Pair-Up: Prototyping Human-AI Co-orchestration of Dynamic Transitions between Individual and Collaborative Learning in the Classroom

Kexin Bella Yang
kexiny@cs.cmu.edu
Human-Computer Interaction Institute, Carnegie Mellon University
Pittsburgh, Pennsylvania, USA

Vanessa Echeverria
Monash University
Clayton, Australia
Escuela Superior Politecnica del Litoral
Guayaquil, Ecuador
vanechev@espol.edu.ec

Zijing Lu
Human-Computer Interaction Institute, Carnegie Mellon University
Pittsburgh, USA
zijinglu@andrew.cmu.edu

Hongyu Mao
Carnegie Mellon University
Pittsburgh, USA
hongyum@andrew.cmu.edu

Kenneth Holstein
Human-Computer Interaction Institute, Carnegie Mellon University
Pittsburgh, USA
kjholste@andrew.cmu.edu

Nikol Rummel
Institute of Educational Research, Ruhr-Universität Bochum
Bochum, Germany
nikol.rummel@rub.de

Vincent Aleven
Human-Computer Interaction Institute, Carnegie Mellon University
Pittsburgh, USA
aleven@cs.cmu.edu

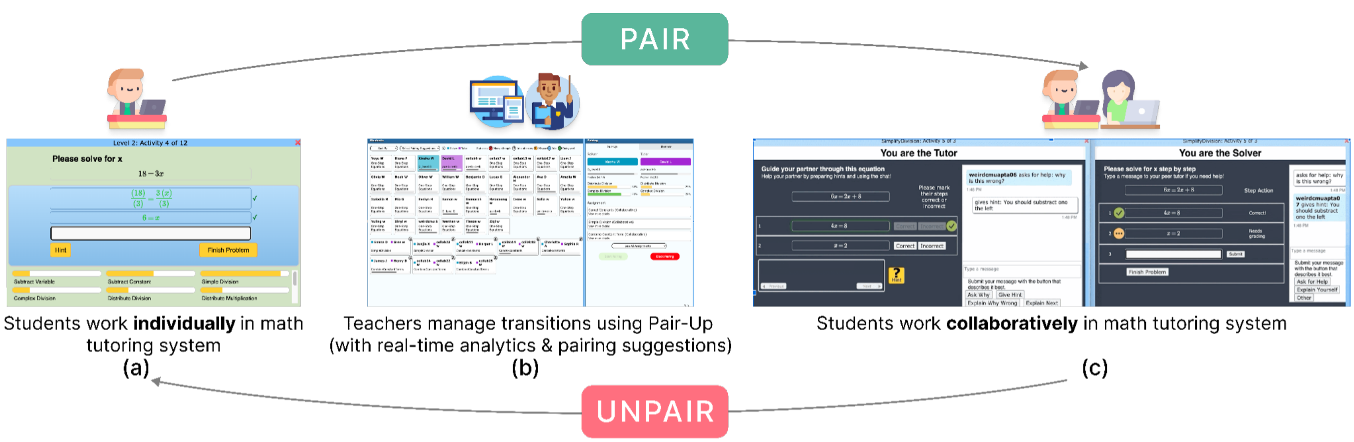

Figure 1: Co-orchestration technology ecosystem for dynamic transition.

ABSTRACT

Enabling students to dynamically transition between individual and collaborative learning activities has great potential to support better learning. We explore how technology can support teachers in orchestrating dynamic transitions during class. Working with five teachers and 199 students over 22 class sessions, we conducted classroom-based prototyping of a co-orchestration technology ecosystem that supports the dynamic pairing of students working with intelligent tutoring systems. Using mixed-methods data analysis, we study the resulting observed classroom dynamics, and how teachers and students perceived and experienced dynamic transitions as supported by our technology. We discover a potential tension between teachers' and students' preferred level of control: students prefer a degree of control over the dynamic transitions that teachers are hesitant to grant. Our study reveals design implications

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04.
https://doi.org/10.1145/3544548.3581398

and challenges for future human-AI co-orchestration in classroom use, bringing us closer to realizing the vision of highly-personalized smart classrooms that address the unique needs of each student.

CCS CONCEPTS

• Human-centered computing → Empirical studies in HCI; **• Applied computing → Interactive learning environments**.

KEYWORDS

Classroom Orchestration, Teacher-supported Tools, Educational Technology, Human-AI Collaboration, Collaborative Learning

ACM Reference Format:

Kexin Bella Yang, Vanessa Echeverria, Zijing Lu, Hongyu Mao, Kenneth Holstein, Nikol Rummel, and Vincent Aleven. 2023. Pair-Up: Prototyping Human-AI Co-orchestration of Dynamic Transitions between Individual and Collaborative Learning in the Classroom. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3544548.3581398

1 INTRODUCTION

Classroom teaching is a dynamic and complex job: it can require teachers to improvise to deliver the most effective instructional support for students at any given moment. Teachers need to continuously observe each learner's current state to flexibly decide how, when and who to support [5, 26, 48]. When multiple students need help simultaneously, a teacher may be unable to attend to their requests and learning needs. Prior classroom observations found that in such moments, teachers may team students up opportunistically, so that students can get help from their peers [21, 22]. We refer to such events as *dynamic transitions* between individual and collaborative learning. Dynamic transitions go beyond the pre-planned team formation strategies that are common in classrooms, where most or all students start or stop collaboration at the same time. In dynamic transitions, by contrast, teachers have students transition between learning modes (individual or collaborative) *when the need arises*, to be maximally responsive to the fact that students learn at their own pace. A teacher thus needs to gauge which students need help, identify suitable partners for these students, and monitor their collaboration to decide when it should end (i.e., when the students should revert back to individual learning). It is not an easy task to manage such dynamic transitions, especially since teachers may need to simultaneously teach, support other students, and attend to unexpected classroom behaviors [21, 22].

The current study prototypes a technology ecosystem that leverages both human and AI system strengths in classroom management. Various educational tools have been designed and developed to support teaching and learning processes mediated by technology [5, 27]. Some of these tools support *classroom orchestration*, which broadly refers to planning, monitoring and real-time management of complex classroom activities [17, 22]. Dillenbourg pointed out [16] that orchestration considers interactions at the level of the broader learning ecosystem, such as a classroom with one or more instructors and groups of learners [25]. At the intersection of education and HCI, there is an emerging emphasis on designing tools from a *human-AI co-orchestration* standpoint, which emphasizes the collaboration and shared control of orchestration responsibility between classroom stakeholders (usually instructor and students), and the AI system [19, 25, 43]. The lens of co-orchestration is important: although assistance from an AI system with managing classroom activities is useful, instructors and students generally like to preserve some level of control over what happens [18, 19, 55]. Adopting the human-AI co-orchestration lens, we created a technology ecosystem with the aim of allowing students to transition dynamically between individual and collaborative learning in a self-paced way. Our technology ecosystem shares orchestration control with human instructors, in that an AI system recommends students' pairings to teachers, and teachers have control over the final orchestration decisions.

Dynamically combining individual and collaborative learning has pedagogical potential, but has yet to be investigated in authentic classrooms. In educational practice, combining individual and collaborative learning activities is very common (e.g., Think-Pair-Share [30] and Jigsaw [6]). From a learning science perspective, individual and collaborative learning each have their own benefits. From the point of view of the Knowledge Learning Instruction (KLI) framework [29], collaborative learning offers opportunities for mutual elaboration and co-construction of knowledge, whereas individual learning may promote induction and refinement as learning mechanisms (cf. [29, 41]). Previous work by Olsen et al. has shown that combining individual and collaborative learning may yield more effective learning for students than either mode alone [42]. Despite some emerging research on transitioning *dynamically* between individual and collaborative learning modes [18, 43], there has not yet been a study that empirically prototypes such processes in the classroom, supported by technology that teachers can use independently without much human assistance.

In this work, we conducted a classroom-based prototyping study of *dynamic transitions*. We have implemented a co-orchestration technology ecosystem consisting of intelligent tutoring systems that support individual and collaborative learning, and a novel teacher-facing co-orchestration tool named Pair-Up. This tool supports human-AI co-orchestration of students' dynamic transitions between individual and collaborative learning activities. It supports teacher awareness of students' learning status by displaying real-time learning analytics. It also provides pairing suggestions to assist teachers with decision-making, but preserves space for teacher agency (i.e., teachers can flexibly override system-suggested pairings). We conducted a classroom prototyping study with five teachers and 199 middle-school students. The overall goal of the prototyping study was to understand the feasibility and desirability of dynamically combining individual and collaborative learning in K-12 classes and to get an initial sense of whether the human-AI co-orchestration technology ecosystem is up to the task. We performed mixed-methods analyses of user transaction log data collected from the technology ecosystem, as well as teacher interview and student survey data.

As discussed next, the current study contributes to HCI research by investigating 1) teachers' orchestration of dynamic transitions and students' learning processes with dynamic transitions, 2) classroom stakeholders' perceived value of dynamic transitions supported by a co-orchestration technology ecosystem, and 3) design opportunities and challenges for future tools that support dynamic transitions. Through this prototyping study, we combine and extend two strands of prior work, respectively, in *human-AI co-orchestration* and *dynamically combining individual and collaborative learning*.

2 RELATED WORK

2.1 Human-AI Co-Orchestration in Classroom

Classroom orchestration broadly refers to planning, monitoring and real-time management of complex classroom activities [17, 22]. Instead of instructors taking the sole responsibility to leverage the complementary strength of humans and AI, it may be beneficial if classroom activities could be *co-orchestrated* by human and AI systems [46]. *Human-AI co-orchestration* is a specific form of human-AI interaction where the humans involved are classroom stakeholders. It refers to educational technology where the orchestration responsibilities might be distributed between (1) instructors and learners, (2) instructors and AI-based instructional agents, or (3) instructors, learners, and AI agents [25]. This distribution of responsibilities may occur via *role splitting*, in which humans and AI agents each take on different roles and tasks within an instructional scenario, or via *role sharing*, in which humans and AI agents play analogous roles and contribute to the same orchestration tasks [25].

Some existing human-AI co-orchestration technologies guide instructors' attention to learners most in need of human attention, while delegating support for other students to AI agents (e.g., [20, 35]). For example, the FACT orchestration system alerts teachers to students who need help, and recommends whom to help and with what. It also suggests instructional responses that teachers can broadcast to a particular group or to the entire class [49]. FACT, while exemplifying co-orchestration, focuses on teacher monitoring and guiding collaboration, instead of transitioning between activities dynamically as in our work.

An emerging vision of human-AI co-orchestration, as outlined by Holstein and Olsen [25], is to combine *instructor-AI* and *instructor-learner* co-orchestration, and share orchestration responsibilities across *instructors*, *learners*, and *AI*. While some initial work has been done in this direction (e.g., [18, 39]), more work is needed to better understand how such shared control can best be designed and function. Without careful design, sharing responsibility between instructors and learners may risk creating a greater load for the instructor, given that learners are often still learning how to take on these responsibilities [25].

Closest to our work, Echeverria et al. conducted a technology probe study in classrooms to gain insight into how teachers, students, and an AI system might co-orchestrate the transitioning between individual and collaborative learning. They explored three ways of distributing control between humans on the AI system: teaming up students based on choices made by students, teachers and the AI system, respectively [18]. They identified a need for hybrid control between students, teachers and AI systems, and adaptivity and/or adaptability for different classroom contexts (e.g., teachers' preference and students' prior knowledge). One difference is that this prior study explored student control, which we do not do in the current study. Another key difference is that in this prior technology-probe study, the pairing actions were performed in a Wizard-of-Oz approach: an additional person played the role of system behind-the-scenes and performed the pairing and, in the system-controlled condition, decided whom to pair up. While the Wizard-of-OZ study gave valuable insights into the needs and desires of shared control in an authentic classroom setting, the current study goes beyond the Wizard-of-Oz study in that the technology ecosystem was fully implemented (i.e., no Wizard was needed). In the current work, the co-orchestration technology shares control with human instructors, in that the AI system recommends students' pairings to teachers, and teachers have control over the final orchestration decisions.

