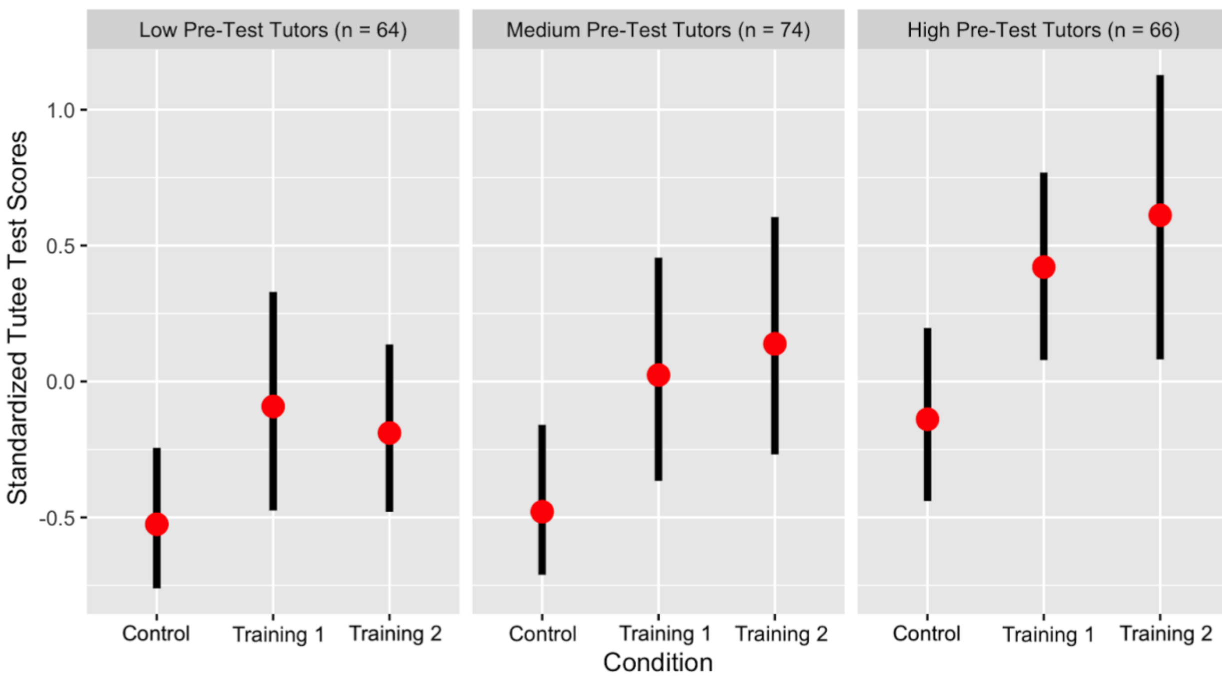
Development of PeerTeach

PeerTeach is a web-based application co-designed by Stanford researchers and middle school math teachers through multiple rounds of iteration over four years. Two Randomized Controlled Trials (RCTs) were previously conducted to establish the efficacy of the PeerTeach trainings, involving 204 student dyads (Rosier, 2023). These RCTs compared two PeerTeach trainings against control groups, with both trainings designed to promote learner-centered teaching behaviors.

Results from these prior studies demonstrated that the PeerTeach trainings significantly increased the frequency with which tutors employed learner-centered strategies in both virtual scenarios and real-life tutoring. These behavioral changes translated to substantial learning gains for tutees across all levels of tutor content mastery. Effect sizes were large: .78 for the first PeerTeach training and .92 for the second, compared to controls (Rosier, 2023). Students coached by PeerTeach-trained tutors showed approximately 30% higher post-test scores compared to those helped by control tutors. As Figure 1 highlights, a key finding was that both tutor training and tutor content knowledge were strong predictors of effective teaching. These findings directly informed our current implementation, particularly the decision to identify individual topics each student had mastered and could tutor, while training all students on effective teaching methods.

Figure 1[1]

Combining both rounds of data, tutor pre-test scores and condition both predict tutee learning



Intervention

Following the previous successful RCTs, we formed a research-practice partnership with a middle school to transform PeerTeach from a training tool into a comprehensive platform that coordinates teaching interactions within a classroom of trained peer tutors. The intervention was implemented in two seventh-grade classes with a combined 57 students over a five-month period. Sessions were held online. All students had prior experience with remote learning due to COVID-19 measures and had access to necessary devices (laptops or tablets). While no formal assessment of technological competencies was conducted, technical support was provided as needed, and all students were able to successfully navigate the platform with minimal assistance.

[1] Data combined from two rounds of data collection by first converting tutee assessment scores and tutor pre-test scores into standardized z-scores. Dots represent means. Lines represent 95% confidence limits for the population mean based on nonparametric bootstrapping of data.

Each weekly session consisted of four components: a brief celebration of measured improvements from the prior week's peer tutoring, breakout rooms where paired students taught one another for 20 minutes, post-surveys, then 5-question assessments to track progress. Sessions were conducted remotely during class hours. New tutoring pairs were created weekly, with the matching algorithm favoring students switching roles each session whenever possible.

Throughout the implementation, teachers played a supportive but limited role in the peer tutoring process. Teachers validated the diagnostic assessments, occasionally observed sessions to provide feedback, and helped facilitate the weekly celebration of improvements. However, the tutoring interactions were primarily student-led, with minimal direct teacher intervention. The PeerTeach platform handled matching, training, and directing the tutoring interactions, minimizing additional teacher workload while maximizing student autonomy.

To understand the consequences of varied inputs to a peer learning ecosystem, we staggered the introduction of new technological components across three phases:

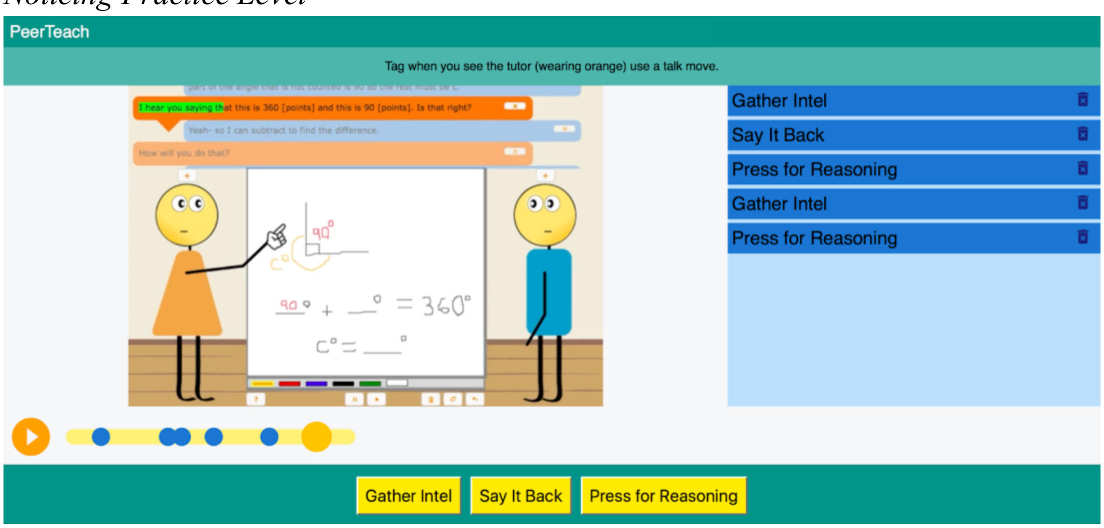
Phase 1 (Weeks 1-3): Naturalistic peer tutoring without technological interventions.

Phase 2 (Week 4): Introduction of the PeerTeach training, where every student completed a 40-minute training focusing on three key teaching moves: asking eliciting questions, revoicing to clarify understanding, and asking probing questions.

The training consisted of three integrated components: watching short videos on teaching moves, engaging in a noticing activity by tagging videos of tutoring interactions when key teaching moves occurred (see Figure 2), and participating in a simulation exercise where they taught a virtual character (see Figure 3). This design was informed by Sherin's (2005) noticing framework, which posits that teachers must attend to important teaching moments, relate them to useful pedagogical frameworks, and act based on pedagogically-sound reasoning.

