Engagement and Learning Benefits of Goal Setting with Rewards in Human-AI Tutoring

Conrad Borchers[0000−0003−3437−8979], Alex Houk[0009−0001−5933−6970], Vincent Aleven[0000−0002−1581−6657], and Kenneth R. Koedinger[0000−0002−5850−4768]

Carnegie Mellon University
{cborcher,ahouk,aleven,krk}@cs.cmu.edu

Abstract. Active learning promises improved educational outcomes yet depends on students' sustained motivation to engage in practice. Goal setting can enhance learner engagement. However, past evidence of the effectiveness of setting goals tends to be limited to non-digital learning settings and does not scale well as it requires active teacher or parent involvement. We study goal setting in a hybrid human-AI tutoring context, where students engage with personalized learning software with feedback and hints while being supported by a human tutor. Each student set a weekly goal (e.g., how much time or how many skills to master) and was rewarded for goal achievement, while human tutors regularly checked in about goal progress. We investigate whether this intervention improves student engagement and skill mastery. We observed 110 middle school students in a hybrid tutoring program over 12 weeks, with goal-setting support integrated into the program after six weeks. Using a quasi-experimental interrupted time series model, we estimated the intervention's impact on engagement and learning. Weekly practice time increased, on average, by about 25% after introducing goal setting, and skills mastered per week increased by about 40%. This effect remained stable over time. These findings suggest that goal-setting support can improve the quantity and quality of practice in hybrid tutoring contexts with minimal added teacher workload. The present study advances the understanding of strategies that can increase student engagement with personalized learning systems in a low-cost and scalable manner, improving the learning benefits of such systems.

Keywords: goal setting, hybrid tutoring, K-12, STEM, active learning

1 Introduction and Related Work

Active learning, a cornerstone of seminal AIED learning systems such as intelligent tutoring systems, teachable agents, and inquiry-based learning, involves directly engaging students through problem-solving rather than passive instruction [17]. If instruction is adequate, active learning enhances learning, regardless of prior knowledge [16]. However, the effectiveness of active learning hinges on students' willingness and motivation to engage in the effortful practice it requires.

This law is empirically accurate across demographic groups and replicated in authentic classroom learning contexts using cognitive modeling [12, 16, 26].

Addressing this challenge—how to support students in sustaining motivation—remains a critical focus in educational technology research. Here, we study how goal setting, a method that has been extensively linked to enhanced effort, can support students in benefiting more from AIED systems. Our study is situated in hybrid tutoring classrooms, where learners are supported by technology and human tutors through remote video conferencing software [28].

Goals are *"object or aim of an action, for example, to attain a specific standard of proficiency, usually within a specified time limit"* [18]. Setting goals as a motivational and performance-enhancing strategy has been extensively studied. Seminal work by Locke and Latham [19] identified key factors contributing to goal achievement: providing goal achievement feedback, fostering goal commitment through rewards, ensuring requisite knowledge and skills to achieve goals, and accounting for situational support, such as teacher or parent involvement [15]. *Homework behavioral contracts*, or contingency contracts, are one common method for implementing goal setting in classroom and homework settings [23]. Cooper [9] describes contingency contracting as a written agreement between two parties. This contract outlines the tasks assigned to each participant, the rewards designated for the student, and the conditions required to earn rewards. Empirical studies have shown that student practice time increases when parents and students jointly set specific goals [15]. More recent literature reviews demonstrate that contingency contracting is generally an effective strategy to support the acquisition of cognitive and non-cognitive skills in students [2].

Integrating goal setting with personalized learning remains underexplored [10, 25]. Goal-setting contracts are typically performed on physical paper, limiting integration with the feedback and data-driven capabilities of learning systems [24]. The combination of goal setting and personalized learning may be particularly effective for several reasons. First, it may outperform traditional goal-setting methods due to the continuous performance feedback provided by technology, which helps students calibrate their efforts and manage overconfidence or underconfidence in their ability to achieve goals [22, 13]. Second, compared to pen-and-paper homework, students may achieve greater learning gains per unit of effort when goal setting is combined with intelligent tutoring, thereby amplifying the engagement benefits traditionally associated with goal setting [15, 16, 12]. Third, prior research suggests that regular goal setting, feedback, and evaluation cycles can strengthen students' self-regulated learning (SRL) skills, which may transfer to other learning tasks, enhancing overall learning effectiveness [8].