2.2 Dynamic Social Transitions for Personalized Learning

Dynamic transitions between individual and collaborative learning have several potential benefits. First, they may achieve more personalized and differentiated learning. Given each learner is unique in their needs, abilities, and learning pace [54], switching learners to learning activities most suitable for them in the given moment may yield better learning. A second potential benefit of dynamic transitions is that they can leverage the complementary benefits of two learning modes. Collaborative learning supports mutual elaboration and co-construction of knowledge, whereas individual learning may promote induction and refinement as learning mechanisms (cf. [17, 20]). Additionally, research showed that combining individual and collaborative learning may be more effective than either mode alone. Students who engaged in a combination of individual and collaborative learning had higher learning gains, made fewer errors and asked for fewer hints, compared to students who exclusively worked individually or exclusively worked collaboratively [42].

Prior research has also studied social transitions between classroom activities [18, 43, 47]. Multiple *social levels* often occur in a classroom, in that teachers may assign students to work alone, in pairs, in large groups, or in the whole class [43]. Fluid social transitions, as defined by Olsen et al. [43], are those that occur asynchronously between students - not all at the same time for everyone in the class. Dynamic transitions in our context are one form of fluid social transitions, specifically, between individual and collaborative learning activities. Without technological support, managing such transitions is difficult for teachers, as they need to monitor the real-time progress of students, decide whom to team up, determine the optimal time point for transitioning, decide what the teamed-up students should work on, and actualize the transitions (e.g., let the students know what they should do). In the field of HCI and education, many of today's tools are specifically designed for supporting a single social level [43], focusing exclusively on individual learning [20], collaborative learning [1, 15, 38], or whole-class learning [45]. Tools that do support multiple social levels generally have limited support for smooth transitions between these levels [43]. Systems that do facilitate transitions, however, often provide support that is related to the timing of activities (e.g., a timer in the system) [24, 34, 36], leaving the majority of the orchestration load for the teachers [43]. Thus, educational technology tools today, to the best of our knowledge, provide limited support for fluid, dynamic transitions between individual and collaborative learning, a gap that our work aims to bridge.

3 RESEARCH QUESTIONS

Building on the idea of fluid social transitions [43], we define dynamic transitions as having three components 1) monitoring students' learning skills, status, and progress, 2) having students transition when the need arises (e.g., when a student is no longer progressing productively in one mode of learning); and 3) pairing up students in ways that are not fully pre-planned. We developed a co-orchestration technology ecosystem that supports dynamic transitions and tested it in a classroom-based prototyping study. Our study aims to answer the following research questions:

(1) Observed classroom dynamics in classes with dynamic transitions between individual and collaborative learning
  - **RQ1** What are students' learning processes?
  - **RQ2** What are teachers' orchestration processes and pairing strategies?

(2) Students' and teachers' perceptions of dynamic transitions between individual and collaborative learning
  - **RQ3** Do teachers and students see value in dynamic transitions?
  - **RQ4** What design opportunities do they see for future tools that might better support dynamic transitions?

Given the exploratory nature of the study, we do not formulate specific hypotheses related to these research questions.

4 METHODS

The technology ecosystem used in the study comprises three components: an individual tutoring system, a collaborative learning system that supports peer tutoring, and Pair-Up, a teacher-facing orchestration tool.

The first component is an AI-based tutoring system Lynnette (Fig. 1, (a) and Fig. 2) that supports *individual* learning. This system provides step-by-step guidance as a student practices equation-solving individually, in the form of adaptive hints and feedback [33]. The second component is an AI-based tutoring system that supports peer tutoring, which is a re-implementation of Erin Walker's adaptive peer tutoring system (APTA) [51]. It supports two students as they practice solving equations *collaboratively* (Fig. 1 , (c) and Fig. 3), with one student in the role of ***Solver*** and one in the role of ***Tutor***. Whereas the tutoring system for individual learning provides guidance directly to the student, in the collaborative tutoring system, the Solver relies on their partner, the Tutor, for tutoring support. The Tutor is supported by the tutoring system in coaching the Solver - the tutoring system does not support the Solver directly. It provides guidance both with the mathematics and with how to be an effective peer tutor. The Tutor can provide feedback through the system interface (marking each step as correct or not) and can provide hints or explanatory messages via chat. The student taking the Tutor role can request hints from the system for any given step, and need to evaluate the correctness of each step the Solver is taking in order for them to move forward. The Solver and the Tutor do not necessarily sit close to each other and can communicate fully online. In the current design, students can only see their partners' alias usernames. Each collaborative assignment contains three math equation-solving problems. Both systems support personalized mastery learning, estimating students' knowledge based on their interaction with the system predicted by Bayesian Knowledge Tracing (BKT) algorithm [14]. Both tutoring systems have a track record of effectively supporting student learning [20, 33, 52].

The third component is a novel teacher-facing orchestration tool named Pair-Up (Fig. 1 (b) and Fig. 4). Given that individual and collaborative tutoring systems have been extensively used and evaluated in studies [20, 33, 52], we focus description on Pair-Up, a new co-orchestration tool. The core function of Pair-Up is to support teachers in monitoring students' progress and state (in individual and collaborative learning mode), pair students to work collaboratively, and unpair them to stop the collaboration.

The design of Pair-Up was undertaken iteratively in phases with increasing fidelity, building on and informed directly by four prior studies: 1) a Wizard of Oz technology probe in the classroom on transitioning between individual and collaboration [18] 2) surveying teachers' preferences and boundaries in the human-AI co-orchestration process [55], 3) data simulation of possible pairing algorithms [56], 4) interview and co-design workshops with teachers [**?**]. From these studies, we have found that teachers prefer system suggestions and have final control over pairing decisions and that it is feasible to pair students based on their knowledge level, (specifically, in-the-moment wheel-spinning status [9]) based on historical log data from intelligent tutoring systems. We also gained insight into the specific design features and layout choices that teachers desire in a real-time orchestration tool.

Pair-Up, a web-based tool, displays students in card format (Fig. 4 , left panel). Pair-up displays students' skills and progress, allows teachers to choose whom to pair up and what collaborative assignment they should work on (Fig. 4 , right panel). To support teachers in monitoring students' learning state, Pair-Up *displays students' learning analytics in real-time*, by means of indicators of their recent learning status (Fig. 4, (a)), which have been co-designed iteratively with teachers [**?**]. As shown in Fig. 4, displaying the learning status on a student card could alert teachers to students' behaviors such as idling, misusing the software, making lots of errors, making many attempts, and doing well. Pair-Up also displays students' progress in problems at hand, as well as their estimated mastery of the knowledge components [2] that students are practicing. Teachers can sort the student cards in various ways, such as alphabetically, based on the number of math problems solved, or based on the status indicators (i.e., idle, misuse, lots of errors, many attempts, doing well). The teacher is in charge of deciding whom to team up with and when. The system design is commensurate with design guidelines for human-AI systems [3], and is grounded in prior findings from user-centered research with teachers [55]. The tool makes suggestions to teachers as to which student might take on the role of Solver or Tutor, and highlights suggested Solvers in teal and suggested Tutors in purple (Fig. 4). The teachers *may choose to follow the system's pairing suggestions or override them*, and pair students based on their own judgment [18]. Informed by prior results on data simulation of pairing algorithms and surveying of teachers' preferred pairing strategies [55, 56], the tool currently has two built-in pairing suggestion algorithms: *random pairing and pairing students by different knowledge levels*. The latter identifies (as candidates for the Solver role) students who are making slow progress on some of the knowledge components and suggests a

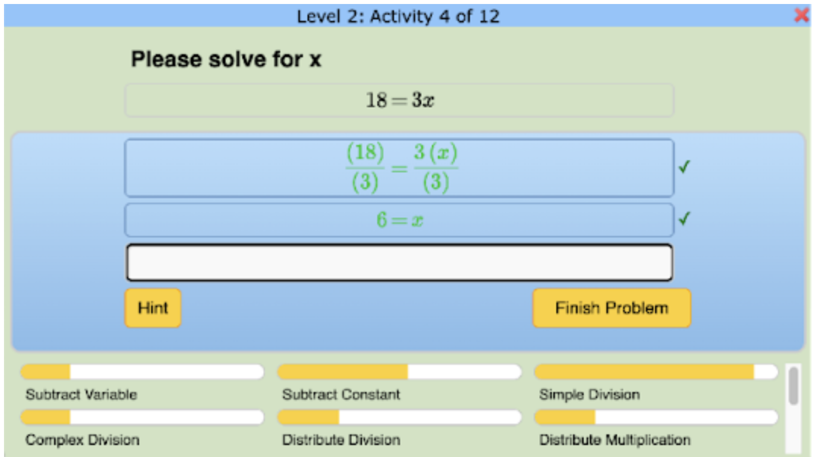

Figure 2: Tutoring software that supports individual learning in algebra.

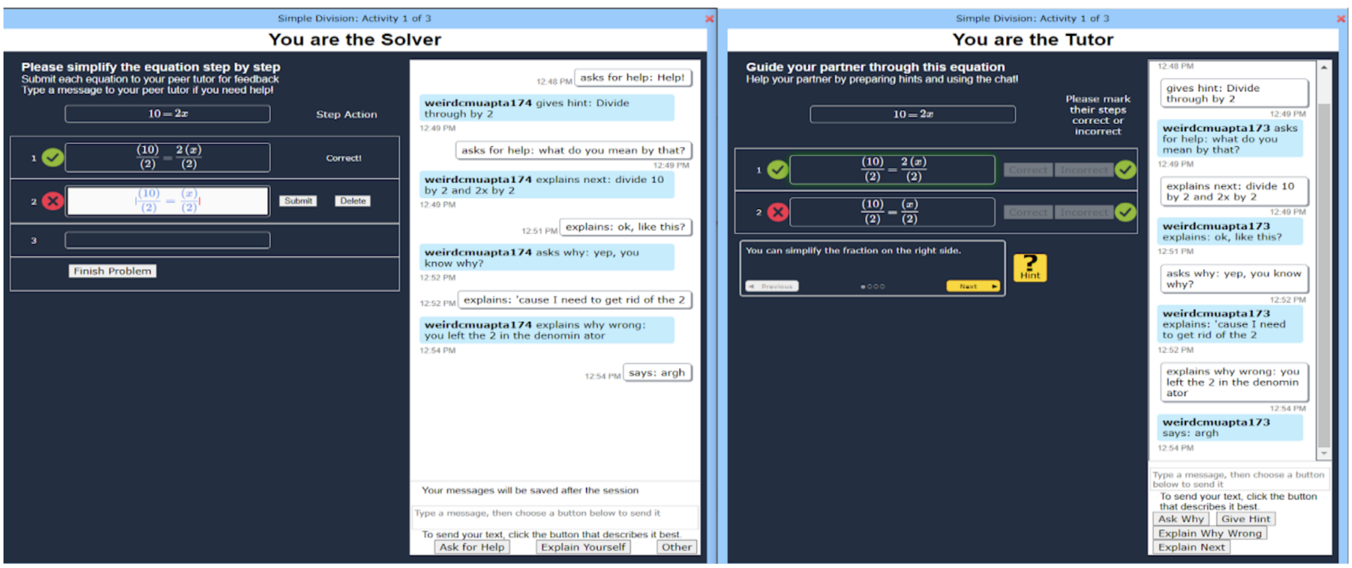

Figure 3: Tutoring software that supports peer tutoring in algebra.

partner (in the Tutor role) who can likely help them with the knowledge components that they are struggling with, based on students' real-time interaction learning data with systems.