Figure 2

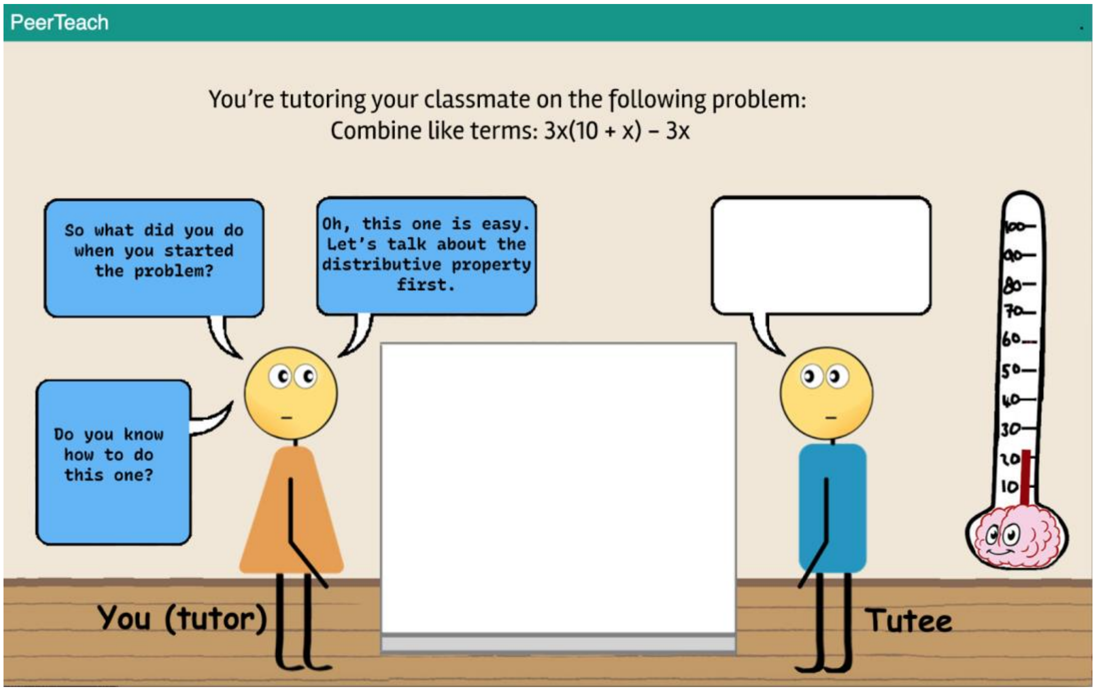
Noticing Practice Level



Note. Students tag the video each time the cartoon tutor uses one of the focal talk moves.

Figure 3

Decision Making Level



Phase 3 (Week 7): Enhancement of the matching algorithm to incorporate student compatibility data alongside topic mastery.

Matching Algorithm

The platform used two iterations of matching algorithms. The initial algorithm (Weeks 4-6) used math pre-assessment data to pair students with specific areas of need with peer tutors who had mastered those topics. The reward function for evaluating match quality increased with higher tutor mastery and lower tutee mastery. To ensure lower-performing students had teaching opportunities, the reward was slightly increased for tutors with lower overall math achievement.

The enhanced algorithm (Weeks 7-20) incorporated both topic mastery data and student compatibility data to optimize pairings. This approach maintained the benefits of mastery-based matching while adding the interpersonal dimension found to be critical for effective collaborative work (Edmondson, 1999). Once pairs were determined, students were informed of their partnerships and given three to four days to prepare peer tutoring lessons.

Sampling

The study involved 57 seventh-grade students from two classes at a middle school, selected through convenience sampling based on an existing partnership with the school and teacher willingness to participate. This sampling approach is consistent with design-based research in educational settings, where the emphasis is on iterative refinement of interventions in authentic contexts rather than large-scale generalization (Anderson & Shattuck, 2012). Students were only excluded if they or their parents opted out of participation in this study, which occurred for only one student. Students across various demographics and special education needs were represented among the final 57 students in this study.

From the full participant pool, twelve case study students were selected purposefully to ensure a wide range of perspectives. The selection balanced performance levels (four high-performing, four average-performing, and four low-performing students) and gender (six boys and six girls). This purposive sampling strategy follows recommendations by Creswell and Plano Clark (2018) for qualitative components within mixed-methods designs, where diversity of experiences is prioritized over statistical representation.

Data Collection Methods & Instruments

Students began the study by taking an online diagnostic assessment consisting of 5 questions on each of 20 topics they had learned in the prior six months. This diagnostic determined each student's mastery level on each math topic, which was essential for the initial pairing algorithm. After each weekly online peer tutoring session, students completed post-session surveys capturing their experiences and perceptions and 5-question assessments to track progress on taught concepts.

Robust pre- and post-surveys were administered at the beginning and end of the study to measure changes in attitudes, perceptions, and self-efficacy. Interview protocols were developed for semi-structured individual interviews with the twelve case study students. These protocols evolved over time in response to emerging themes from ongoing data analysis. Starting in week 7, students also completed surveys indicating how well they worked with each of their peers, specifically focusing on collaboration quality rather than personal preference.

Assessment items were aligned with the school's mathematics curriculum and reviewed by classroom teachers to ensure content validity. Survey instruments were piloted with a small sample of students similar to our study population to ensure clarity, appropriateness, and consistency. Weekly assessments were drawn from a standardized item bank and were reviewed to ensure comparable difficulty levels across versions. Interview protocols were developed iteratively with feedback from educational researchers and practitioners. Refinements were made to the interview protocols as themes emerged, with the goal of improving question clarity and ensuring a more consistent experience across interviews.

Data Collection Procedure

Data collection was interwoven with the intervention implementation, allowing each phase to inform subsequent modifications. Prior to the intervention, diagnostic assessments and pre-surveys were administered to establish baselines and inform initial pairings. During weeks 1-3, weekly post-session surveys and assessments collected during the naturalistic peer tutoring phase provided baseline data on untrained peer tutoring interactions.

The data from weeks 4-7 informed the decision to incorporate compatibility metrics into the matching algorithm. During the mid-intervention period, semi-structured interviews with case study students were conducted, with interview questions informed by trends observed in the weekly survey data. Throughout weeks 7-20, weekly data collection continued with the enhanced matching algorithm, allowing comparison between different intervention phases. Individual semi-structured interviews with each of the twelve case study students were conducted at three strategic points: at the beginning of the project (establishing baseline perceptions), middle of the project (capturing experiences with the training intervention), and close of the project (reflecting on the complete experience including the enhanced matching algorithm).

Data Analysis

Given the limited research on implementing peer learning environments centered around students teaching one another, this study utilized grounded theory approaches (Charmaz, 2014). We operationalized this concept by writing short analytic memos following each interview to capture key ideas, then analyzing themes across memos when constructing subsequent interview protocols. This allowed ideas meaningful to subjects to influence subsequent conversations, enabling others to confirm, disconfirm, and elaborate on emerging concepts.

With a dataset of 34 interviews, we employed "consensus" coding (Harry et al., 2005) rather than quantitative checks of Intercoder Reliability. Our analysis process followed an adaptation of Strauss and Corbin's (1998) levels of analysis.

For Level One (Open Coding), both authors collaboratively developed open codes by watching video recordings of each interview and coding its respective transcript. When authors applied codes differently, that moment of the interview was revisited and discussed until consensus was achieved. This process resulted in 176 'open codes.' In Level Two (Conceptual Categories), open codes were sorted into 10 'conceptual categories' (Harry et al., 2005).

Our mixed-methods design (Creswell & Plano Clark, 2018) integrated quantitative and qualitative data throughout the analysis process. Survey results informed subsequent interview protocols, while findings were validated through triangulation of student self-reports with tutee performance data. Qualitative insights provided explanatory context for quantitative results, particularly regarding how compatibility-based matching influenced tutoring interactions.

Ethical Considerations

This study was conducted in accordance with the ethical guidelines established by the university's Institutional Review Board, which approved the research protocol prior to implementation (Protocol #45383). Informed consent was obtained from parents/guardians of all participating students, and student assent was secured at the beginning of the study.