Although past AIED systems have supported learners in setting goals (e.g., [11]), these studies focused on process goals and improving metacognition *during* learning rather than performance goals typical for goal-setting contracts [23]. Additionally, prior research on SRL goal support has usually been limited to short-term instructional interventions (e.g., of a few hours [11]). In contrast, the present study examines interventions over several weeks, allowing for observing engagement changes and persistence, which we model through linear time trends.

The present study adopts a goal-setting approach that minimizes the need for teacher involvement. We worked with teachers integrating goal setting and rewards into their hybrid tutoring classrooms. We observed whether integrating goal-setting classroom practices into hybrid tutoring is feasible and has tangible student learning benefits. Using an interrupted time series design [21], we estimated the influence of our intervention on practice time. Further, to validate whether students learn more, or merely engage more [16, 12, 26], we monitor estimated skill mastery during the goal-setting. Finally, to confirm if intervention benefits are long-lasting rather than short-term, we model linear time-related engagement trends. We investigate the following research questions:

RQ1: Do students engage in more practice during hybrid tutoring sessions after completing goal-setting contracts compared to before?
RQ2: Does student engagement remain stable over time during goal support in hybrid tutoring?
RQ3: Are changes in engagement during hybrid tutoring with goal setting reflected in skill acquisition?

This study contributes to the growing research body on human-AI-supported learning [14]. It examines the immediate impact of goal-setting contracts, support, and rewards on practice time and their effects on skill mastery, providing insights into strategies for enhancing learning in AIED learning settings.

2 Methods

2.1 Sample and Recruitment

We analyzed data from a hybrid tutoring program running for 12 weeks between October and December 2024. Data from Thanksgiving break (week of November 27th) was excluded, as no school activities occurred. Students (N=110) were from a charter school in the Mid-Atlantic United States, serving grades 6-9. All students enrolled at the school were invited to participate in the tutoring program, and those who provided consent were included in the sample. Students were all male, nearly all African American, and all from low-income backgrounds.

As part of the hybrid tutoring program, online human tutoring is available to students during one class session per week, during which they engage in math practice with the IXL software. A researcher and the school's two math teachers facilitated the goal-setting activities in person. Each classroom session was 43 minutes. Eight classes were served, four working with the 6th and 7th-grade teacher and four with the school's 8th and 9th-grade teacher.

The researcher and one tutor supervisor facilitated tutor participation and training. The tutors were fifteen university student workers who had worked with the student participants since the beginning of the 2024-25 school year as part of the hybrid tutoring program. They provided as-needed mathematics support and goal progress check-ins via the Pencil video conferencing software. Human tutor support was initiated by either the student or the tutor based on student learning needs based on standardized test scores. Tutors varied in how frequently they tutored, from one period per week to five periods per week ($M = 2$ weeks).

2.2 Materials

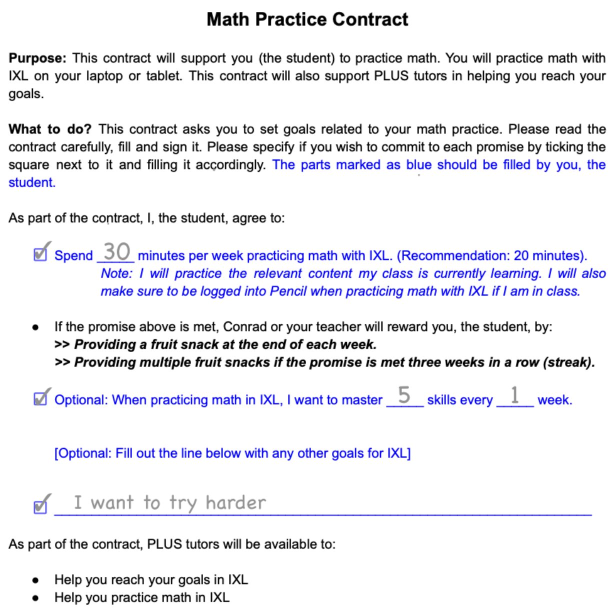
Learners practiced using the IXL Math software, an adaptive online learning platform designed to support personalized math practice. The effectiveness of the IXL math curriculum has been documented in past research, finding significant improvements in learning relative to comparable non-IXL schools throughout a three-year intervention in grades 3-8 [4]. IXL is often used by teachers in the United States [20] and provides a comprehensive curriculum covering a wide range of topics aligned with the Common Core State Standards. The platform uses real-time analytics to adapt problem difficulty based on a student's performance using knowledge tracing and mastery learning using a proprietary algorithm, ensuring a tailored learning experience that matches their skill level.