4.1 Participants and Procedure of Classroom Prototyping

We conducted a prototyping study in in-person classrooms in a suburban public school in the U.S. Five middle school math teachers (three males (T1, T4, T5) and two females (T2, T3)), and a total of 199 students from 11 classes voluntarily participated, including five 8th-grade classes, five 7th-grade classes, and one 6th-grade advanced class. The average class size is 19.5 (Min = 7, Max = 25). T3 teaches 7th-grade special education with a smaller class size of 7 students, all of whom have an Individual Education Program (IEP). Most have a specific learning disability and may be 1-2 grade levels behind most 7th graders.

Each class participated for 2 class periods, resulting in a total of 22 study sessions. Each study session had a duration of 33-37 minutes, so that each student participated for a total of 70-75 minutes. The first session of each class started with a 7-min tutorial video that briefed the students about how to interact with the AI-based tutoring systems. All students started out on their Chromebook working with the individual tutoring system. A subset of the students (assigned dynamically by their teacher) collaborated using the collaborative tutoring system.

The teachers monitored and paired up students at any time during the class session, using the orchestration tool Pair-Up (Fig. 4) either through a desktop computer (T1, T2, T5) or a tablet (T1, T4). When pairing students, teachers were able to see students'

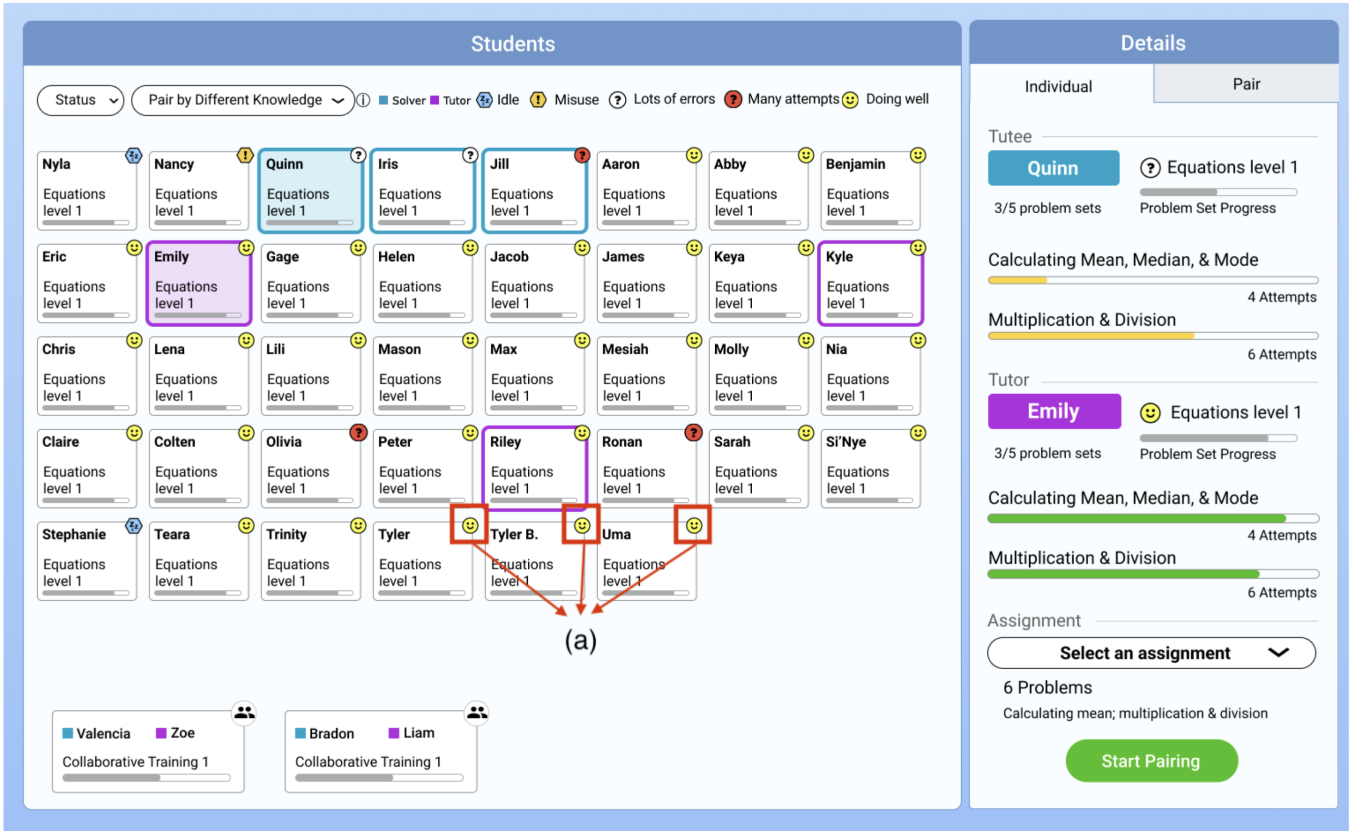

Figure 4: Pair-Up, a teacher-facing co-orchestration tool for managing dynamic transitions. The panel on the left displays cards that represent students working individually (upper part) and collaboratively (lower part). The top-right corner of each card (marked "(a)") shows an icon that represents a student's real-time learning status, indicating whether the student is idling, misusing the tutoring software, making many errors, making many attempts, or doing well. The panel on the right supports the teaming up of students. It displays the skills and progress of a candidate for the Solver role and one for the Tutor role, both selected by the teacher, to help the teacher assess whether these two students might make a good collaborative team. As well, in this panel, the teacher can choose a collaborative assignment for paired-up students to work on.

data, including the skill bars and the number of problems finished. In addition to the data provided via Pair-Up, teachers were also walking around the room to keep an eye on students' pace and monitor their learning situation. The students who were paired up often did not sit next to their partner and communicated with their partner via chat, which was built into the web-based collaborative tutoring system. When students who were paired up were done with the collaborative assignment, or when they were unpaired by the teacher, they switched back to individual work. After the classroom sessions, all teachers participated in a one-on-one semi-structured interview.

4.2 Data Collection and Analysis

To study the students' learning processes (***RQ1***) and teachers' classroom orchestration processes (***RQ2***) during the dynamic transitions, we analyzed the log data collected from the two tutoring systems and Pair-Up, which consisted of a total of 19633 transactions. Metrics related to students' learning process 1 include students' error rate, correct and incorrect attempts on each step, step duration, hint request frequency, and students' chat (in the collaborative tutoring system only). Metrics related to the teachers' orchestration process, include the pairing algorithms that teachers selected, the students that the system suggested to pair up, and teachers' pairing actions. To study how teachers and students perceive the value of dynamic transitions, as supported by the co-orchestration technology ecosystem (***RQ3***), and what design opportunities they see to better support such dynamic transitions (***RQ4***), we conducted and analyzed teacher interviews and student surveys. We now describe our data analysis methods.

4.2.1 Teacher Interview and Student Survey. The aim of the teacher interviews was to understand whether teachers see value in dynamic transitions (***RQ3***) and what design opportunities teachers see for future tools to better support such dynamic transitions (***RQ4***). The interviews lasted on average 45 minutes. The researcher leading the interview asked teachers about 1) the pairing strategy they used, 2) their experience with dynamic transitions and 3) how supported they felt while using the tool. The teacher interviews were

Table 1: Relevant metrics on students learning from transaction data logged by intelligent tutoring systems

Measure	Description
Error Rate	The percentage of students that asked for a hint or were incorrect on their first attempt.
Assistance Score	For a given opportunity, the number of incorrect attempts plus hint requests equals the assistance score.
Number of Incorrects	Total number of incorrect attempts by the student on the step
Number of hints	Total number of hints requested by the student for the step.
Step Duration	Step Duration is the total length of time spent on a step, calculated by adding all of the durations for transactions that were attributed to a given step.
Correct Step Duration	The step duration if the first attempt for the step was correct. Correct Step Duration might also be described as "reaction time" since it's the duration of time from the previous transaction or problem start event to the first correct attempt.
Error Step Duration	The step duration if the first attempt for the step was an error (an incorrect attempt or hint request).

video-recorded and transcribed verbatim. Based on the interview transcripts of five teachers, two coders collaboratively conducted iterative affinity diagramming [12] to surface common themes in the interviews. To this end, they clustered the teachers' statements were iteratively into themes.

Additionally, the goal of our post-study student survey was to understand whether students see value in dynamic transitions (***RQ3***) and study what design opportunities students see for future tools to better support such dynamic transitions (***RQ4***). A total of 171 participating students filled out the survey. The survey questions were adapted from a study by Lin et al. that evaluated students' perceived enjoyment and effectiveness in learning [32]. The survey contained seven five-point Likert scale questions about students' *enjoyment and perceived effectiveness* in learning individually and collaboratively (where 1 means "strongly disagree" and 5 "strongly agree") 5, two multiple choice questions on *students' preferred level of agency* in the pairing process, and three open-ended questions about 1) how students liked the experience of switching between individual and collaborative learning, and what they 2) liked or 3) did not like about the experience. We analyzed students' responses to the multiple-choice questions by computing the percentage of students who chose each option. We analyzed their responses to the Likert scale questions by computing means and standard deviations. To test if students' perceived the effectiveness and enjoyment of individual and collaborative learning modes differently, we performed paired t-tests on students' responses to Likert scale questions. Two coders coded students' responses to the three open-ended questions through thematic analysis [11] to identify common themes in students' opinions.

4.2.2 Collaborative Learning Conversation Content. To understand students' learning processes resulting from the dynamic transitions (***RQ1***) and to probe deeper into students' collaborative learning processes, we performed a manual content analysis of students' conversations during the collaborative learning activities. In this study, students communicated with each other via the chat window. Similar to Wang et al., who analyzed students' chat data from peer learning (cf. [37, 53]), we analyzed students' conversations in chat, logged by the collaborative tutoring software. We adopted the coding scheme from Mawasi et al. [37] where they coded peer help-giving behaviors in middle school math classrooms, with three categories for student conversation: minimal contribution, facilitative contribution and constructive contribution, under which we developed 13 subcategories (2 and supplementary materials).

4.2.3 Analyzing Teachers' Pairing Strategies. To understand teachers' orchestration processes and strategies for pairing up students (***RQ2***), we extracted from log data the number of times teachers paired and unpaired students, how frequently they used each pairing algorithms, and whether they followed the system's suggestions for whom to team up. Additionally, we investigated whether students who were assigned to different collaborative roles by their teacher exhibited different patterns in their learning. Specifically, we compared the learning curves (c.f. [13]) of students who were assigned to take on different roles (Solver, Tutor, both, or neither). One motivation behind this analysis is to confirm whether teachers' actual pairing selection aligned with their intended pairing strategy. Learning curves are auto-generated line graphs based on the transaction data from students' interaction with tutoring systems in Datashop[1]. A learning curve depicts students' performance over successive opportunities to practice a specific learning objective. Our analysis procedure contains the following steps:

- *Step 1.* Depending on the role teachers assigned to students when they paired them up, we grouped students into four categories: students who only served in the role of Tutor (Tutor Only, N = 53), students who only had the role of Solver (Solver Only, N = 61), students who had both roles, though at different times (Both Solver and Tutor, N = 64), and students who did not do any collaboration (No Collaboration, N = 20).
- *Step 2.* We created four data subsets, one for each of the four categories defined in Step 1. We emphasize that these groups were formed after the classroom sessions in the analysis process based on teachers' decisions. Thus, the groups are not experimental conditions with randomly assigned students.