Several measures were taken to ensure ethical data collection and student wellbeing. All data were de-identified during analysis, with pseudonyms used in reporting to protect student privacy. Tutoring sessions were structured to minimize academic risk, with teacher oversight to ensure that inaccurate information was not propagated. The intervention was designed to enhance rather than disrupt normal classroom activities, with scheduling coordinated in consultation with classroom teachers to minimize interruption to regular instruction.

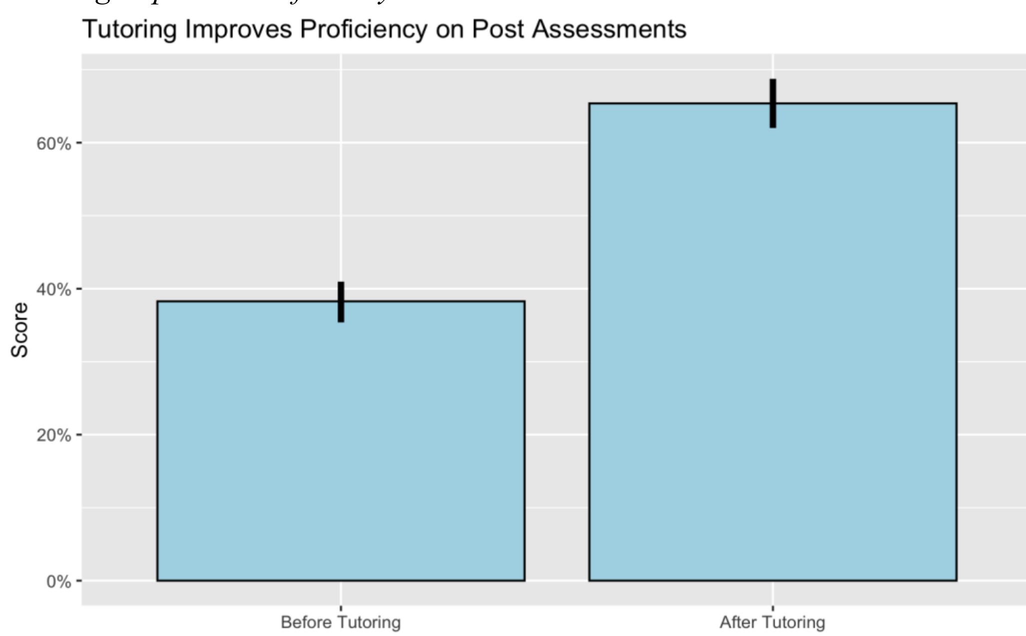
Results

Several indicators confirmed the efficacy of this tech-mediated training and matching system in creating productive, meaningful learning experiences. Five-question pre-assessments preceded 208 peer tutoring sessions, followed by equally challenging five-question post-assessments. Peer tutoring sessions were approximately twenty minutes and, as shown in Figure 4, they took students from a mean of 38% proficiency (SD = 20%) to a mean of 65% proficiency (SD = 25%). This improvement was statistically significant (paired t-test, $p < .01$) with a large effect size (Cohen's d = 1.20). Beyond measured learning gains, interviews yielded compelling revelations on how students thought about themselves, school, and peer learning.

While the previous development of PeerTeach utilized randomized controlled trials to establish efficacy (Rosier, 2023), the current implementation employed a design-based research approach without randomization or control groups. These pre/post assessment data should therefore be interpreted as promising evidence of effectiveness within this implementation context rather than causal confirmation.

Figure 4

Tutoring Improves Proficiency



Qualitative Analysis Overview

Qualitative analysis of the 34 interview transcripts yielded 10 conceptual categories reflecting students' experiences with technology-mediated peer tutoring. Through our iterative coding process, we identified 176 open codes that were ultimately organized into these conceptual categories. Table 1 presents these categories along with their most common subcodes and frequencies of occurrence. While these frequencies indicate topics that emerged prominently in our conversations with students, they reflect both student interests and our interview questioning. The substantive content of students' reflections, rather than mere frequency counts, provides the most valuable insights into how students experienced and made meaning of tech-mediated peer tutoring.

Table 1 Conceptual Categories from Interviews and the Most Common Subcodes for Each
Conceptual Categories from Interviews and the Most Common Subcodes for Each

Conceptual Categories	**Most Common Subcodes (with Frequency of Subcode Occurence)**	**Number of Subcodes**	**Number of Instances[2]**
Training Impact on Sessions	1. Training generally improved sessions (11) 2. Now I elicit my peer's thinking (10) 3. Now I check for understanding (9)	42	90
Conceptions of Good Tutoring	1. Patience is very important (7) 2. Good tutors prepare (7) 3. Good tutors explain well (6)	33	77
Comfort and Compatibility	1. Feels comfortable with peers (14) 2. More comfortable with friends than peers (9) 3. Comfort with peers over adults (6)	16	45
Benefits of Peer Tutoring	1. Learns by teaching (12) 2. Personalized attention is helpful (9) 3. Happy when making an impact (4)	15	41
Conceptions of Bad Tutoring	1. Bad tutors read off the slide (3) 2. Bad tutors rush (3)	17	22
Tutor Effort and Creativity	1. Inventing teaching strategies (4) 2. Gratitude messages to good tutors (3)	13	19
Preference for Tutor/Tutee Role	1. Prefers even tutor-tutee balance (5) 2. Prefers tutoring (5) 3. Enjoys being tutor and tutee (3)	8	19
Challenges of Remote Learning	1. Absent students hurt the experience (3)	9	17
Self-perceptions of Abilities	1. Self-efficacy with explaining (3) 2. Positive social self-perception (3)	6	10
Drawbacks of Peer Tutoring	No subcodes repeated	4	4

[2] Instances represent moments in interviews when a particular idea emerged. Whether mentioned in one sentence or several, the interview segment was coded as a single instance. In this table, the three most common subcodes are listed for each conceptual category. Subcodes with fewer than three instances were excluded

The rich descriptions students provided reveal nuanced understandings of their peer tutoring experiences. When discussing comfort with peers, students articulated how this psychological safety enabled risk-taking in learning, with one student explaining they could 'speak freely' with peers in ways impossible with adults. Students who described 'learning by teaching' detailed specific cognitive mechanisms, such as how preparation forced deeper understanding of material and how explaining concepts solidified their own knowledge. Similarly, students who reported 'eliciting peer thinking' and 'checking for understanding' described concrete changes in their tutoring approach following training, moving from immediate correction of mistakes to guiding peers toward self-discovery. These substantive insights form the foundation for our detailed findings presented in the following sections, organized by their relationship to our research questions.

Outcomes from Training: Addressing RQ1 and RQ2

Without preparation and reinforcement, students often engage in unproductive helping behaviors. Training is a necessity for successfully implementing a peer co-learning community, but is challenging and burdensome for teachers. To overcome these challenges, we used game-based modules to help students internalize evidence-based pedagogy. Students took up a variety of moves and every focal student reported employing at least one core teaching strategy from the training. Of these moves, the most adopted were eliciting peer's thinking and checking for understanding. Most interviewed students reported that they rarely checked for understanding prior to training yet now did so consistently.

In initial interviews with students, correcting mistakes was mentioned by five students as an important tutor role. Following the training and after accumulating more tutoring experience, students reported avoiding immediately correcting mistakes, instead encouraging tutees to find their own mistakes, and using mistakes as an opportunity to probe for misconceptions. As one student commented, "Before, the first thing that came to my mind was correcting the tutee and now the first thing is walking through the tutee's shoes and understanding why he or she is saying that this is the right method."

Another common trend was the shift from an overly didactic approach to a question-focused one. Before training, the most common response to "what makes a good tutor?" sounded like this answer: "First explaining, maybe a few times, and then giving the tutee a chance to clarify anything they didn't understand." After the training, the same question most yielded responses like, "First you have to ask questions to understand how much they know and after that you need to start explaining if they don't understand."

Similarly, students reported they began engaging in more perspective-taking, as they worked harder to understand tutees' beliefs and misconceptions. One student even connected his efforts to understand the tutee with his own confidence: "[before] I would immediately start explaining and then when the person would say, 'Hey, I don't know this concept' then your confidence will be like zero because I explained everything and that person still does not know anything." Following the training and his maturation as a tutor, he reported much higher confidence when explaining because he knew "what the other person knows about a particular topic." Confidence emerged as a key training consequence. As a second focal student expressed, "After the training, I could interact with the student with more confidence."