Students receive immediate feedback on their answers to math problems, typically requiring a single solution step. Brief motivational messages are displayed after correct attempts. The system provides an explanation or a step-by-step problem walkthrough for incorrect responses. IXL may occasionally offer a worked example to support students in mastering concepts. Further, IXL tracks skill progress and generates reports, enabling educators to monitor individual performance and identify areas needing support. Motivational elements, such as achievements and milestones, are integrated to sustain student engagement.

2.3 Procedures

During all hybrid tutoring sessions, before and after the goal-setting phase (six weeks each), students attended their regular 43-minute math class and signed into IXL and Pencil, a video conferencing program, on their assigned Chromebooks. Students were greeted by a tutor and sent to individual breakout rooms, where they worked through assigned problems in IXL Math and could ask for assistance from tutors. On Pencil, students would share their screen with tutors so that tutors would view the students' IXL work and provide content support as needed. Tutors also provided motivational support, such as praising the students' efforts. Each class period was attended by approximately the same set of tutors each week. Students were introduced to all tutors supporting their class session at the beginning of the school year in order to build rapport. On a weekly basis, a student may have interacted with any of the tutors in their class session's set of tutors or another tutor if one of the regular tutors was absent. The classroom teacher and a research assistant helped students resolve any technical issues during the lesson. Additionally, we observed that teachers usually shared an IXL leaderboard with students during practice, which continued during the goal-setting phase, which allows students to see how many minutes they practiced in the current week. We observed that some students would occasionally check the leaderboard to gauge if they had reached their goal during practice.

Students completed a math practice goal contract in the first week of goal setting, handed out by a research assistant at the beginning of their math practice period (Fig. 1). Students completed the contract independently and chose their goal for minutes practiced and, optionally, the number of skills mastered per week and a third, custom goal. The contract was a design adapted from research recommendations described in Peacock et al. [23]. Specifically, it actively engaged the student in setting goals and detailed rewards associated with reaching them every week. Finally, it highlighted the importance and purpose of the goal with detailed instructions at the top and highlighted that remote tutors would help the student achieve their goals and learn math at the bottom.



Math Practice Contract

Purpose: This contract will support you (the student) to practice math. You will practice math with IXL on your laptop or tablet. This contract will also support PLUS tutors in helping you reach your goals.

What to do? This contract asks you to set goals related to your math practice. Please read the contract carefully, fill and sign it. Please specify if you wish to commit to each promise by ticking the square next to it and filling it accordingly. The parts marked as blue should be filled by you, the student.

As part of the contract, I, the student, agree to:

- [x] Spend 30 minutes per week practicing math with IXL. (Recommendation: 20 minutes).
  Note: I will practice the relevant content my class is currently learning. I will also make sure to be logged into Pencil when practicing math with IXL if I am in class.

- If the promise above is met, Conrad or your teacher will reward you, the student, by:
  >> Providing a fruit snack at the end of each week.
  >> Providing multiple fruit snacks if the promise is met three weeks in a row (streak).

- [x] Optional: When practicing math in IXL, I want to master 5 skills every 1 week.

[Optional: Fill out the line below with any other goals for IXL]

- [x] I want to try harder

As part of the contract, PLUS tutors will be available to:

- Help you reach your goals in IXL
- Help you practice math in IXL

Fig. 1: Illustration of the goal-setting process with an example contract.

Students received a biweekly printout goal achievement report, *if* they over- or underachieved by a margin of 33% or more, handed out by a research assistant (e.g., "Average: 30 minutes (150% goal achievement)") and were able to adjust their goal. Further, remote tutors had access to their set goals and occasionally reminded them about their goals. The two goal categories (i.e., achieving a certain number of practice minutes or the number of skills mastered) were determined through discussions with the school's STEAM coordinator and math teachers. If students met one of their goals, they would get a reward the next week, which was distributed by a research assistant at the beginning of the tutoring sessions. Students were free not to choose any goals, and 24 (22%) did choose not to do so or were absent on the first week of goal setting, though the class research assistant offered all students the opportunity to complete a goal contract in the first subsequent week they were in class. All students continued participating in the hybrid tutoring program as usual. To avoid selection bias, students without a goal contract were not excluded from the analysis.