We inspected the learning curve for each group of students (Solver, Tutor, both, and neither, defined above), and compared the

[1]A data analysis repository and analytics infrastructure for the learning sciences community (https://pslcdatashop.web.cmu.edu/).

Table 2: Coding scheme: for collaborative learning conversation content analysis

High level categories	Definition	Sub-categories
Minimal Contribution	Behavior that involves little to no domain content knowledge, e.g., greeting, confirming partner's identity, chatting or conversation related to usability and features of the tutoring systems	- Confirm partner identity - Minimal social behaviors - System usability related conversation
Facilitative Contribution	Behaviors that involve domain content knowledge, and facilitate the collaboration by moving the conversation forward, but are of limited help on building transferable skills for the domain knowledge	- Conversation that is generated by the tutoring software - Solver request for answer - Tutor give answers - Solver ask for correctness feedback - Tutor give correctness feedback - Other transactive social behavior
Constructive Contribution	A statement involving reasoning and explanation of content knowledge. For example, answering a question with an explanation, correcting others with explanation, or asking a specific clarification question to help partner build transferable skills	- Solvers ask Tutors for explanations / clarification - Tutors ask the solver for explanation - Tutors give explanations / clarification

following metrics for each group: Error Rate, Number of Incorrect (Steps), Number of Hints, Step Duration, Correct Step Duration, and Error Step Duration 1. Below we report where we find salient differences in the learning curves of these groups (i.e., indicated by one line of a group is consistently higher or lower than another comparison group(s)).

5 RESULTS: STUDENTS' LEARNING AND TEACHERS' ORCHESTRATION PROCESSES DURING DYNAMIC TRANSITIONS (RQ1 AND RQ2)

In this section, we report students' learning processes that occurred as they switched back-and-forth between individual and collaborative learning (RQ1), based on log data collected by the tutoring systems and Pair-Up. Students solved on average 23.8 math problems while working individually and 1.7 problems while working collaboratively. Students spent on average 39.8 minutes (76% of total time) on individual problems and 12.0 minutes (24% of the time) on collaborative problems 3. As shown in Table 3, during the 22 class sessions, a total of 210 collaboration activities (defined as two students teamed up to work collaboratively on one assignment) happened, with an average of 18 activities in each class. This considerable number of collaboration activities is an initial piece of evidence that teachers were well-supported in orchestrating dynamic transitions by the co-orchestration technology ecosystem.

5.1 Students' Collaborative Learning Processes (RQ1)

In this section, we report results on students' observed learning processes as they dynamically combined individual and collaborative learning (RQ1), based on collaborative learning conversation content analysis 4.2.2.

As mentioned, we analyzed the content of students' chat dialogues using the coding scheme shown in 2 to see what kind of conversational behavior students engaged in during their collaboration (4.2.2). Of all coded behaviors in collaborative learning conversations, 43.1% were minimal contributions, 43.7% were facilitative contributions and 13.1% were constructive contributions. Typical conversations are shown in 4. The standard deviation of these code frequencies tended to be large, suggesting that different student pairs exhibit different learning behaviors. The top subcategories within each major category are: Confirming partner identity (Minimal), transactive social behaviors (Facilitative) and Tutor gives explanation/ clarification (Constructive). The results from this analysis (details in Fig. 5) show that students engaged mostly in minimal conversational behaviors, such as building rapport or confirming their partners' identity, as well as in facilitative behavior, such as giving answers directly to their partners without scaffolding. They engaged less frequently in constructive behaviors that may deal with conceptual knowledge and hold the potential to build transferable mathematics skills.

5.2 Teachers' Orchestration Processes and Pairing Strategies (RQ2)

In this section, we report results on how teachers orchestrate the dynamic transitions and on their strategies for pairing up students (RQ2), based on results from the teacher interviews log data collected from Pair-Up, and informal classroom observations (4.2.3).

From classroom observations, we found that two teachers (T2, T3) teamed up all students at the same time, and three (T1, T4, T5) paired students up at different times. From the interview data, we ascertained that T3 (who taught special education students, who usually have some kind of learning disability)thinks pairing and unpairing everyone at the same time would make it easier for her students to manage the transitions, and easier for herself to manage the classroom since that way she does not need to give instructions multiple times. Although we did not have a chance to ask T2 why she paired up everyone at the same time, we speculate that she prefers more synchronized classroom management because

Table 3: Results from the analysis of log data collected by tutoring software and the orchestration tool (n =19633 rows)

Analysis	Mean	SD	Min	Max
Number of collaboration activities per class (over 2 class sessions)	18.0	8.5	6	37
Math problems solved individually per student	23.8 (93%)	17.2	0	70
Math problems solved collaboratively per student	1.7 (7%)	2.0	0	7
Time spent on individual problems per student (in min)	39.8 (76%)	14.5	0	80
Time spent on collaborative problem per student (in min)	12 (24%)	7.9	0	32.8
Number of unpairing actions by teacher per session	2.8 (22%)	2.5	0	10
Number of times teachers followed the pairing suggestions	6.4 (48%)	4.9	1	20

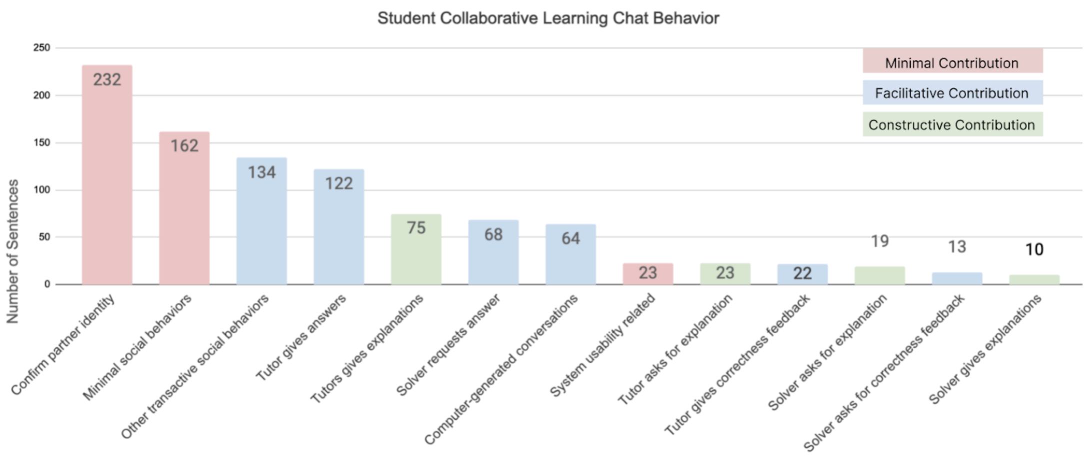

Figure 5: Coding results of students' chat behaviors during collaborative learning.

Table 4: Representative examples of student conversation for minimal, facilitative, or constructive contribution (Students' names are alias)

Minimal	Facilitative	Constructive
T: who are you S: Ash who are you T: Angela S: OKay T: soooo am i supposed to like tell u what to do S: yeah T: Ok T: oh wait no ur supposed to tell me your steps and i tell u if its right	S: hello blozo T: Who are you S: enzo T: Wrong S: why tho T: You need a = S: bruh how it wrong T: Idk T: type 2*x+2*-3+6=14 [Tutor gives answers] T: put that as your answer.. thats what i was told to tell you	S: it probably wrong T: it is a negative number T: would you like me to explain the steps of the problem? S: yeah plz im confused T: alright! just gimme a second T: so first, you'd subtract the 5 from the 1 so everything is on either side [Tutor gives explanations] T: let me know what you get when you subract 5 from 1 S: -4 T: okay, so you take that -4 and divide it by 4 T: what do you get when you divided -4 by 4? [Tutor asks for explanation] S: so it would be -1 T: yes :) T: x=-1

she taught the largest number of classes (five sessions a day) and students (104 in total).

It is interesting to ask what information teachers paid attention to as they made decisions about pairing up students. One teacher mentioned that they would read the room: *"I mean, uh, some classes,* *like some of my periods, they don't require a lot from me. Mm-hmm they don't need as much guidance or assistance or help. And so then I would just kind of monitor and walk around and see, or watch on the screen to see when I should pair or when it looks like they're ready to be paired or ready to switch roles"* (T2). Another teacher mentioned that they would look at students' skill bar and progress on Pair-Up, before checking the system's pairing suggestions: *"I was looking for kids that were not doing so well. And just natural instinct right away, I wanted to pair them with my dashboard somebody had their bars up high. I would feel safer pairing when a student has someone as a tutor that has more knowledge, at least on the bars there. Then I started looking at the suggestions from the tutor itself"* (T1).

We also looked at the degree to which teachers followed the suggestions given by Pair-Up regarding which students to team up. As mentioned, the Pair-Up tool gave pairing suggestions under two pairing algorithms. Teachers could select the pairing algorithm to use and had the final say over whether to follow, ignore, or override these suggestions. Analysis of the log data from the Pair-Up tool (Table 3) indicates that teachers followed the tool's pairing suggestions 6.4 times per class session, comprising 48% of their total pairings. Teachers followed the suggestions of the Different Knowledge Level algorithm (82%) more frequently than those of the Random algorithm (18%). There was substantial variability among teachers in the degree to which they followed the tool's pairing suggestions, ranging from 23% of all pairing decisions to 86%. Three teachers (T1, T2, T3) found the system's pairing suggestions to be helpful *"some of the suggestions were good suggestions, I think."* (T2), and *"I also liked the one that, um, [the system] gave the options for who to pair up, like the suggestions, I like that."* (T3); while one teacher (T4) preferred to manually pair up students, as he thought he was more familiar with students' situations and can pair students in a more flexible and personalized way: *"I didn't really use any of the pairing suggestions for that class. Just because I've been working with them for quite a while so like I kind of already know the students. I go more with just the manual because I take all that other stuff into consideration too when I'm pairing them up. I like to kind of leave it open and just kind of look and see who needs to be paired up and who needs help with what, rather than worry about suggestions"* (T4). Teachers may reject the system's suggestions, for example when the suggested students do not get along (*"The only one I was considering changing was Tom and Roy (alias), just because of, again, social interactions that they've had. They don't get along"* (T3)), or if teachers want to see how different students interact: *"Sometimes I override it just to see how certain students would work together, to see how these students would interact with each other"* (T1).

Based on the teacher interviews, we found that teachers generally wished to pair students who are doing well with those who are not doing well (i.e., struggling). For example, teachers mentioned that *"I try to pair the students who were really good at a topic with someone who maybe made some mistakes on that topic before, I use the students who know farther ahead and obviously had already mastered a lot of those topics and skills as the tutor"* (T4) and *"... if you see somebody who, you know, has really got it and then somebody who's maybe struggling on the problems a little bit, then I pair them up"* (T3). Teachers also mentioned taking into account students' interpersonal relationships: *"A lot of them who are paired are friends. They are more familiar with each other. They can be more direct in their evaluation of what the kid is doing"* (T5). Pairing up friends may ensure they are comfortable working together and increase interaction between collaboration partners. *"Especially in a short period of time. I know I just had two periods. If this is over the course of a school year, I may mix them all up. I just knew I had limited time, and I wanted to get as much interaction as I could. I've seen this where there's not much communication going on. It's just two individuals sitting beside each other sort of doing their own thing. You know, the collaboration is good, but the kids have to buy into it."* Teachers also said they take into account personalities and physical proximity in class. These intended pairing criteria (as stated in interviews) generally aligned with prior research [56].