The training was many students' first opportunity to actively consider *how* they help others. This novelty spurred a degree of thoughtfulness and introspection that gave way to a multitude of metacognitive benefits beyond implementing particular teaching strategies. As one student explained, "Before the training I really didn't think much about what I can do better. But the training made me think about what I could do." Asking that question led to the development of new teaching strategies for several focal students that were not covered in the training.

These findings directly address RQ1 by identifying interactive training modules as a critical technological feature for overcoming the barrier of students' limited pedagogical skills. The significant shift in tutoring approaches—from didactic explanation to eliciting questions, from immediate correction to facilitated discovery—demonstrates that well-designed technological training can effectively develop specific tutoring behaviors without adding to teacher workload. Additionally, these findings address RQ2 by showing how this training contributed to sustainable practices through the development of metacognitive skills and increased confidence, which appear to support the emergence of what might be called a "helper identity" among students.

Outcomes from Compatible Matching: Addressing RQ1 and RQ2

In Week Seven, we began matching tutoring pairs based on both mastery differentials and compatibility. This shift dramatically affected student attitudes and comfort levels. As one student put it, after starting to match with compatibility, "Everybody is comfortable with the people they are getting... it has really helped." In fact, every focal student mentioned feeling comfortable with peers in one-on-one teaching sessions and this comfort, they pointed out, allowed for more question-asking and risk-taking. In general, students most often cited being able to "speak freely" as a key advantage to the matching shift.

Pairing students based on compatibility understandably raises concerns regarding possible off-task behavior. Critically, with this increased compatibility, students did not experience any slump in their after-tutoring assessments, suggesting that matching friends did not weaken learning despite students being essentially unmonitored in those sessions. Increased comfort between students had little downside and significant upside. For instance, one student commented, "The first reason [I felt comfortable] is that I was familiar with all three of my tutees, they were good friends of mine." Another explained, "Peer teaching has been easy because my partners are my friends and I know them and I've been talking to them for a few years."

The compatibility among partners increased noticeably over time because the matching algorithm became responsive to students' self-reported levels of compatibility. One unexpected byproduct of this shift was an increased frequency of same-gender pairs, which occurred in 47% of pairs before we initiated compatible matching, and in 70% of pairs after. Gender arose during interviews as a relevant factor for student comfort. One male subject noted that it can be "difficult to talk" to his female peers because he has less familiarity with them, making it a challenge to "speak as freely." Two other male subjects reported feeling comfortable partnering with girls themselves but commented that their male and female peers occasionally resented being paired across-gender. One said, "But if the tutor is a boy and he gets a tutee that's a girl, then he is like, 'I just got a tutee who was a girl and she's just dumb.'" These issues of same-gender comfort (and occasional downright sexism) could partly explain why, in previous studies, same-gender tutoring pairs tend to learn more. This was a key finding in a recent peer tutoring meta-analysis (Leung, 2015), suggesting teachers should think critically about competing goals of cross-gender interaction, learning, and peer interaction comfort.

These findings address RQ1 by identifying compatibility-based matching algorithms as a key technological feature for addressing implementation barriers. By automating the complex process of creating optimal tutoring pairs—considering both content knowledge differences and interpersonal compatibility—the technology reduced the logistical burden on teachers while creating conditions for productive learning interactions. The findings also address RQ2 by demonstrating how these compatibility-based matches contributed to sustainable practices by creating the psychological safety necessary for risk-taking and vulnerable learning behaviors, which Edmondson (1999) identifies as essential for effective collaborative learning.

Demonstrated Role of Technology: Addressing RQ1, RQ2, and RQ3

Technology was integral to fostering this co-learning community where students embraced the collective and collaborative aspects of math, a subject often perceived as individualistic. One pathway was removing the complex and time-consuming challenge of training effective peer helpers, accomplished through PeerTeach training modules. An additional technology affordance was automatically measuring and recognizing the learning progress which students were achieving through collective effort. To accomplish this, PeerTeach automatically generated post-assessments for learners, tabulated the growth facilitated by a peer, and then presented the class with success stories at the beginning of each class.

Four students reported that their "happiest moments" emerged when they were recognized for the learning progress which they had helped others achieve, which occasionally spurred messages of gratitude from helped students. Because this learning model automated the pre- and post-assessment process, it was easy for students to immediately grasp how consequential their interactions were for the learning of the tutee. An anecdote from one student underscores why such recognition of their tutoring success matters: "When he did the test, he called me and said, 'Thank you!' He was really happy. His percentage had grown very much. He felt the test went much easier… I've never taught that properly to anyone before, so it was a great feeling." This remark highlights the key motivational importance of making visible the impact from more formalized helping. It was often the greatest source of reported happiness for students, and it prompted gratitude sharing, which was additionally motivating. Finding ways to facilitate more gratitude sharing, in conjunction with the act of making growth more visible to each student, emerged as clear and realistic goals for future R&D work.

Teachers often report that the process of thoughtfully matching students can be very time-consuming. Beyond that, it is rarely driven by data informing which students are prepared to help and which students need help. None of this is surprising. The class wide combinatorics of tutor-tutee-topic triplets are all but infinite. Hence, predicting the most productive pairs is a data science puzzle of sizable magnitude with tremendous potential. By producing an optimization algorithm and fine-tuning its variable weights, PeerTeach became responsive to students' compatibility levels. As a result, our upgraded matching algorithm appeared to demonstrably improve students' comfort and risk taking.

These findings address all three research questions. For RQ1, they identify automated assessment and recognition systems as key technological features for addressing implementation barriers by reducing teacher workload while making learning progress visible. For RQ2, they demonstrate how these systems promote sustainable practices by creating positive interdependence—a key element of productive collaborative learning identified in our theoretical framework. For RQ3, they highlight how technology can enhance valued aspects of peer tutoring by making visible the impact of helping others, which students consistently identified as a significant source of satisfaction and motivation.

Why Students Value This Form of Tech-Mediated Peer Learning: Addressing RQ3

Co-learning communities can only evolve and sustain if members see their collective practices and mission as meaningful, producing outcomes which they care about (Wenger, 1999). A central focus of this study was to understand what students value about peer learning. Such insights can—as they did in this case—inform how practitioners and developers of interventions design peer-learning experiences and frame them to yield buy-in. Four months of consistently interviewing focal students offered a surprisingly clear and consistent set of outcomes students value: 1) learning as a byproduct of teaching, 2) deeper and more sustained attention when needing support, and 3) freedom to take risks among peers.

Learn by Teaching

Students quickly recognized the value of peer tutoring for their learning. This was the most frequently cited benefit of PeerTeach. Without direct prompting, nine of twelve focal students described that they value how much they learn when they teach, mentioning this equally often at the start and end of the project. This continuity suggests that the learning benefits of tutoring are obvious, a priori, and reinforced through experience.

Teaching is a powerful learning mechanism confirmed in numerous studies—often finding tutors have learned more than their tutees (Topping, 2017). There are several key theories for what drives tutor learning: preparing to teach (Bargh and Schul, 1980), verbalizing one's understandings (Coleman, Brown, & Rivkin, 1997), and scaffolding while monitoring tutee errors (Topping & Ehly, 2001; Topping, 2005). While there are several theories regarding how teaching drives learning for the tutor, students in this study repeatedly reported that it was the preparation phase that pushed them to understand the math content more deeply. One student explained, "When the tutor has to teach, the tutor also has to study the topic very properly." Another observed, "While making the PowerPoint Presentations, we have to find some real-life examples so that makes us research more; then we make examples ourselves and that also makes the tutor learn more." These statements echo Bargh and Schul (1980), who first established this powerful connection between learning and preparing to teach. Students also cited other mechanisms well-documented in the literature—one noted, "When I check that the answers are correct, my math gets revised." This comment echoes what Topping and Ehly (2001) describe as learning through scaffolding and error management.