During the goal-setting phase, hybrid tutoring continued in the same format with the addition of students' individual goal contracts. Students who met their goal would receive a reward at the start of the next tutoring session, distributed by the teacher and a research team member. To integrate goal setting into the hybrid tutoring program effectively, we followed recommendations and requests from the school's STEAM coordinator and math teachers. Accordingly, each week of goal achievement was rewarded with a packet of healthy fruit snacks.

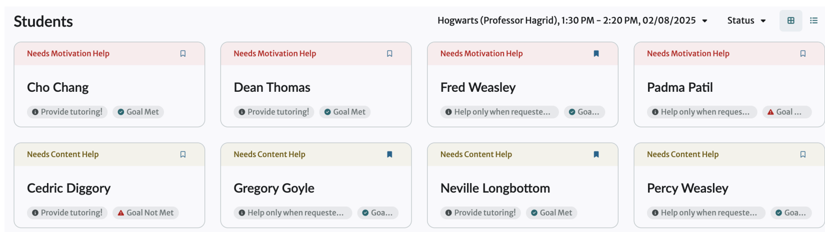


Fig. 2: Example student data dashboard as seen by remote tutors.

To support student goal achievement, tutors could view whether a student had met their goal in the previous week via a dashboard (Fig. 2). The dashboard included recommendations for which students tutors should initiate tutoring interactions based on their inferred support needs (which is not the subject of this study). Tutors were trained to praise students who achieved their goals for their efforts. If a student had not met their goal, tutors were trained to provide advice tailored to why the student felt they did not meet it. For example, if a student expressed inadequacy in math, the tutors were asked to empathize with their struggles and motivate them to persevere. These evidence-based practices are a regular and effective part of hybrid tutoring beyond math support [27].

2.4 Measures and Data Preprocessing

A researcher compiled a weekly report detailing the number of minutes each student practiced and the number of skills each student mastered, practiced, and was proficient in. In IXL, the threshold for estimated knowledge mastery of skills was 80% for proficient and 100% for mastered, referred to as "SmartScore" in the student- and teacher-facing application. This dataset was used to populate the student data dashboard used by tutors and for data analysis (Section 2.5).

We computed the following measures: a **time** variable denoted the number of weeks (to estimate the general trend in the outcome variable over the entire observation period), a **goal setting time** week indicator (a separate week count to model time trends specifically after goal setting had been introduced, capturing changes practice trends compared to before goal setting), and a binary **goal setting indicator** distinguishing pre- and during-goal setting phases).

2.5 Data Analysis

To investigate our three research questions, we employed an interrupted time series design paired with linear mixed-effects modeling to analyze trends in practice time and skill acquisition. The unit of analysis was individual student weeks. Each row presents a student's weekly practice outcome regarding the number of minutes practiced in IXL and the number of practiced and mastered skills.

Given that goal setting was introduced in the middle of the school term, an interrupted time series design naturally emerged in the study data. Interrupted time series analysis is a quasi-experimental approach used to assess the effect of an intervention by analyzing data trends before and after its implementation [21]. This method is particularly suited for evaluating interventions where randomized control trials are not feasible or desirable. In our context, a student-level randomized assignment of the goal-setting intervention was undesirable as it could cause students to get frustrated about not getting the opportunity to earn rewards, potentially weakening the control condition and student morale. By comparing learning before and during the goal-setting phase, the design helps identify immediate and sustained changes attributable to goal-setting while accounting for underlying longitudinal trends in student effort and learning. As shown in Fig. 3, the interrupted time series design of this study segmented the data into two phases: (1) a pre-goal setting phase baseline phase and (2) a during-goal setting phase, each spanning six weeks. By examining practice and learning differences between these phases within students (i.e., we compare each student's engagement and outcomes before and after), we sought to infer the effectiveness of the goal-setting activity regarding student practice behaviors.

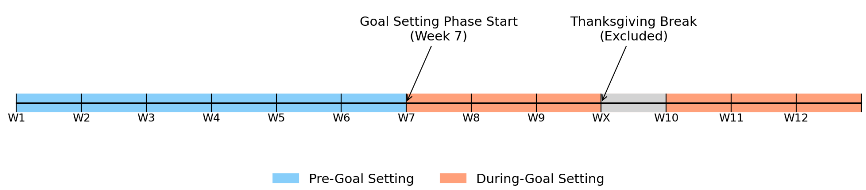


Fig. 3: Schematic Study Timeline

To evaluate the impact of goal setting in hybrid tutoring, we fit a linear mixed-effects model with scaled weekly practice time as the dependent variable. This method is considered quasi-experimental because the introduction of goal setting was not randomized but implemented naturally as part of the program, allowing for comparison of outcomes before and after the intervention without random assignment [21]. Since the practice outcomes (e.g., minutes practiced, skills mastered) are not independent within students, potentially biasing p-values of model coefficients toward significance, we included a random student intercept in the models. This adjustment accounted for each student's baseline practice behavior [5]. The model further included fixed effects for time, time since the introduction of goal setting, and a main effect of the goal setting (see Section 2.4). The model of weekly student engagement (Y_{ij}) was specified as follows:

$$Y_{ij} = \beta_0 + \beta_1(\text{Week}_j) + \beta_2(\text{Week Goal}_j) + \beta_3(\text{Goal (Yes/No)}_j) + u_i + \epsilon_{ij} \quad (1)$$

where u_i represents the random intercept for student i, index j denotes week counts, and ϵ residual error. Linear time trends in engagement are separately modeled via week counts for the overall period (Week variable) and the intervention period (Week goal variable, starting at Week 7) to detect changes in engagement trends related to the intervention. Finally, the binary coefficient of `Goal (Yes/No)`, set to 1 during the intervention, estimates the intervention effect on the outcome Y after adjusting for time trends and individual student differences through u_i. This effect is important because it captures the immediate level shift in student engagement attributable to the onset of goal setting.

To address **RQ1**—whether students engage in more practice during hybrid tutoring sessions with goal-setting contracts than without—we examined the main effect of goal setting (`Goal (Yes/No)`) using the mixed-effects model. This effect, estimated in standard deviation units after standardizing the outcomes, quantifies the average difference in practice time between the pre-and during-goal setting phases. We visualized weekly practice time trends using a centrality plot to complement this analysis. The plot displayed mean practice minutes per week, with reference lines indicating pre- and during-goal setting averages.

For **RQ2**, which examines whether students' practice achievement increases over time following the introduction of goal setting, we interpreted the time trends after goal setting was introduced using the separate week count (`Week_Goal`). This allowed us to evaluate whether goal-setting contracts led to any significant longitudinal increase or decrease in practice time, captured as linear trends.

To study **RQ3**—whether changes in practice time during goal setting are reflected in skill acquisition—we descriptively analyzed aggregated summary statistics of skill mastery and compared them to results for **RQ1**. This analysis leveraged multiple mastery measures (described in the Measures Section 2.4) to examine patterns in skill acquisition before and during goal setting.

3 Results

3.1 Descriptive Differences in Time Spent

Fig. 4 illustrates the average number of minutes practiced per week, with dashed lines representing the average minutes practiced before and during goal setting. During goal setting, students exhibited a higher average practice time ($M = 21.9$, $SD = 20.7$) than before ($M = 17.6$, $SD = 25.1$). This difference reflects a 24.43% increase in mean practice time during goal setting. The blue line represents weekly fluctuations in the mean minutes practiced across all participants.

3.2 Interrupted Time Series Modeling

RQ1 related to whether students engaged in more weekly math practice during hybrid tutoring during the intervention. The fixed effects estimates in Table 1

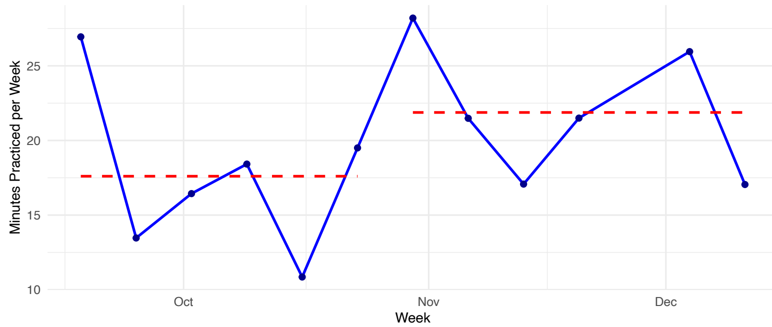


Fig. 4: Average number of practice minutes across weeks with dashed reference lines showing the average outcomes pre-and-during goal setting.

Table 1: Fixed effects estimates from the mixed effects interrupted time series model of the weekly practice time outcome (scaled to standard deviations).