As for the timing of when teachers paired up students and the conditions under which they did so, some teachers (T3, T4) paired up students based on their observation and monitoring of student's progress in the individual activities, and paired up students who are continually struggling (*"As students've progressed a little bit, I basically look at how they're progressing and if I see students continually struggling, then that's where I would be, pairing them up"* (T4)). For the teachers who paired up students dynamically at different times, they either saw the timing of pairing as something that can be flexibly adapted to students' unique situation (*"But I pair them up when I feel they need it. There is not a set time that I do it"* (T5)), or they viewed the timing as not critical (*"I don't see a bad time"* (T1)).

Based on the log data, we find that on average, teachers unpaired students 2.8 times per session, representing 22% of all the pairs they formed. In all other cases, the unpairing happened automatically when the students finished their collaborative assignment. (These assignments were relatively short, compared to the individual work.) As to the timing or conditions under which teachers unpair students, the five teachers in the study had different preferences. Two teachers (T1, T4) liked to unpair students when their knowledge level increases (*"[I unpair students] whenever they have started to start to show mastery of a new concept if they were paired up for that reason"* (T1)), when their *"misconceptions are gone"*, or if they needed no additional help (*"If I saw the student was able to get through the first two questions with the help and they didn't need any additional help, then I wouldn't pair them and let them go back to individual"* (T4)). One teacher (T5) said they would unpair students based on how the collaborative activity is progressing, for example when it becomes social time ("You hear the conversation, and then the conversation is not focused on what they're doing anymore. And it just becomes, you know, then just chit chatting then that's when I know we gotta get away from this" (T5)). This teacher also stated *"... you can't do it too much. because I think too much collaboration, ... it turns into social time. It's not productive, there's a fine line"* (T5). Some teachers mentioned they did not want to limit the collaboration to a fixed number of problems, or *"haven't really thought honestly about what would be the appropriate time to unpair them if we were doing something (learning activities) after"* (T4).

Regarding the question of whether teachers' actual pairing decisions aligned with their intended pairing strategy, the quantitative analysis of the teachers' pairing strategies (4.2.3) produces three relevant findings:

1) The *No Collaboration* group performed worse than the other groups in their individual work. Students who were not selected by their teacher for collaborative assignments (i.e., the No Collaboration group, who only worked individually) were worse than the other groups on all metrics (see Table 2): They made more mistakes (in their work on the individual tutoring system) than the students in the other groups, all of whom were at one point or another involved in collaboration (Fig. 6, left; the *No Collaboration* group has the highest error rate across learning opportunities). We also found from the learning curve that the No Collaboration students spent more time per step than the other three groups on both correct and incorrect steps.

We conducted statistical tests comparing the error rate of the No Collaboration group with the other groups. The error rate of the No Collaboration group was significantly different from that of the three groups combined (M = 0.23 vs. 0.15, SD = 0.18 vs. 0.16), t(194) = 2.06, p = 0.02. Similarly, we conducted a t-test which shows that the No Collaboration group spent significantly more time in steps in individual work (in seconds) than the other three groups (M = 34.26 vs. 22.78, SD = 20.06 vs. 28.57), t(194)=1.69, p = 0.04).

This evidence suggests that the students who were left to work individually by their teachers could use help from an instructor or from peers. It may indicate that there is room for improvement in the co-orchestration process in that it could better help teachers to identify students who might benefit from being paired up. While we cannot know the specific rationale teachers have for leaving the No Collaboration student without collaboration, we can infer, from the interviews, three factors that may have played a role: 1) ***student behavioral***: the teacher expects that certain students might not work well with others, 2) ***student preference***: the teacher thinks the certain students might prefer learning by themselves, (e.g., *"Um, I have a few, I have some students that do not work well with other students and want work on their own and [I like] just having that opportunity for them to work on the solo assignment."* (T2)) or 3) ***the teacher has helped them individually*** (e.g., *"[I would]First provide individual help, then the pairing serves as the additional help"* (T4)).

2) *Tutors Only* performed somewhat better than *Solvers Only.* Students who on one or more occasions were selected to take on the role of Tutor but were never selected for the role of Solver (i.e., the Tutor Only group) had better performance on the individual learning activities than those who, in their collabotative activites, only had the role of Solver. The learning curves derived from the log data from the tutoring system suggest that the *Tutor Only* group had a lower error rate than the *Solver Only* group (Fig. 6, right). There were no obvious visual differences in the learning curve on step duration for the Tutor Only and Solver Only groups. Neither the difference between the Tutor Only group and the Solver Only group in error rate (M = 35.61 vs. 22.65, SD = 20.06 vs. 28.57), t(110)=0.23, p = 0.41), nor that in step duration (M = 27.73 vs. 22.5, SD = 28.80 vs. 40.45), t(110) =0.80, p = 0.21) was statistically significant, however. Thus, although there was some evidence that teachers preferred (and were able to select) higher-performing students in the role of Tutors, though the performance difference between students selected to only be Tutors vs. those selected to only be Solvers was not statistically significant. They may view these students as more capable of helping students with weaker learning or slower progress. This finding aligns with most teachers' intended pairing strategy as stated during the interviews (i.e., to pair those who learn well with those who are learning less effectively), as well with the most common strategy teachers generally adopted as found in prior research [55, 56].

6 RESULTS: TEACHERS' AND STUDENTS' PERCEIVED VALUE AND DESIGN OPPORTUNITIES (RQ3 AND RQ4)

This section reports results regarding teachers' and students' perceived value of dynamic transitions (RQ3), as well as the opportunities for future tool design expressed by them (RQ4), based on data from the teacher interviews and student surveys.

6.1 Teachers' Perceived Value of Dynamic Transitions (RQ3)

In this section, we report interview results on teachers' experience with and the value they perceived in dynamic transitions (RQ3). All five teachers (including the special education teacher) reported being likely to use technology such as the Pair-Up ecosystem in their regular classrooms.

When asked during the interviews, the teachers generally saw pedagogical value in dynamically combining individual and collaborative learning activities. As one teacher expressed: *"I definitely like it (the dynamic transitions) because it lets you differentiate the learning a little bit better because some kids don't need to be paired up"* (T4), and *"I do like that aspect of having it where you can pair up some or have some work individually, and you can mix it up as needed"* (T4). One teacher liked that this self-paced learning can help distribute resources, support peer learning, and allow students in a class to catch up: *"You got to separate the kids, kids who all understand everything, some students in the middle. Got to use those who get it very quick to try to help the others"* (T5). Another teacher (T2) thought it was useful *"to let students who do not work well with other students and want to work on their own just having the opportunity to work on solo assignments"*. She also stated that dynamic transitions are useful for students who are *"at home due to COVID for an extended period of time, and might not be logging on at the same time"* (T2). However, the special education teacher (T3) had some reservations about dynamic transitions for her class. Although she thought her students liked a mix of different activities and *"enjoyed (switching) back and forth instead of just doing the same thing the same way"*, she viewed dynamic transitions as useful only *"for students who can handle the transitions"*. She was concerned that some of her students with learning disabilities may find the transitions back and forth between different types of activities to be difficult. She still found the dynamic transitions worthwhile and mentioned that *"students learn better when they switch (between learning modes), so I think the pros outweigh the cons"* (T3). One teacher thought that switching back and forth too frequently may hinder students' learning as they may get distracted by who their partners are (T5). Teachers also think that more advanced classes may manage dynamic transitions better. For example, one teacher (T4) selected, for participation in the study, a class he deemed to be more advanced than the other classes, as he thought the students with less experience with equation solving would struggle with

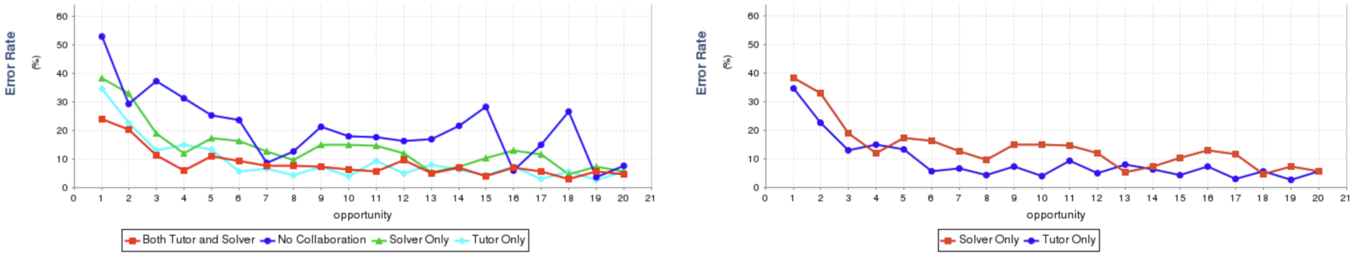

Figure 6: Comparing learning curves that show error rates between: The *No Collaboration* group and the other three groups (left); and the Solver Only and Tutor Only groups (right); The X-axis shows the opportunity count, which captures the number of practice opportunities students had with a given knowledge component in the domain content.

the math content knowledge and might not have the headroom to simultaneously take on the challenge of navigating smoothly between the two tutoring systems.

As for teachers' perceived classroom management load compared to the regular load, the two teachers who used the orchestration tool on *tablet computers* mentioned it was easier or similar to their regular classes. One teacher (T4) expressed that it was easier because he can multitask (walking around the classroom while pairing students), which saves them time so they have more time to help individual students. *"Because it was easier to work with more students and see what they were doing and I can do a quick pair up instead of maybe taking five minutes like I would have to normal class to help certain students"*. This teacher thought it would be harder if he had to use a static device such as a desktop computer *"I think it enhances that if it's mobile. I think it would totally take (me) away from the classroom management if I have to sit at the desk"* (T4). The special education teacher (who used a desktop computer) described it was harder to manage the classroom.

Overall, teachers mentioned they liked the experience, for four main reasons: 1) it allows for self-paced learning that is personalized to students' learning progress and preferences, 2) it allows teachers who use the orchestration tool on the mobile device to multitask (e.g., pairing, monitoring and helping students), 3) it provides a different way of teaching compared with their regular lecturing, 4) the software provides support for peer tutoring, in addition to support for math learning.

6.2 Students' Perceived Value of Dynamic Transitions (RQ3)

In this section, we report quantitative and qualitative survey results on students' experience with and the value they perceived in dynamic transitions (RQ3). In their responses to the *five-point Likert Scale questions* in the students' survey (Table 5), over half of the 171 students who responded rated a mix of individual and collaborative learning as helpful for their math learning (54%). 30% were neutral and 16% of students disagreed to a greater or lesser degree (M=3.3 out of 5, SD=1.05). Students reported higher enjoyment in individual learning than in collaborative learning and the difference was statistically significant (M=3.5 v.s. 3.3, SD = 1.16 vs. 1.23, t(169) = 2.18, p=0.03). 60% of students liked working individually using the online math tutor, and 49% of students liked working collaboratively (chose "like somewhat" or "like a great deal"). Students also perceived individual learning to be more effective, although the difference was not statistically significant (M=2.9 v.s. 2.7, SD=0.99 v.s. 1.06), t(169) =1.36, p=0.08. More students preferred working individually all the time than working collaboratively all the time (M=2.8 v.s. 2.6, SD=1.27 v.s. 1.23). This difference was statistically significant t(169) = 2.06, p = 0.04.