More Personalized Attention

Focal students consistently described this peer learning experience as providing more attention than they ever receive in the classroom. While students expressed genuine understanding of teachers' limitations to address all student needs, most focal students characterized teacher attention as limited while peer tutors could offer their full attention. As one put it, "Sometimes it happens where a specific student is paid more attention to by a teacher, not purposefully, but it happens. With one-on-one, the tutor has total attention on the tutee, so it's better." In similar fashion, another student commented, "A teacher can't reach out to all the students who have [problems] wrong. He can only do it one-by-one and a peer can always be there." Similar sentiments arose repeatedly. Several students additionally noted that a key benefit to having a peer's full attention was the freedom to ask more questions. As one explained, "We can't just ask anything to teachers. Like that would waste their time and also teachers can't keep repeating stuff for us again and again."

Increased Comfort

Our grounded theory approach was committed to uncovering what students found most meaningful about a peer learning community. Though not an explicit research topic initially, 'comfort' emerged as one of the most salient constructs, 'compatibility' its antecedent, and free inquiry, questioning, and risk-taking its consequences.

Many students felt that being taught by peers is much more comfortable than being taught by teachers. Students noted that it is much easier to ask peers questions, because they are given personalized attention, and since it is less intimidating to seek help from someone their age. While students generally perceived their current teachers as kind, past experiences gave them reason for fear in whole group contexts. For instance, one student recounted, "I've had some teachers who, even for a little difficult question, get really angry and are quite short tempered… but when it comes to a peer, I'm never afraid to ask questions." Similar stories were shared by two others. As student comfort increased during this study, the impact of peer teaching sessions extended beyond the classroom as students reported forming chat groups outside of school to discuss math.

These findings directly address RQ3 by identifying learning through teaching, receiving personalized attention, and experiencing psychological safety as the aspects of peer tutoring students value most. They also show how technology enhanced these valued elements through facilitated preparation processes, dedicated time for one-on-one interaction, and compatibility-based matching that promoted comfortable learning environments. The technological system effectively amplified these naturally occurring benefits of peer tutoring by removing organizational barriers and creating optimal conditions for their emergence.

Discussion

Theoretical Insights

Vygotskian Mechanisms in Peer Tutoring Interactions

Our findings provide empirical support for the Vygotskian mechanisms theorized to underlie effective peer tutoring. The shift in tutoring behaviors reported by students—from immediate correction to facilitated discovery, from didactic explanation to eliciting questions—exemplifies the transition from traditional transmission approaches to scaffolding within the Zone of Proximal Development. As students became more aware of their tutees' thinking processes, they were better able to provide the contingent support necessary for effective learning within the ZPD.

These behavioral shifts align with Chi's (2009) interactive-constructive-active-passive (ICAP) framework, which posits that interactive learning activities yield deeper conceptual understanding than constructive, active, or passive approaches. The PeerTeach training pushed students from primarily passive or active tutoring behaviors (direct explanation, rote demonstration) toward more interactive approaches (dialogic questioning, co-construction of meaning). This shift not only benefited tutees but also deepened tutors' own understanding through the elaborative processes required for effective explanation (Fiorella & Mayer, 2013).

The reported metacognitive benefits of training further support Vygotsky's (1978) concept of internalization. As students externalized their thinking about effective teaching, they began to develop an internal model of effective pedagogical practice—what Sherin (2005) describes as "professional vision" for teaching. This developing vision appeared to transfer beyond the specific strategies taught in the training, as evidenced by students' reports of inventing new teaching approaches. This emergent capacity for pedagogical problem-solving suggests that the training facilitated not just the adoption of specific strategies but the beginnings of a more general ability to recognize and respond to learning needs—a key aspect of Vygotskian instruction.

Lewinian Dynamics in Peer Learning Communities

The dramatic impact of compatibility-based matching on student comfort and risk-taking provides compelling evidence for Lewin's emphasis on psychological safety as a prerequisite for productive learning interactions. As Edmondson (1999) has demonstrated in organizational contexts, psychological safety—the shared belief that one can take interpersonal risks without negative consequences—is essential for the vulnerable acts of asking questions, admitting confusion, and engaging in productive dialogue. Our findings suggest that this principle operates similarly in educational settings, with student reports of "speaking freely" and increased question-asking directly linking comfort to learning behaviors.

The importance of making learning progress visible and celebrating collective achievement directly supports Lewin's concept of positive interdependence. By creating technological structures that demonstrated the mutual benefits of the tutoring relationship, we were able to address one of the primary motivational challenges of peer tutoring identified in our theoretical framework. Students' reports of happiness when seeing tutees' progress suggests that these structures successfully created what Johnson and Johnson (2009) call "promotive interaction," where participants actively encourage and facilitate each other's efforts toward group goals.

The gender dynamics observed in compatibility-based matching offer an important nuance to our understanding of field forces in peer learning contexts. The increased frequency of same-gender pairings following the introduction of compatibility metrics, coupled with student reports of discomfort in mixed-gender interactions, suggests that gender remains a powerful social force in classroom interactions. This finding aligns with research on gender homophily in classroom networks (Mehta & Strough, 2009) and highlights the complex interplay between social identity, interpersonal comfort, and learning outcomes that must be considered in peer learning design.

Addressing Research Questions

Design Features for Implementation Barriers

Our primary research question asked how technology-mediated peer tutoring environments should be designed to overcome key implementation barriers. The findings suggest three critical design features: (1) interactive training that builds specific pedagogical skills, (2) data-driven matching algorithms that balance content knowledge and interpersonal compatibility, and (3) automated assessment and recognition systems that make learning progress visible.

The PeerTeach training modules addressed the barrier of students' limited pedagogical skills by providing explicit modeling, guided practice, and feedback on key tutoring behaviors. The particular effectiveness of simulation-based training elements, where students practiced teaching a virtual character, aligns with Gee's (2003) principles of effective learning environments, where learners can experiment with new identities and approaches in low-risk contexts before transferring these skills to higher-stakes real-world interactions.

The matching algorithms addressed both logistical and pedagogical barriers by automating the complex process of creating optimal tutoring pairs while ensuring that all students had opportunities to serve as tutors. This balanced approach to role assignment appears to have contributed to the high levels of student engagement and the absence of the "free rider" problem often observed in collaborative learning contexts (Salomon & Globerson, 1989). The integration of compatibility data further enhanced these matches by creating conditions for psychological safety and productive risk-taking.

The automated assessment and recognition systems addressed motivational barriers by making the reciprocal benefits of peer tutoring visible to all participants. By providing immediate feedback on tutoring effectiveness through pre/post assessment comparisons, these systems created what Slavin (1983) calls "equal opportunities for success," where all students could experience recognition for contributing to others' learning regardless of their overall achievement level.

Development of Sustainable Practices

Our second research question asked how technology-mediated design solutions influence the development of sustainable peer tutoring practices. The findings suggest that sustainability emerges from the interplay of skill development, positive emotional experiences, and structural support.

The metacognitive benefits reported by students—increased confidence, deeper reflection on helping strategies, development of new approaches—suggest that the training facilitated not just the adoption of specific techniques but the development of what might be called a "helper identity." This identity development is crucial for sustainability, as it shifts peer tutoring from an externally imposed activity to an internalized practice consistent with students' self-concept. As Wenger (1998) argues in his communities of practice framework, sustainable learning communities emerge when participants see the practices as meaningful extensions of their identities.

The positive emotional experiences reported by students—from the satisfaction of seeing tutees improve to the comfort of working with compatible peers—further contributed to sustainability by creating what Fredrickson (2001) calls "upward spirals" of positive emotion and engagement. Students' reports of continuing mathematical discussions outside of class and forming study groups suggest that these positive experiences were beginning to generalize beyond the formal tutoring sessions, potentially creating self-sustaining peer learning communities.

The structural support provided by the technology—from automated matching to recognition of achievement—reduced the organizational burden that typically limits the sustainability of peer tutoring initiatives (Greenwood et al., 2002). By handling the logistical aspects of peer tutoring implementation, the technology allowed teachers and students to focus on the substantive aspects of the learning interactions, potentially increasing the likelihood of continued implementation.

Valued Aspects of Peer Tutoring

Our third research question asked what aspects of peer tutoring students value most and how technology can enhance these elements. The findings reveal three primary values: learning through teaching, receiving personalized attention, and experiencing psychological safety. Students' consistent emphasis on learning through teaching aligns with research on the "protégé effect" (Chase et al., 2009), which suggests that the responsibility of teaching others can motivate deeper engagement with content than learning for oneself. The particular emphasis students placed on preparation as a learning mechanism extends previous research on tutor learning by highlighting the motivational power of the "responsibility to teach" as a driver of thorough preparation.