Predictors	**Estimates**	**95% CI**	**p-value**
Intercept (β_0)	0.65	0.30 – 0.99	<.001
Week (β_1)	-0.06	-0.10 – -0.02	.008
Week Goal (β_2)	0.02	-0.03 – 0.08	.404
Goal (Yes/No) (β_3)	0.48	0.28 – 0.69	<.001

summarize the key findings of the interrupted time series analysis. In line with the descriptive data above, goal setting was associated with a significant increase in the number of minutes practiced per week after adjusting for time trends ($\beta_3 = 0.48$, $p < .001$), corresponding to an average effect of about 0.5 SD.

Answering whether student engagement remained stable throughout the goal-setting phase (RQ2), weekly practice time showed a small but significant decline ($(\beta_1) = -0.06$, $p = .008$), suggesting a downward trend prior to goal setting. The post-time effect was not significant ($\beta_2 = 0.02$, $p = .404$), indicating that the rate of change during goal setting did not significantly differ from the previous trend. These results suggest that goal-setting led to a significant improvement, though it could not disrupt the existing downward trend in engagement. It also suggests that the goal-setting effect was robust across time and did not recede.

3.3 Did Students Also Learn More During the Intervention?

One concern is that students will not learn more when setting goals primarily related to effort (e.g., minutes spent), which could be achieved by practicing with less effort [3]. Therefore, we investigated skill mastery estimated from the IXL learning software before and during goal setting (RQ3). Table 2 summarizes the descriptive statistics for skills practiced, proficient, and mastered across two 6-week periods: pre-goal setting and during-goal setting.

Table 2: Descriptive statistics of average skills practiced, proficient, and mastered over two 6-week periods: pre-goal and during-goal settings.

Metric	Pre-Goal Setting	During-Goal Setting
Total Skills Practiced	7.46 (SD: 6.75)	11.49 (SD: 9.01)
Total Skills Proficient	3.42 (SD: 4.08)	4.84 (SD: 5.57)
Total Skills Mastered	2.97 (SD: 3.64)	4.12 (SD: 5.18)

The results show a substantial increase in all skill mastery metrics following the introduction of goal setting into hybrid tutoring. The total skills practiced increased by 53.97%, from an average of 7.46 (SD: 6.75) pre-goal setting to 11.49 (SD: 9.01) during goal setting. Similarly, the total skills proficient increased by 41.52%, from 3.42 (SD: 4.08) to 4.84 (SD: 5.57), while the total skills mastered rose by 38.72%, from 2.97 (SD: 3.64) to 4.12 (SD: 5.18).

4 Discussion

This study investigated the impact of goal-setting contracts, goal support, and rewards on students' practice behaviors and skill acquisition in a hybrid tutoring program. Motivated by the challenge of sustaining student engagement in AIED systems, this study aimed to explore whether structured goal-setting mechanisms could address motivational gaps and enhance learning outcomes.

4.1 Engagement in Practice Improved Through Goal Setting (RQ1)

The analysis revealed a significant increase in average weekly practice time during the goal-setting period. This finding highlights the potential of goal-setting contracts with rewards to enhance student engagement. The observed increase in practice time suggests that students responded positively to the intervention, aligning with prior research on the motivational benefits of goal setting [15, 2]. However, unlike past research, our results are of significance because of the novel context of hybrid tutoring and personalized learning, which enables novel, data-driven forms of goal support that are more scalable [1, 2]. Specifically, we advance the theory of who can function as an accountable partner in goal-setting interventions: hybrid tutors and, to a lesser extent, the personalized learning technology itself (which reminded students of their goal progress through a data-driven leaderboard, for instance, enabling goal monitoring).

From a practical standpoint, this result implies that hybrid tutoring systems can incorporate goal-setting mechanisms to boost engagement without necessitating constant teacher or parental oversight, unlike past research [15]. Schools and educators may consider embedding similar contracts into digital learning platforms to foster sustained student participation to the degree that resources to monitor, support, and reward goal completion are available.

4.2 Changes in Practice Time Remained Stable Over Time (RQ2)

Students maintained higher practice levels during the intervention. This boost in engagement did not significantly change over time. This finding suggests that while contracts may effectively sustain engagement, they do not automatically promote continuous improvement (although one past study found such virtuous cycles in goal achievement [29]). Still, variation in achievement could lead to a separation of performance levels within a class, subject to future research, and may be studied through student-level performance histories. Finally, we observed a significant trend whereby student practice time faded throughout the 12 weeks, possibly due to unobserved factors such as increased testing.