The three optional *open-ended* questions, which asked (1) how students liked the experience of switching between individual and collaborative learning, and what they 2) liked or 3) did not like about the experience), respectively yielded 48, 50, and 47 students' responses. Based on these responses, we found that 39% of students liked switching between individual and collaborative learning (N=19). Reasons that students liked dynamically combining individual and collaborative learning include *"It was helpful so you could do things on your own but also be able to have someone to be partnered with,"* or *"It was very cool and different from other sites like this"*. 31% of students rated the experience as neutral (N=15), and 25% reported not liking it (N=12). Analyzing why these 12 students did not like the switching between activities, it turned out to be mainly because they preferred a single learning mode (N=7), e.g., *"I don't like the switch because I was doing good on my own"*. While most students are fine with switching to a collaborative activity while solving a problem individually, a minority of students reported not liking the unanticipated switching (N=2) e.g., *"It was fine, a little annoying at times because I would be in the middle of a question (and I had to leave it)"*, indicating design opportunities for the transitions to happen in a smoother and less intrusive way.

Students enjoyed collaborative learning with dynamic transitions to different degrees. Asked what they enjoyed about the study experience, among the 50 students' responses, 40% of students (N=20) reported that they enjoyed collaborative learning and having a partner to work with, and that they liked *"working with random people from across the room"* or *"working with a friend"*. Only 6% expressed that the thing they enjoy is working alone (N=3). When asked to state one thing they did not like about the experience, among the 47 responses, 30% of students mentioned some aspects of the collaborative learning experience they did not enjoy (N=14). The main complaints students have about the collaborative learning experience include their partner taking too long to answer or not doing the work (N=5), students not liking their assigned Tutor (N=4), or not being able to pick their partners (N=4).

Table 5: Survey responses from 171 students on experience of dynamic transitions

Survey Items (Self-report by students in five-point Likert Scale)	M	SD
How much did you like working individually using the online math tutor?	3.5	1.16
How effective did you learn math knowledge through individually using the online math tutor?	2.9	0.99
Would you prefer to work individually all the time?	2.8	1.27
How much did you like working collaboratively using the online math tutor?	3.3	1.24
How effective did you learn math knowledge through collaboratively using the online math tutor?	2.7	1.06
Would you prefer to work collaboratively all the time?	2.6	1.23
Do you think a mix of individual and collaborative learning is helpful for your math learning?	3.5	1.04

6.3 Design Opportunities for Future Tools Grounded in Data (RQ4)

In this section, we report design opportunities for future orchestration tool design (RQ4), based on results from both the student surveys and the teacher interviews.

6.3.1 Students' Opinions on Student Control and Agency in the Dynamic Transitions. The survey results indicate that many students prefer to know in advance when and with whom they will get paired. Only 14% of students reported they are okay with not knowing this information. 37% of the students said they would like to know both when and with whom they will be paired. Students cared slightly more about with whom they are paired (27%) than exactly when they will be paired (22%). In addition, the majority of the students (59%) *wished they could have chosen their partner*, with 33% reporting "Maybe", and 8% reporting "No". These findings suggest that co-orchestration tools for transitioning between learning activities may need to support greater awareness on the part of students. They may also need to grant learners some degree of agency regarding the choice of their partner and the timing of transitions.

6.3.2 Teachers' Opinions on Student Control and Agency in the Dynamic Transitions. One teacher mentioned he would like to give students some control over when to collaborate, but not full control as some students *"will just mess around"* (T1). Most teachers preferred that students not know who their partner is (N=3). Results from affinity diagramming reveal that teachers thought anonymous pairing may 1) avoid some social issues (e.g., students do not get along with their partner), 2) avoid some hard-to-control student behaviors that make classroom management harder (e.g., shouting across the room), 3) prevent students from off-topic chatting, 4) protect students' privacy, and 5) in the special education class, prevent students from making fun of smart students and teasing them about being the "Tutor". One teacher, however, acknowledged that it is hard to prevent students from knowing their partners' identity. (Students could reveal their identify in the chat.)

6.3.3 Teachers' General Design Feedback. Teachers made several suggestions for improving the tool. 1) Currently, the tool suggests several possible partners for any given Solver. Two teachers expressed a preference for even more efficient pairing, with the tool automatically pairing up students and the teacher only having to review and confirm these pairings. 2) Currently, the roles of Solver and Tutor are fixed for any collaborative activity. One teacher instead desired a simple way to have students do reciprocal peer tutoring, to experience being both a Solver and a Tutor. (Indeed, the first version of the collaborative tutoring system was meant to support such reciprocal peer tutoring [52].) 3) Currently, the system assigned students to a new collaborative problem when they are paired. One teacher strongly preferred to just pair students to work collaboratively on the problem they were initially working on (struggling) individually because it would be more effective to help struggling students who may just need "a little additional nudge or help" to get unstuck.

Two teachers (T1, T4) wanted to be able to see detailed deep-dive information about individual students (e.g., [20]) so that they could provide personalized help and pair students in a more customized way. One of these teachers suggested that the system would suggest collaborative activities for given students based on their misconceptions. The special education teacher made slightly different suggestions than the other teachers. In her eyes, her students may easily lose focus and are very "reward-motivated." Thus, she suggested adding visual or audio rewards (e.g., a "Ding" sound) when they did something right to keep them engaged, and make the transitioning between activities as smooth as possible. She would also appreciate more guidance on when to pair or unpair students.

7 DISCUSSION

In this work, we developed and prototyped a novel co-orchestration technology ecosystem that focuses on supporting *dynamic transitions* between individual and collaborative learning. *Dynamic transitions* go beyond the pre-planned, whole-class strategies that are common in classrooms; the term refers to students' transitioning between individual and collaborative learning, in a flexible, personalized, self-paced way. It is not an easy task for teachers to manage such dynamic transitions, especially when they need to attend simultaneously to other management tasks. This prototyping study was designed to test the feasibility of dynamically combining individual and collaborative learning, as well as the desirability, in the eyes of teachers and students. We also wanted to get an impression of the resulting classroom dynamics. In this section, we discuss the main answers to the four research questions. We also consider design implications, ethical implications, limitations, and future work.

7.1 Learning and Orchestration Process of Dynamic Transitions, and New Classroom Dynamics

We found students in collaborative activities engaged mostly in minimal conversational behaviors, such as building rapport or confirming their partners' identity, as well as in facilitative behaviors, such as giving answers directly to their partners without scaffolding. They engaged less frequently in constructive behaviors that may deal with conceptual knowledge and hold the potential to build transferable mathematics skills (RQ1). We found some evidence that teachers preferred and were able to identify higher-performing students in the role of Tutors, so they could, in the teachers' eyes, facilitate their partner's learning. However, we also found that students who only worked individually (i.e., who were not assigned to collaborative learning by their teacher) performed worse than those who on one or more occasions were assigned to collaborate (RQ2). Potential reasons that teachers did not select students for collaborative learning activities include behavioral reasons (certain students might not work well with others), student preference (the teacher thinks the students might prefer learning by themselves) or that teacher would like to help them individually. Still, it is possible that teachers overlooked some students who might have benefited from peer tutoring. Assuming peer tutoring can be an effective way to help get students unstuck, leaving students in individual mode may risk prolonging or exacerbating their struggle, especially if teacher help is not available or adequate. Thus it may be fruitful for future orchestration tools to support teachers even more effectively in identifying students who need to be paired up.

With this initial exploration of dynamic transitions in the classroom, *new classroom dynamics arose*. Firstly, the students switched dynamically between (AI-supported) individual learning and (AI-supported) collaborative learning. Secondly, teachers were able to successfully orchestrate this complex process while multitasking. They walked around the classroom to keep an eye on individual students while monitoring the progress of all students. They quickly paired students up using the orchestration tool, sometimes accepting the tool's suggestions for whom to team up, at other times substituting other choices. Designing orchestration tools that allow teachers to multitask has been explored in past research (c.f. [27]). It is an exciting venue for design exploration for future tools, as it fits the practical needs in teachers' classroom practice [5, 22]. Thirdly, teachers expressed that dynamic transitions could be helpful given COVID-19 constraints. As students may join remotely or log in at different times, it allows them to be paired up and collaborate more flexibly without prolonging the waiting time for partners.

Overall, the prototyping study provides evidence that, with appropriate technical support, dynamic combinations of individual and collaborative learning are feasible – the orchestration is manageable, and teachers appear to be content with the teams they form, which tended to align with their stated preference to assign higher-performing to the role of Tutor. Students were able to switch between the two tutoring systems without great problems – even if there were some relatively minor complaints regarding the suddenness of transitions and less agency over the choice of partners.

7.2 Experiences with and Perceived Value of Dynamic Transitions

For RQ3, we found students' perception of dynamic transitions can be affected by their collaborative learning experience. A considerable number of students report certain aspects of collaborative learning (usually concerning the partner they are paired with), to be the one thing they particularly like (40%) or dislike (30%) during the entire study experience. This finding indicates that collaboration can be a polarizing experience. Students perceptions may be inseparably tied to, and greatly reflect their perceived enjoyment and effectiveness of *working with a particular partner*, instead of collaborative learning activities itself. This finding may also explain why students desire more agency in the pairing process (i.e., knowing when and with whom they are paired, and choosing their own partner).

From teachers' perspective, we found initial evidence that teachers felt well supported by the Pair-Up tool in orchestrating dynamic transitions. As well, they liked being able to monitor students and pair them up based on their progress in individual activities. Teachers reported a positive view towards dynamic transitions and expressed they were likely to use the technology in their regular teaching in the future, as it allows for more differentiated learning and it helps ensure that students receive help in moments of struggle. Nevertheless, our findings also turned up some concerns from special education teachers or teachers of weaker classes, namely, that students with weaker domain knowledge or learning disabilities may not handle switching between activities well.

Most teachers (four out of five) suggested that the classroom management load when using the technology ecosystem is similar to or less than that in their regular practice. In general, teachers were receptive to the idea of dynamic transitions and recognized the pedagogical value of dynamically transitioning students from individual to collaborative learning and back. This finding confirms results from a prior co-design study [**?**]. Over half of the students thought a mix of individual and collaborative learning is helpful for their math learning.

While it was not the primary focus of this study to evaluate which type of device affects the orchestration load of dynamic transitions, teachers perceived that a mobile device (e.g., a tablet) yields a lower orchestration load compared to a static device (e.g., desktop) when orchestrating dynamic transitions, as it allows them to multitask in their usual practice. Prior research in the HCI and AIED communities has demonstrated that mobile, portable or wearable devices in classrooms hold great potential to support more efficient teaching practices and classroom orchestration (c.f. Keeping watch [44], ClassBeacon [4], Lumilo [21] and others [5, 10, 40]), which this work provides additional evidence for.

7.3 Design Implications and Challenge

For RQ4, in addition to the specific suggestions pointed out by teachers and students in surveys and interviews, we provide the following *design implications* based on study findings, intended for the HCI and educational research communities, especially researchers building human-in-the-loop technologies to support teaching and learning.

Firstly, future tools for supporting chat-based collaboration activities (whether dynammically combined with other activities or not) should focus on cultivating positive and constructive collaboration behavior (which involves reasoning and explanation of content knowledge and may help build transferable skills). Constructive contributions in collaborative dialogues, such as reasoning or explanation in terms of conceptual knowledge, can often support students in building transferable skills [37**?**]. In the current study, constructive contributions were rare in students' dialogues, however. Scaffolding students in better peer tutoring, so that they engage in more constructive behaviors such as explaining and clarifying conceptual knowledge, may lead to more productive collaborative learning experiences. In this kind of short-term collaboration activity, it can also be important to ensure that the students are responsive to their partners. One effective way of nudging students to engage in timely interactions may be for the system to display a timer countdown notification when one party is idle for a long time.