The value placed on personalized attention reflects research on the benefits of individualized instruction (Bloom, 1984) and highlights a key advantage of peer tutoring over traditional classroom instruction. Students' appreciation for the "full attention" provided by peers underscores what Cohen (1994) identifies as the "resource dilution" problem in traditional classrooms, where teacher attention must be distributed across many students.

The importance of psychological safety—expressed through students' emphasis on comfort, freedom to ask questions, and lack of judgment—underscores the social-emotional dimensions of effective learning environments. The salience of comfort in students' reflections challenges the common emphasis on cognitive aspects of learning environments to the exclusion of affective dimensions. As Immordino-Yang and Damasio (2007) argue, emotion and cognition are inextricably linked in learning, with emotional processes guiding attention, engagement, and memory formation.

Design Challenges

Despite the overall success of the intervention, several design challenges emerged that warrant consideration in future iterations:

Role Balance and Identity

Students noted several drawbacks of this peer learning setup. One tutor mentioned fear of giving wrong information. Supporting this, another student asserted, "It's really difficult to doubt what the teacher is saying but, when I'm getting tutored by a peer, I know that there is a high chance of the peer making mistakes." As a result, we transitioned from the language of "tutor teaching tutee" to "coach helping player," to alleviate pressure to be all-knowing or faultless. These titles resonated with students as they imbued status to both parties; coach as well as player carries positive connotations, and a sense of guidance rather than perfect accuracy.

Additionally, some students complained that they were not tutors more often. Following this revelation, a question was added to the post-survey to gauge how often students wanted to serve in each role. On average, students wanted to tutor 58% of the time; only one out of 51 respondents indicated wanting to be tutor fewer than three times out of ten. This common desire for balanced role-taking suggests matching algorithms should strongly weight balanced roles. Fine-tuning the PeerTeach matching algorithm to create better balance—without sacrificing tutoring quality—is a technical challenge for future iterations. In addition to wanting parity, two students who primarily tutored recognized that being tutored could improve their own tutoring abilities. One explained, "I would like to give other people a chance as well to teach me and I could get innovative ideas from them." While there are clear learning benefits to selecting tutors with topic mastery (Rosier, 2023), there is also clear value to maintaining co-learning communities where all fully engage in both roles.

This finding connects to research on the importance of identity development in peer learning contexts (Wortham, 2006). The opportunity to take on the role of "expert" or helper appears to be valuable not just for the learning benefits it provides but also for the positive identity it confers. This suggests that technology-mediated peer tutoring environments should be designed to provide balanced opportunities for all students to experience this identity, regardless of their overall achievement level.

Gender Dynamics and Social Complexity

The emergence of gender as a factor in compatibility-based matching highlights the complex social dynamics that influence peer learning interactions. While the increased frequency of same-gender pairs appeared to enhance comfort and risk-taking, it also potentially reinforced existing social divisions and limited opportunities for cross-gender collaboration. This tension between immediate psychological safety and broader social integration goals represents a significant design challenge.

This challenge connects to broader questions about how educational technologies can effectively navigate complex social dynamics without inadvertently reinforcing problematic patterns. As Baker et al. (2022) have argued, seemingly neutral algorithmic systems can amplify existing social biases if these biases are not explicitly addressed in the design process. Future iterations of compatibility-based matching algorithms might need to incorporate mechanisms for gradually expanding students' comfort zones while still maintaining sufficient psychological safety for productive learning interactions.

Balancing Structure and Agency

The structured nature of the PeerTeach intervention—with its predetermined roles, topics, and assessment processes—appeared to support effective implementation but potentially limited student agency in shaping their own learning interactions. This tension between structure and agency represents a recurring challenge in educational technology design (Scardamalia & Bereiter, 2014).

Future iterations might explore how to maintain the benefits of structured support while providing increasing opportunities for student choice and self-direction as they develop greater competence in peer tutoring. This progression could potentially follow a scaffolding approach, with technological structures gradually fading as students internalize effective tutoring practices and develop greater autonomy in managing their own learning interactions.

Implications for Practice

This study demonstrates technology can play a central role in organizing and amplifying the power of peer tutoring, helping teachers overcome key challenges of Group Learning (GL): students lacking interactional skills, groups lacking positive interdependence, and the logistical challenges of facilitating complex interpersonal dynamics.

Student Collaboration Skills

The game-based PeerTeach training modules laid the foundation for effective helping during this design-based research study. The training had clear take-up for critical pedagogical moves, successfully shifting students from overly didactic teaching to student-centered learning. Beyond specific skills, training also promoted a sense of confidence, making the experiences more gratifying. With few existing digital tools for training effective interactional behavior, the frontiers are open for technologists and learning scientists to develop increasingly responsive training systems. With advancing technology in the areas of automated transcription and natural language processing, it is realistic to envision an intelligent system unobtrusively monitoring and evaluating students' interactional behaviors and assisting their development as tutors through real-time support or targeted micro trainings.

Positive Interdependence

Assessments within cooperative learning are risky if uncritically implemented. Webb, Pierce, and Mastergeorge (2010) found that grading the direct products of GL activities promoted giving answers, copying, and shifting focus away from mastery. Performance orientation, as opposed to mastery orientation, has been shown to hinder intrinsic motivation, learning interest, and interactional behaviors (Elliot & Dweck, 2005; Topping, 2017). The present study suggests that if peer tutoring work itself is ungraded, subsequent assessment does not negatively impact collaboration dynamics. This parallels Slavin's (1983) assertion that evaluation is motivating and not detrimental to collaboration so long as end-products could not be individually completed—providing what Cohen and Lotan (1997) call group-worthy tasks. In this sense, peer tutoring followed by assessment is inherently group worthy as the task cannot be successfully completed without both students successfully executing their roles.

The responsibility of helping others learn can engender a powerful sense of duty—what Chase et al. (2009) dub the protege effect—and motivation is amplified by systems recognizing the degree to which peer tutors successfully helped tutees. This assessment loop and the act of making progress visible are key technology affordances promoting student buy-in.

Being able to predict compatible pairings between students using algorithms and reward functions yields affective experiences that sustain peer learning and promote comfort and risk-taking, each fostering learning. The construct of 'comfort' should not be taken lightly as a superficial aspect of collaboration with minor implications. For many students, it is a key determinant for their learning. For instance, in a study analyzing the key factors predicting learning in small groups, Theobald et al. (2017) found that students who "felt more comfortable" in their group showed a 28% increase in content mastery over those who did not.

Implementation

For teachers who attempt to create thoughtful heterogeneous groups, the time burden is problematically large (Jolliffe & Snaith, 2017). Additionally, when teachers play matchmaker by deputizing some students to help others, they typically only choose high-performing students for receiving this privilege. This bias not only denies lower-performing students the learning benefits of teaching, it also imbues and reifies status for a select few, in a Matthew Effect of "rich get richer" (Rigney, 2010). Data-driven matching algorithms identify topics that generally-struggling students can teach, promoting collaboration that is empowering to all students. It seems likely that employing algorithmic matching is the only way to break free of teachers' unconscious favoritism when promoting peer helping and to create more balanced helping opportunities while still utilizing data to inform who needs help and who is ready to give it.

Limitations

This study has several methodological limitations that should be considered when interpreting the findings. The convenience sampling approach limits generalizability beyond similar educational contexts. The participating students were from an Indian middle school, and cultural factors specific to this setting may have influenced implementation and outcomes in ways that might not transfer to other contexts.

The focus on mathematics as the subject domain represents another limitation. Mathematics content tends to be more structured and sequential than other subjects, potentially facilitating peer tutoring processes in ways that might not translate to more open-ended domains like literature or social studies. The online implementation due to COVID-19 restrictions created both affordances and constraints that may differ from in-person peer tutoring.

Our reliance on self-reported data for understanding students' tutoring behaviors introduces potential biases related to social desirability and recall accuracy. While we triangulated these reports with performance data, direct observation of tutoring interactions would have provided more objective measures of behavioral changes. Additionally, the pre-post assessments focused on immediate content mastery rather than long-term retention or transfer, leaving questions about the durability of learning gains.