4.3 Skill Acquisition Benefits Exceed Engagement Benefits (RQ3)

The goal intervention yielded substantial increases in skill mastery, with gains in skills practiced (54%), proficient (42%), and mastered (38%) exceeding those in practice time (25%). This disproportionate improvement suggests that the intervention may have enhanced not only the quantity but also the quality of practice. Our finding is contrary to the idea that students would simply maximize unproductive practice time to achieve engagement goals with little effort [3].

Notably, most students chose to set practice time goals, not skill mastery goals, aligning with past research, noting that it is easier for middle school students to express goals in the former, more familiar metric [24]. Hence, initially, expressing goals in terms of the number of minutes worked may be sufficient for students to achieve mastery learning goals in AIED systems, as considerable design research in explainable AI is required to make mastery-based problem selection intuitive for middle school students [6].

4.4 Limitations and Future Work

Several limitations warrant consideration. First, our analysis cannot disentangle the effects of goal-setting mechanisms, reward structures, and hybrid tutor goal support on extrinsic and intrinsic motivation. Meta-analytic evidence demonstrates that intrinsic motivators (e.g., goal-directed achievement) and extrinsic incentives (e.g., performance-contingent rewards) simultaneously influence performance, rather than necessarily undermining each other, including in education [7]. Therefore, future work may focus on *isolating* the distinct contributions of extrinsic and intrinsic motivation, for example, via modeling of local effects of intrinsic goal-setting and adjustment events from extrinsic reward events.

Second, the six-week goal-setting period may have been insufficient to capture long-term effects, including potential impact of receding engagement if goal-setting were to be taken away. Third, given the observational nature of the study design, which closely works around existing teacher and program practices, evidence remains quasi-experimental and the intervention partially confounded with the (personalized) learning content. Fourth, the chosen linear mixed-effects model does not account for potential non-linear time trends in practice behavior.

Future research could study differential goal trajectories in and out of classrooms. Past research predicts that students' self-efficacy benefits of meeting goals may propel them into a virtuous cycle of increased achievement [29]. Lastly, a research assistant facilitated the goal-setting procedure in this study. While we believe that human tutors could independently implement goal setting and monitoring in the future, thus requiring neither extra teacher effort nor classroom support, this remains an untested assumption. The minimized teacher involvement observed in this study was feasible due to the hybrid tutoring context, where remote tutors assumed responsibility for monitoring and discussing goal progress with students. These responsibilities may not translate directly to traditional classroom settings without hybrid support, as the same workload would otherwise fall on classroom teachers. Future research should examine whether human tutors can effectively manage goal-setting responsibilities without additional support and explore the scalability of such interventions across contexts.

5 Conclusion

Although active learning in AIED systems is widely recognized as effective for improving educational outcomes, research on supporting students to initiate, sustain, and achieve practice efforts is limited. Drawing from goal-setting research from non-digital homework research, we studied the integration of goal-setting contracts with contingent rewards into a middle school hybrid tutoring program, where human tutors supported students working with a personalized learning system. The intervention significantly boosted weekly practice time and skill mastery based on interrupted time series modeling over 12 weeks. Students spent about 25% more time on task during the goal-setting phase, offsetting a prior downward trend in engagement. Skill development rose disproportionately: total skills practiced jumped by about 50%, while proficiency and mastery increased by about 40%. Notably, the intervention's immediate impact did not change over time despite an overall semester-wide decline in engagement.

The results of this study have practical implications for designing scalable improvements to learning and engagement in resource-constrained educational environments, such as hybrid tutoring with as-needed human support in AIED learning systems. By leveraging goal-setting contracts and integrating goal support into existing tutoring programs, schools can enhance student engagement and learning outcomes with minimal additional burden on teachers.

Acknowledgments

This work was made possible with the support of the Learning Engineering Virtual Institute. The opinions, findings and conclusions expressed in this material are those of the authors. This research was funded by the Institute of Education Sciences (IES) of the U.S. Department of Education (Award #R305A220386).