Secondly, teachers would like to be given analytics about student work that more directly inform them of the optimal timing of pairing and unpairing actions. While students are working *individually*, teachers desire more deep-dive information on students' misconceptions and current problems to help them make more informed pairing decisions. In *collaboration*, given the considerable differences in students' experiences and the large variance in their collaborative learning quality, it is important to help teachers monitor the progress in and quality of students' collaborative activities, so as to optimize the learning experience and outcome for students. The design of collaborative learning analytics can draw from prior work on positive indications of collaboration including sharing resources, joint actions, mutual planning, equal participation, communication and reaching consensus [27]. Additionally, one (system-generated) analytic that might guide teachers decide when to unpair students is whether students have improved substantially on the knowledge skills for which they were paired up initially. Similarly, it would be useful if the system could detect when students' misconceptions had been resolved.

Thirdly, tools for dynamic transitions may need to be *customizable and support teachers' varied preferences depending on their goals for pairing*. Educational technology works best when aligned with teachers' values, instructional goals, and teaching practices [23, 28]. Echoing the technology probe study by Echeverria et al., who found that the design of dynamic transitions may need to be adaptable for different classroom contexts (e.g., teacher's preference and students' prior knowledge) [18], the current study uncovered that the tool design may need to be adaptable to the *goals, purposes, and values* teachers hold for initiating collaboration between their students. For example, teachers whose goal for pairing students is *to have reciprocal peer-tutoring* prefer a role-switching feature, so students can switch between the roles of Tutor and Solver. In contrast, teachers whose goal for pairing is *to help struggling students get unstuck* wished for a feature that allows struggling students to continue working on the problems they were stuck on when they get teamed up with a partner, instead of getting new problems. Some other teachers, who teach large classes may prefer a more efficient and automated way of forming pairs. They may be more willing to give up some degree of agency and control in the human-AI co-orchestration process in exchange for greater efficiency [3].

Our study also reveals a *design challenge*: we observed a tension between teachers' and students' preferred level of control over the transition process. Previous studies in human-AI co-orchestration found that teachers prefer to have the final say in the pairing process [18, 55]. However, in our study, we found that students would also like to have control over when and with whom they will be paired. It is an interesting design challenge to creatively incorporate more student agency without threatening teachers' authority and preferred control of the classroom. As Holstein and Olsen pointed out, sharing the orchestration responsibilities between instructors and learners may be beneficial, but may also risk creating a greater load for the instructor [25]. Still, it is an emerging vision for co-orchestration systems to combine instructor-AI and instructor-learner co-orchestration [25]. Features that may help address this challenge include a student-informed list of preferred and blocked partners, as well as allowing students to request a different partner or to request that the collaboration is halted when it is of poor quality. Additionally, when deciding the proper partner, timing, and content for a collaborating pair, one potential way to incorporate both teachers' and students' agency and control can be to have one party propose ideas and to let the other party makes the final decision (e.g., teachers propose content for the collaborative activity and students choose, or student propose several possible partners and teachers select one). Another general approach is to co-design with teachers and students to generate design ideas that might balance their preferred level of control in the classroom.

7.4 Ethical Implications

With the adoption of tools that support dynamic transitions can come ethical implications, as revealed in the current study. Firstly, dynamic transitions, which inherently require students to switch between different learning activities, may pose additional challenges for students with learning disabilities, as revealed in this study. Since these students may have impaired attention spans [8, 50], and may have difficulty focusing on and staying engaged in one learning activity, switching between different ones may compromise their learning effectiveness. Secondly, as compared to traditional group formation which usually involves the whole class so that every student is part of a group, the nature of dynamic transitions encourages instructors to only pair students for whom they deem collaboration to be beneficial. While teachers express that they oftentimes let students who in their judgment do not need help work by themselves, it is possible that their judgments are not always accurate and/or up-to-date. This is evidenced in our analysis, which found that students left in individual mode have lower performance than other groups. Given dynamic transitions as supported in this current technology ecosystem rely quite heavily on teachers' subjective judgment, dynamic transitions may risk keeping struggling students in individual learning states. Thus, given that peer tutoring holds the potential to scaffold students, keeping them in individual learning activities may unnecessarily prolong their struggle, without careful design such as support to help teachers spot struggling students or allow students more control in the process [18, 25, 46]. Thirdly, given that teachers may be inclined to pair students who they know would work well with their peers, one potential risk of dynamic transitions is that it will not provide a fair and equal opportunity for collaboration, for students whom teachers consider to "have social behaviors issues" or to "not work well with others". Future tools for supporting dynamic transitions should take these ethical implications into consideration.

7.5 Limitation and Future Directions

One limitation of the current study is it only evaluates dynamic transitions in the math algebra context for students using online intelligent tutoring systems, so the new insights it provides may be domain-specific. We do not claim that the tools used in the current study are comprehensive (e.g., they do not provide for (partial) student control), as we are still honing the technology ecosystem. Future work can explore how to balance teachers' and students' preferred levels of control in the co-orchestration process. Additionally, this prototyping study was not designed to test whether dynamic combinations of individual and collaboration lead to better experiences or outcomes for students or teachers, compared to effective alternative forms of instruction. With a classroom technology ecosystem of high complexity, it is appropriate to first do prototyping studies to test feasibility and desirability, as we did in the current study. Evaluating whether dynamic combinations of individual and collaborative learning have better outcomes than other forms of instruction is an interesting and important open question for future research.

8 CONCLUSION

We conducted a classroom prototyping study of a human-AI co-orchestration technology ecosystem, to get a sense for whether dynamically combining individual and collaborative learning can be feasible and helpful. The study shows that dynamic transitions between different activities have pedagogical value in the eyes of teachers and, to a lesser degree, of students. Teachers view dynamic transitions as a way to achieve differentiated learning that can be customized to each student's individual pace and preferred learning mode. Supporting dynamic transitions in actual classrooms is worthwhile from a pedagogical perspective, yet not without challenges such as finding the optimal timing, content, and partner for transitions to collaborative learning. The challenge may lie in 1) supporting teachers in identifying opportune moments and suitable partners, without overloading them in classroom management, 2) ensuring students are responsible, responsive, and contribute in a constructive way during collaboration, 3) allowing smooth, non-disruptive transitions that switch students to different learning activities in moments that they may be most helpful, without collaborative activities turning into unproductive social or idle time. The current prototyping study builds a case that with Pair-Up and the two AI-based tutoring systems that support individual and collaborative learning, dynamic combinations of individual and collaborative learning are feasible in classrooms. The study also highlights a number of ways in which the design of the technology ecosystem could be improved.

The current study is the first prototyping study in authentic classrooms on dynamic transitions as supported by a human-AI co-orchestration technology ecosystem. The findings, insights, and design implications in this study are not limited to the specific design case of Pair-Up or the co-orchestration technology ecosystem. It may generalize to other educational systems that uses complementary strength of human and AI systems, and collaborative learning systems where students communicate via chat. A secondary goal of our study is to spur more work at the intersection of HCI and education that adopts the lens of human-AI co-orchestration. We believe there are many possibilities to embody the insights from the classroom study and the design implications in future tools that support classroom learning and teaching, and leverage complementary strengths of human and AI systems synergistically.

ACKNOWLEDGMENTS

We thank NSF funding agency, all research team members, and participating teachers and students for their contribution to this research.

REFERENCES

[1] Hamed Alavi and Pierre Dillenbourg. 2012. An Ambient Awareness Tool for Supporting Supervised Collaborative Problem Solving. *Learning Technologies, IEEE Transactions on* 5 (07 2012), 264–274. https://doi.org/10.1109/TLT.2012.7

[2] Vincent Aleven and Kenneth R Koedinger. 2013. Knowledge component (KC) approaches to learner modeling. *Design Recommendations for Intelligent Tutoring Systems* 1 (2013), 165–182.

[3] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-AI Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300233

[4] Pengcheng An, Saskia Bakker, Sara Ordanovski, Ruurd Taconis, Chris L.E. Paffen, and Berry Eggen. 2019. Unobtrusively Enhancing Reflection-in-Action of Teachers through Spatially Distributed Ambient Information. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300321

[5] Pengcheng An, Kenneth Holstein, Bernice d'Anjou, Berry Eggen, and Saskia Bakker. 2020. The TA Framework: Designing Real-Time Teaching Augmentation for K-12 Classrooms. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3313831.3376277

[6] Elliot Aronson. 1978. *The jigsaw classroom.* Sage.

[31]]Bartel-undated-mj A. N. Bartel, E. M. Silla, N. A. Vest, T. Nagashima, V. Aleven, and M. W. Alibali. [n. d.]. Reasoning about equations with tape diagrams: Insights from math teachers and college students. *Proceedings of the 15th International Conference of the Learning Sciences - ICLS 2021* ([n. d.]). https://doi.org/10.22318/icls2021.685

[8] B Bateman. 1992. Learning disabilities: the changing landscape. *J. Learn. Disabil.* 25, 1 (Jan. 1992), 29–36. https://doi.org/10.1177/002221949202500105

[9] Joseph Beck and Yue Gong. 2013. Wheel-Spinning: Students Who Fail to Master a Skill, Vol. 7926. 431–440. https://doi.org/10.1007/978-3-642-39112-5_44

[10] Matthew Bernacki, Jeffrey Greene, and Helen Crompton. 2019. Mobile Technology, Learning, and Achievement: Advances in Understanding and Measuring the Role of Mobile Technology in Education. *Contemporary Educational Psychology* 60 (11 2019), 101827. https://doi.org/10.1016/j.cedpsych.2019.101827

[11] Virginia Braun and Victoria Clarke. 2016. (Mis)conceptualising themes, thematic analysis, and other problems with Fugard and Potts' (2015) sample-size tool for thematic analysis. *International Journal of Social Research Methodology* 19 (06 2016), 1–5. https://doi.org/10.1080/13645579.2016.1195588

[12] Reviewed by Elmar Hashimov. 2015. Qualitative Data Analysis: A Methods Sourcebook and The Coding Manual for Qualitative Researchers. *Technical Communication Quarterly* 24, 1 (2015), 109–112. https://doi.org/10.1080/10572252.2015.975966 arXiv:https://doi.org/10.1080/10572252.2015.975966

[13] Hao Cen, Kenneth Koedinger, and Brian Junker. 2006. Learning Factors Analysis – A General Method for Cognitive Model Evaluation and Improvement, Vol. 4053. 164–175. https://doi.org/10.1007/11774303_17

[14] Albert T Corbett and John R Anderson. 1995. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Model. User-adapt Interact.* 4, 4 (1995), 253–278. https://doi.org/10.1007/BF01099821

[15] Sébastien Cuendet, Quentin Bonnard, Frederic Kaplan, and Pierre Dillenbourg. 2011. Paper Interface Design for Classroom Orchestration. (05 2011). https:

//doi.org/10.1145/1979742.1979904
[16] Pierre Dillenbourg. 2013. Design for classroom orchestration. *Computers & Education* 69 (2013), 485–492.
[17] Pierre Dillenbourg and Patrick Jermann. 2010. *Technology for Classroom Orchestration.* Springer New York, New York, NY, 525–552. https://doi.org/10.1007/978-1-4419-5716-0_26
[18] Vanessa Echeverria, Kenneth Holstein, Jennifer Huang, Jonathan Sewall, Nikol Rummel, and Vincent Aleven. 2020. Exploring Human–AI Control Over Dynamic Transitions Between Individual and Collaborative Learning. In *Addressing Global Challenges and Quality Education*, Carlos Alario-Hoyos, María Jesús Rodríguez-Triana, Maren Scheffel, Inmaculada Arnedillo-Sánchez, and Sebastian Maximilian Dennerlein (Eds.). Springer International Publishing, Cham, 230–243.
[19] Kenneth Holstein, Vincent Aleven, and Nikol Rummel. 2020. *A Conceptual Framework for Human–AI Hybrid Adaptivity in Education.* 240–254. https://doi.org/10.1007/978-3-030-52237-7_20
[20] Kenneth Holstein, Bruce Mclaren, and Vincent Aleven. 2018. *Student Learning Benefits of a Mixed-Reality Teacher Awareness Tool in AI-Enhanced Classrooms.* 154–168. https://doi.org/10.1007/978-3-319-93843-1_12
[21] Kenneth Holstein, Bruce Mclaren, and Vincent Aleven. 2019. Co-Designing a Real-Time Classroom Orchestration Tool to Support Teacher–AI Complementarity. *Journal of Learning Analytics* 6 (08 2019), 27–52. https://doi.org/10.18608/jla.2019.62.3
[22] Kenneth Holstein, Bruce Mclaren, and Vincent Aleven. 2019. *Designing for Complementarity: Teacher and Student Needs for Orchestration Support in AI-Enhanced Classrooms.* 157–171. https://doi.org/10.1007/978-3-030-23204-7_14
[23] Kenneth Holstein, Bruce M. McLaren, and Vincent Aleven. 2017. Intelligent Tutors as Teachers' Aides: Exploring Teacher Needs for Real-Time Analytics in Blended Classrooms. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (Vancouver, British Columbia, Canada) *(LAK '17)*. Association for Computing Machinery, New York, NY, USA, 257–266. https://doi.org/10.1145/3027385.3027451
[24] Stian Håklev, Louis Faucon, Thanasis Hadzilacos, and Pierre Dillenbourg. 2017. Orchestration Graphs: Enabling Rich Social Pedagogical Scenarios in MOOCs. https://doi.org/10.1145/3051457.3054000
[25] Kenneth Holstein And Jennifer K Olsen. 2021. Human–AI Co-orchestration: The Role of Artificial Intelligence in Orchestration. In *Handbook of Artificial Intelligence in Education.* Edward Elgar Publishing.
[26] Trevor Kerry and Carolle A. Kerry. 1997. Differentiation: teachers' views of the usefulness of recommended strategies in helping the more able pupils in primary and secondary classrooms. *Educational Studies* 23, 3 (1997), 439–457. https://doi.org/10.1080/0305569970230309 arXiv:https://doi.org/10.1080/0305569970230309
[27] Ahmed Kharrufa, Sally Rix, Timur Osadchiy, Anne Preston, and Patrick Olivier. 2017. Group Spinner: Recognizing and Visualizing Learning in the Classroom for Reflection, Communication, and Planning. 5556–5567. https://doi.org/10.1145/3025453.3025679
[28] Paul Kirschner. 2015. Do we need teachers as designers of technology enhanced learning? *Instructional Science* 43 (03 2015). https://doi.org/10.1007/s11251-015-9346-9
[29] Kenneth Koedinger, Albert Corbett, and Charles Perfetti. 2012. The Knowledge-Learning-Instruction (KLI) Framework: Toward Bridging the Science-Practice Chasm to Enhance Robust Student Learning. *Cognitive science* 36 (04 2012), 757–98. https://doi.org/10.1111/j.1551-6709.2012.01245.x
[30] Aditi Kothiyal, Sahana Murthy, and Sridhar Iyer. 2014. Think-pair-share in a large CS1 class: Does learning really happen? *ITICSE 2014 - Proceedings of the 2014 Innovation and Technology in Computer Science Education Conference* (06 2014). https://doi.org/10.1145/2591708.2591739
[31]]Lawrence2022-je L. Lawrence, B. Guo, K. Yang, V. Echeverria, Z. Kang, V. Bathala, C. Li, W. Huang, V. Aleven, and N. & Rummel. [n. d.]. Co-designing AI-based orchestration tools to support dynamic transitions: Design narratives through conjecture mapping. *Computersupported collaborative learning* 15 ([n. d.]). https://par.nsf.gov/biblio/10366101
[32] Chun-Yu Lin, Chung-Kai Huang, and Chao-Jung Ko. 2020. The impact of perceived enjoyment on team effectiveness and individual learning in a blended learning business course: The mediating effect of knowledge sharing. *Australasian Journal of Educational Technology* 36 (01 2020), 126–141. https://doi.org/10.14742/ajet.4446
[33] Yanjin Long and Vincent Aleven. 2013. Supporting Students' Self-Regulated Learning with an Open Learner Model in a Linear Equation Tutor, Vol. 7926. https://doi.org/10.1007/978-3-642-39112-5_23
[34] Roberto Martinez-Maldonado, Andrew Clayphan, and Judy Kay. 2015. Deploying and Visualising Teacher's Scripts of Small Group Activities in a Multi-surface Classroom Ecology: a Study in-the-wild. *Computer Supported Cooperative Work (CSCW)* 24 (06 2015). https://doi.org/10.1007/s10606-015-9217-6
[35] Roberto Martinez-Maldonado, Andrew Clayphan, Kalina Yacef, and Judy Kay. 2015. MTFeedback: Providing Notifications to Enhance Teacher Awareness of Small Group Work in the Classroom. *Learning Technologies, IEEE Transactions on* 8 (04 2015), 187–200. https://doi.org/10.1109/TLT.2014.2365027
[36] Roberto Martinez-Maldonado, Yannis Dimitriadis, Judy Kay, Kalina Yacef, and Marie-Theresa Edbauer. 2013. MTClassroom and MTDashboard: Supporting analysis of teacher attention in an orchestrated multi-tabletop classroom. *Proc. CSCL2013* 1, 119–128.
[37] Areej Mawasi, Ishrat Ahmed, Erin Walker, Shang Wang, Zeynep Marasli, Amanda Whitehurst, and Ruth Wylie. 2020. Using Design-Based Research to Improve Peer Help-Giving in a Middle School Math Classroom.
[38] Emma Mercier. 2016. Teacher orchestration and student learning during mathematics activities in a smart classroom. *International Journal of Smart Technology and Learning* 1 (01 2016), 33. https://doi.org/10.1504/IJSMARTTL.2016.078160
[39] Inge Molenaar, Anne Horvers, and Ryan Baker. 2019. Towards Hybrid Human-System Regulation: Understanding Children' SRL Support Needs in Blended Classrooms. *LAK19: Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 471–480. https://doi.org/10.1145/3303772.3303780
[40] Hannelore Montrieux, Ruben Vanderlinde, Tammy Schellens, and Lieven Marez. 2015. Teaching and Learning with Mobile Technology: A Qualitative Explorative Study about the Introduction of Tablet Devices in Secondary Education. *PloS one* 10 (12 2015), e0144008. https://doi.org/10.1371/journal.pone.0144008
[41] Dejana Mullins, Nikol Rummel, and Hans Spada. 2011. Are two heads always better than one? Differential effects of collaboration on students' computer-supported learning in mathematics. *I. J. Computer-Supported Collaborative Learning* 6 (09 2011), 421–443. https://doi.org/10.1007/s11412-011-9122-z
[42] Jennifer Olsen, Nikol Rummel, and Vincent Aleven. 2019. It is not either or: An initial investigation into combining collaborative and individual learning using an ITS. *International Journal of Computer-Supported Collaborative Learning* 14 (10 2019), 1–29. https://doi.org/10.1007/s11412-019-09307-0
[43] Jennifer Olsen, Nikol Rummel, and Vincent Aleven. 2020. Designing for the co-Orchestration of Social Transitions between Individual, Small-Group and Whole-Class Learning in the Classroom. *International Journal of Artificial Intelligence in Education* 31 (10 2020). https://doi.org/10.1007/s40593-020-00228-w
[44] Rebecca Quintana, Chris Quintana, Cheryl Madeira, and James D. Slotta. 2016. Keeping Watch: Exploring Wearable Technology Designs for K-12 Teachers *(CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 2272–2278. https://doi.org/10.1145/2851581.2892493
[45] Mirko Raca and Pierre Dillenbourg. 2013. System for Assessing Classroom Attention. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (Leuven, Belgium) *(LAK '13)*. Association for Computing Machinery, New York, NY, USA, 265–269. https://doi.org/10.1145/2460296.2460351
[46] Mike Sharples. 2013. Shared Orchestration Within and Beyond the Classroom. *Computers & Education* 69 (11 2013), 504–506. https://doi.org/10.1016/j.compedu.2013.04.014
[47] Mike Tissenbaum and Jim Slotta. 2019. Supporting classroom orchestration with real-time feedback: A role for teacher dashboards and real-time agents. *Int. J. Comput. Support. Collab. Learn.* 14, 3 (Sept. 2019), 325–351. https://doi.org/10.1007/s11412-019-09306-1
[48] Carol Ann Tomlinson. 2014. *The Differentiated Classroom: Responding to the Needs of All Learners.* ASCD.
[49] Kurt Vanlehn, Hugh Burkhardt, Salman Cheema, Seokmin Kang, Daniel Pead, Alan Schoenfeld, and Jon Wetzel. 2018. *The Effect of Digital Versus Traditional Orchestration on Collaboration in Small Groups.* 369–373. https://doi.org/10.1007/978-3-319-93846-2_69
[50] Sharon Vaughn and Sylvia Linan-Thompson. 2003. What Is Special About Special Education for Students with Learning Disabilities? *Journal of Special Education - J SPEC EDUC* 37 (11 2003), 140–147. https://doi.org/10.1177/00224669030370030301
[51] Erin Walker, Nikol Rummel, and Kenneth Koedinger. 2013. Adaptive Intelligent Support to Improve Peer Tutoring in Algebra. *International Journal of Artificial Intelligence in Education* 24 (10 2013). https://doi.org/10.1007/s40593-013-0001-9
[52] Erin Walker, Nikol Rummel, and Kenneth R Koedinger. 2014. Adaptive Intelligent Support to Improve Peer Tutoring in Algebra. *International Journal of Artificial Intelligence in Education* 24, 1 (Jan. 2014), 33–61. https://doi.org/10.1007/s40593-013-0001-9
[53] Xu Wang, Meredith Thompson, Kexin Yang, Dan Roy, Kenneth Koedinger, Carolyn Rose, and Justin Reich. 2021. Practice-Based Teacher Questioning Strategy Training with ELK: A Role-Playing Simulation for Eliciting Learner Knowledge. *Proceedings of the ACM on Human-Computer Interaction* 5 (04 2021), 1–27. https://doi.org/10.1145/3449125
[54] Haoran Xie, Hui-Chun Chu, Gwo-Jen Hwang, and Chun-Chieh Wang. 2019. Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017. *Computers & Education* 140 (06 2019), 1–16. https://doi.org/10.1016/j.compedu.2019.103599
[55] Kexin Yang, Luettamae Lawrence, Vanessa Echeverria, Boyuan Guo, Nikol Rummel, and Vincent Aleven. 2021. *Surveying Teachers' Preferences and Boundaries Regarding Human-AI Control in Dynamic Pairing of Students for Collaborative Learning.* 260–274. https://doi.org/10.1007/978-3-030-86436-1_20
[56] Kexin Bella Yang, Vanessa Echeverria, Xuejian Wang, LuEttaMae Lawrence, Kenneth Holstein, Nikol Rummel, and Vincent Aleven. 2021. Exploring Policies for Dynamically Teaming up Students through Log Data Simulation. *International Educational Data Mining Society* (2021).