As researchers who developed the PeerTeach platform, our dual roles may have introduced bias in both implementation and interpretation. Though we employed strategies to enhance objectivity, including consensus coding and triangulation across data sources, our investment in the platform's success should be acknowledged.

Strengths

This study contributes to educational technology and peer learning literature through several key strengths. The research applies established theoretical frameworks to address persistent implementation challenges in peer tutoring. The design effectively operationalizes Vygotskian and Lewinian principles in technological features that support effective tutoring interactions and promote psychological safety.

The mixed-methods approach provides complementary perspectives on both intervention outcomes and implementation processes. The quantitative data demonstrates learning gains, while the qualitative data illuminates the experiential dimensions that contributed to these outcomes. The longitudinal qualitative component offers insights into how student perceptions and practices evolved throughout the intervention.

The implementation within an authentic classroom setting enhances the ecological validity of the findings compared to laboratory studies. The iterative, design-based research approach allowed for responsive refinement based on emerging findings, creating a more robust intervention while generating design principles with potential applicability beyond the immediate context.

The purposive sampling strategy for case study students ensured representation across performance levels and gender, providing insight into how the intervention affected diverse learners. This approach offers a more nuanced understanding of how different students experienced and benefited from technology-mediated peer tutoring.

A distinctive contribution of this study is the development and classroom deployment of a compatibility-informed matching algorithm, which intelligently pairs students using both content mastery and peer-reported collaboration data. This approach advances existing work on grouping strategies by offering a scalable, data-driven alternative to teacher intuition or random assignment. By integrating social dynamics into algorithmic decision-making, the system promotes psychological safety and productive peer interactions—factors shown to significantly impact learning outcomes but rarely operationalized in educational technologies.

Conclusion

Studies have shown that effective school-based tutoring ideally occurs 1) often and consistently, 2) during the school day with active classroom integration, and 3) in tandem with data collection for more tailored instruction (Nickow, Oreopoulos, & Quan, 2020; Robinson et al., 2021). Each factor allows tutors to build on classroom experiences to reinforce relevant, timely skills through just in time learning (Novak et al., 1999). But school-based tutoring via adult professionals creates enormous expense and logistical challenges and precludes valuable learning opportunities that emerge when students teach.

PeerTeach offers proof of concept that teachers can facilitate peer tutoring in their own classrooms when technology supports the complex processes of training, assessing, matching, and facilitating sessions. Technology-enabled peer tutoring produced large learning gains (Cohen's $d = 1.20$) and consistent post-session improvement across 208 peer tutoring sessions. Additionally, qualitative findings revealed a powerful shift in how students viewed themselves and each other—as confident helpers, motivated learners, and reflective thinkers. Peer tutors began adopting learner-centered strategies, tutees felt more comfortable asking questions, and students across performance levels reported a stronger sense of belonging and purpose.

By automating time-consuming instructional tasks and promoting evidence-based collaboration skills, PeerTeach appears to meaningfully reduce teacher burden while addressing long-standing barriers to Group Learning. These outcomes signal that technology can do more than enhance instruction—it can orchestrate peer learning environments that are scalable, equitable, and deeply human. This model offers a compelling path forward for classrooms seeking to transform peer interaction from an afterthought into a core engine of academic and social growth.

References

Abrami, P. C., Poulsen, C. & Chambers, B. (2004). Teacher motivation to implement an educational innovation: Factors differentiating users and non-users of cooperative learning. *Educational Psychology*, 24(2), 201–216.

Allor, J., & McCathren, R. (2004). The efficacy of an early literacy tutoring program implemented by college students. *Learning Disabilities Research & Practice*, *19*(2), 116-129.

Ander, R., Guryan, J., & Ludwig, J. (2016). *Improving academic outcomes for disadvantaged students: Scaling up individualized tutorials.* The Hamilton Project–Brookings.

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher*, 41(1), 16-25.

Aronson, E. (1978). *The Jigsaw Classroom*. Sage.

Bargh, J. A., & Schul, Y. (1980). On the cognitive benefits of teaching. *Journal of Educational Psychology*, *72*(5), 593.

Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. *International journal of artificial intelligence in education*, 1-41.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6), 4-16. https://doi.org/10.3102/0013189X013006004

Bowman-Perrott, L., Davis, H., Vannest, K., Williams, L., Parker, R., & Greenwood, C. (2013). Academic benefits of peer tutoring: A meta-analytic review of single-case research. *School Psychology Review*, 42(1), 39-55.

Bryant, D. P., Bryant, B. R., Roberts, G., Vaughn, S., Pfannenstiel, K. H., Porterfield, J., & Gersten, R. (2011). Early numeracy intervention program for first-grade students with mathematics difficulties. *Exceptional children*, *78*(1), 7-23.

Buchs, C., Filippou, D., Pulfrey, C., & Volpé, Y. (2017). Challenges for cooperative learning implementation: Reports from elementary school teachers. *Journal of education for teaching*, *43*(3), 296-306.

Chapin, S. H., & Anderson, N. C. (2013). *Classroom discussions in math: A teacher's guide for using talk moves to support the Common Core and more*. Math Solutions Publications.

Charmaz, K. (2014). *Constructing grounded theory*. Sage.

Chase, C. C., Chin, D. B., Oppezzo, M. A., & Schwartz, D. L. (2009). Teachable agents and the protégé effect: Increasing the effort towards learning. *Journal of science education and technology*, *18*(4), 334-352.

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1(1), 73-105. https://doi.org/10.1111/j.1756-8765.2008.01005.x

Chi, M. T. H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, 25(4), 471-533. https://doi.org/10.1207/s15516709cog2504_1

Clarke, A. E. (2005). Pushing and being pulled around the postmodern turn. *Situational analysis: Grounded theory after the postmodern turn*, pp. 1-36. Sage.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, 64(1), 1-35. https://doi.org/10.3102/00346543064001001

Cohen, E. G., & Lotan, R. A. (1997). *Working for equity in heterogeneous classrooms: Sociological theory in practice*. Teachers College Press.

Cohen, E. G., Lotan, R. A., Scarloss, B. A., & Arellano, A. R. (1999). Complex instruction: Equity in cooperative learning classrooms. *Theory into practice*, 38(2), 80-86.

Coleman, E. B., Brown, A. L., & Rivkin, I. D. (1997). The effect of instructional explanations on learning from scientific texts. *The Journal of the Learning Sciences*, 6(4), 347-365.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 32(1), 5-8. https://doi.org/10.3102/0013189X032001005

Dietrichson, J., Bøg, M., Filges, T., & Klint Jørgensen, A. M. (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.

Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), Collaborative learning: Cognitive and computational approaches (pp. 1-19). Elsevier.

Falchikov, N. (2001). *Learning together: Peer tutoring in higher education*. Routledge.

Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology*, 38(4), 281-288. https://doi.org/10.1016/j.cedpsych.2013.06.001

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, 56(3), 218-226. https://doi.org/10.1037/0003-066X.56.3.218

Fryer Jr, R. G. (2017). The production of human capital in developed countries: Evidence from 196 randomized field experiments. *In Handbook of economic field experiments* (Vol. 2, pp. 95-322). North-Holland.

Edmonds, J. (1965). Paths, trees, and flowers. *Canadian Journal of mathematics*, 17, 449-467.

Edmondson, A. C. (1999). *Psychological safety and learning behavior in work teams*. Administrative Science Quarterly, 44(2), 350–383. https://doi.org/10.2307/2666999

Elliot, A. J., & Dweck, C. S. (2005). Competence and motivation. *Handbook of competence and motivation*, 3-12.

Escueta, M., Nickow, A. J., Oreopoulos, P., & Quan, V. (2020). Upgrading education with technology: Insights from experimental research. *Journal of Economic Literature*, *58*(4), 897-996.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. Palgrave Macmillan.

Gillies, R. M. & Ashman, A. F. (1996). Teaching collaborative skills to primary school children in classroom-based work groups. *Learning and Instruction*, *6*(3), 187–200.

Goldman, S., Pea, R., & Maldonado, H. (2004). Emerging Social Engineering in the Wireless Classroom. In Kafai, Y. B., Sandoval, W. A., Enyedy, N., Nixon, A. S., & Herrera, F. (Eds.), International Conference of the Learning Sciences 2004: Embracing Diversity in the Learning Sciences (pp. 222-229). Santa Monica, CA: Lawrence Erlbaum Associates.

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology*, 9(6), 495-522.

Grubbs, N. (2009). The effects of the peer tutoring program: An action research study of the effectiveness of the peer tutoring program at one suburban middle school. *Georgia School Counselors Association Journal*, 16(1), 21-31.

Harry, B., Sturges, K. M., & Klingner, J. K. (2005). Mapping the process: An exemplar of process and challenge in grounded theory analysis. *Educational researcher*, 34(2), 3-13.

Immordino-Yang, M. H., & Damasio, A. (2007). We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, Brain, and Education*, 1(1), 3-10. https://doi.org/10.1111/j.1751-228X.2007.00004.x

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher*, 38(5), 365–379. https://doi.org/10.3102/0013189X09339057

Jolliffe, W. (2015). Developing cooperative learning pedagogy in initial teacher education. In R. Gillies (Ed.), Collaborative learning: Developments in research and practice (pp. 175–200). New York: Nova Science.

Jolliffe, W., & Snaith, J. (2017). Developing cooperative learning in initial teacher education: indicators for implementation. *Journal of education for teaching*, 43(3), 307-315.

Kim, E. J., Goodman, J., & West, M. R. (2025). Kumon in: The recent, rapid rise of private tutoring centers. *Education Finance & Policy*. Advance online publication. https://doi.org/10.1162/edfp_a_00438

King, A. (1997). ASK to THINK-TEL WHY: A model of transactive peer tutoring for scaffolding higher level complex learning. *Educational Psychologist*, 32(4), 221-235.

Kobayashi, K. (2019). Learning by preparing-to-teach and teaching: A meta-analysis. *Japanese Psychological Research*, *61*(3), 192-203.

Lepper, M. R., & Woolverton, M. (2002). The wisdom of practice: Lessons learned from the study of highly effective tutors. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 135-158). Academic Press.

Leung, K. C. (2015). Preliminary empirical model of crucial determinants of best practice for peer tutoring on academic achievement. *Journal of Educational Psychology*, 107(2), 558–579.

Lewin, K. (1948). Resolving social conflicts; selected papers on group dynamics.

Lewin, K. (1951). *Field theory in social science: Selected theoretical papers*. Harper & Brothers.

Mahenthiran, S., & Rouse, P. J. (2000). The impact of group selection on student performance and satisfaction. *International Journal of Educational Management,* 14(6): 255-265

Maldonado, H., Klemmer, S., & Pea, R. (June, 2009). When is collaborating with friends a good idea? Insights from design education. Proceedings of the 9th International Conference on Computer Supported Collaborative Learning.

Mehta, C. M., & Strough, J. (2009). Sex segregation in friendships and normative contexts across the life span. *Developmental Review*, 29(3), 201-220.

Michaels, S., & O'Connor, C. (2015). Conceptualizing talk moves as tools: Professional development approaches for academically productive discussion. *Socializing intelligence through talk and dialogue*, 347-362.

Mitchell, S. N., Reilly, R. C., Bramwell, G., Slonosky, A., & Lilly, F. R. (2004). Friendship and choosing groupmates: Preferences for teacher-selected versus student selected groupings. *Journal of Instructional Psychology*, (31), 20-32.

Nickow, A., Oreopoulos, P., & Quan, V. (2020). The impressive effects of tutoring on PreK-12 learning: A systematic review and meta-analysis of the experimental evidence. (NBER Working Paper No. 27476). National Bureau of Economic Research.

Novak, G. M., Patterson, E. T., Gavrin, A. D., & Christian, W. (1999). *Just in time teaching*. Prentice–Hall.

Pianta, R. C., Belsky, J., Houts, R. & Morrison, F. (2007). Opportunities to learn in America's elementary classrooms. *Science*, *315*, 1795–1796. doi:10.1126/science.1139719.

Rigney, D. (2010). The Matthew effect: How advantage begets further advantage. Columbia University Press.

Robinson, C. D., Kraft, M. A., Loeb, S., & Schueler, B. E. (2021). Accelerating Student Learning with High-Dosage Tutoring. EdResearch for Recovery Design Principles Series. *EdResearch for Recovery Project*.

Robinson, Carly D., and Loeb, Susanna. (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

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. Oxford University Press.

Roscoe, R. D., & Chi, M. T. H. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research*, 77(4), 534–574.

Rosier, S. (2023). PeerTeach: Teaching Learners to Do Learner-Centered Teaching. In: Noroozi, O., De Wever, B. (eds) *The Power of Peer Learning: Social Interaction in Learning and Development*. Springer, Cham. https://doi.org/10.1007/978-3-031-29411-2_11

Saldaña, J. (2021). *The coding manual for qualitative researchers*. Sage.

Salomon, G., & Globerson, T. (1989). When teams do not function the way they ought to. *International journal of Educational research*, 13(1), 89-99.

Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 397-417). Cambridge University Press.

Sharan, Y. (2010). Cooperative learning for academic and social gains: Valued pedagogy, problematic practice. *European Journal of Education*, 45(2), 300-313.

Sharan, Y., & Sharan, S. (1992). Expanding cooperative learning through group investigation (Vol. 1234). New York: Teachers College Press.

Sherin, M., & van Es, E. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of technology and teacher education*, 13(3), 475-491.

Slavin, R. E. (1978). Student teams and achievement divisions. *Journal of research and development in education*, 12(1), 39-49.

Slavin, R. E. (1983). When does cooperative learning increase student achievement?. *Psychological bulletin*, 94(3), 429.

Slavin, R. E. (1986). Learning Together. *American Educator: The Professional Journal of the American Federation of Teachers*, 10(2), 6-11.

Slavin, R. E. (2006). Educational Psychology Theory and Practice (8th edition). In *USA: Library of congress Cataloging in Publication Data*.

Slavin, R. E. (2015). Cooperative learning in elementary schools. Education 3-13, 43(1), 5-14. https://doi.org/10.1080/03004279.2015.963370

Stahl, G., Koschmann, T., & Suthers, D. (2014). Computer-Supported Collaborative Learning. In R. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (Cambridge Handbooks in Psychology, pp. 479-500). Cambridge: Cambridge University Press. doi:10.1017/CBO9781139519526.029

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research techniques*. Sage.

Theobald, E. J., Eddy, S. L., Grunspan, D. Z., Wiggins, B. L., & Crowe, A. J. (2017). Student perception of group dynamics predicts individual performance: Comfort and equity matter. *PloS one,* 12(7), e0181336.

Thurston, A., Duran, D., Cunningham, E., Blanch, S., & Topping, K. (2009). International on-line reciprocal peer tutoring to promote modern language development in primary schools. *Computers & Education*, *53*(2), 462-472.

Topping, K. J., & Ehly, S. W. (2001). Peer assisted learning: A framework for consultation. *Journal of Educational and Psychological Consultation*, 12(2), 113-132.

Topping, K., Buchs, C., Duran, D., & Van Keer, H. (2017). *Effective peer learning: From principles to practical implementation*. Routledge.

Topping, K. J. (2005). Trends in peer learning. *Educational psychology*, *25*(6), 631-645.

Topping, K., Buchs, C., Duran, D., & Van Keer, H. (2017). *Effective peer learning: From principles to practical implementation*. Routledge.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, Mass.: Harvard University Press.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, 53(4), 5-23. https://doi.org/10.1007/BF02504682

Webb, N. M., & Farivar, S. (2014). Developing productive group interaction in middle school mathematics. In *Cognitive perspectives on peer learning* (pp. 117-149). Routledge.

Wenger, E. (1999). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press.

Weston, C., Gandell, T., Beauchamp, J., McAlpine, L., Wiseman, C., & Beauchamp, C. (2001). Analyzing interview data: The development and evolution of a coding system. *Qualitative sociology*, 24(3), 381-400.

Wortham, S. (2006). *Learning identity: The joint emergence of social identification and academic learning*. Cambridge University Press.

Zajac, R. J., & Hartup, W. W. (1997). Friends as coworkers: Research review and classroom implications. *The Elementary School Journal*, 98(1), 3-13.