References

1. Adams, M.A., Hurley, J.C., Todd, M., Bhuiyan, N., Jarrett, C.L., Tucker, W.J., Hollingshead, K.E., Angadi, S.S.: Adaptive goal setting and financial incentives: a 2× 2 factorial randomized controlled trial to increase adults' physical activity. BMC public health **17**, 1–16 (2017)
2. Alwahbi, A.: The use of contingency contracting in educational settings: A review of the literature. Educational Research and Reviews **15**(6), 327–335 (2020)
3. Baker, R., Walonoski, J., Heffernan, N., Roll, I., Corbett, A., Koedinger, K.: Why students engage in "gaming the system" behavior in interactive learning environments. Journal of Interactive Learning Research **19**(2), 185–224 (2008)
4. Bashkov, B.M.: Assessing the impact of ixl math over three years: A quasi-experimental study. ESSA Research Report. Online Submission (2021)
5. Bolker, B.M.: Linear and generalized linear mixed models. Ecological statistics: contemporary theory and application pp. 309–333 (2015)
6. Borchers, C., Ooge, J., Peng, C., Aleven, V.: How learner control and explainable learning analytics about skill mastery shape student desires to finish and avoid loss in tutored practice. In: Proceedings of the 15th International Learning Analytics and Knowledge Conference. pp. 810–816 (2025)
7. Cerasoli, C.P., Nicklin, J.M., Ford, M.T.: Intrinsic motivation and extrinsic incentives jointly predict performance: a 40-year meta-analysis. Psychological bulletin **140**(4), 980 (2014)
8. Chang, D.H., Lin, M.P.C., Hajian, S., Wang, Q.Q.: Educational design principles of using ai chatbot that supports self-regulated learning in education: Goal setting, feedback, and personalization. Sustainability **15**(17), 12921 (2023)
9. Cooper, J.: Applied behavior analysis. Pearson/Merrill-Prentice Hall (2007)
10. DeMink-Carthew, J., Olofson, M.W., LeGeros, L., Netcoh, S., Hennessey, S.: An analysis of approaches to goal setting in middle grades personalized learning environments. RMLE Online **40**(10), 1–11 (2017)
11. Duffy, M.C., Azevedo, R.: Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. Computers in Human Behavior **52**, 338–348 (2015)
12. Gold, G., Borchers, C., Carvalho, P.F.: Students' academic performance and goal orientation relate to initial knowledge but not learning rate. In: Companion Proceedings of the 14th International Conference on Learning Analytics & Knowledge (LAK). Kyoto, Japan (2024)
13. Hadwin, A.F., Webster, E.A.: Calibration in goal setting: Examining the nature of judgments of confidence. Learning and instruction **24**, 37–47 (2013)
14. Holstein, K., Aleven, V., Rummel, N.: A conceptual framework for human–ai hybrid adaptivity in education. In: Artificial Intelligence in Education: 21st International Conference, AIED 2020, Ifrane, Morocco, July 6–10, 2020, Proceedings, Part I 21. pp. 240–254. Springer (2020)
15. Kahle, A.L., Kelley, M.L.: Children's homework problems: A comparison of goal setting and parent training. Behavior therapy **25**(2), 275–290 (1994)
16. Koedinger, K.R., Carvalho, P.F., Liu, R., McLaughlin, E.A.: An astonishing regularity in student learning rate. Proceedings of the National Academy of Sciences **120**(13), e2221311120 (2023)
17. Koedinger, K.R., Kim, J., Jia, J.Z., McLaughlin, E.A., Bier, N.L.: Learning is not a spectator sport: Doing is better than watching for learning from a mooc. In: Proceedings of the second (2015) ACM conference on learning@ scale. pp. 111–120 (2015)

18. Locke, E.A., Latham, G.P.: Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. American psychologist **57**(9), 705 (2002)
19. Locke, E.A., Latham, G.P.: The development of goal setting theory: A half century retrospective. Motivation Science **5**(2), 93 (2019)
20. Matthews, C.D.M.: Middle School Mathematics Teachers' Use of Diverse Mathematics Educational Technology Tools: IXL, i-Ready, and/or MATHia. Ph.D. thesis, Walden University (2025)
21. McDowall, D., McCleary, R., Bartos, B.J.: Interrupted time series analysis. Oxford University Press (2019)
22. Muis, K.R., Winne, P.H., Ranellucci, J.: The role of calibration bias and performance feedback in achievement goal regulation. International Education Research **4**(1), 14–36 (2016)
23. Peacock, G.G., Ervin, R.A., Daly, E.J., Merrell, K.W.: Practical handbook of school psychology: Effective practices for the 21st century. Guilford Press (2009)
24. Peng, C., Borchers, C., Aleven, V.: Designing homework support tools for middle school mathematics using intelligent tutoring systems. In: Proceedings of the 18th International Conference of the Learning Sciences-ICLS 2024, pp. 1730-1733. International Society of the Learning Sciences (2024)
25. Shogren, K.A., Mazzotti, V.L., Hicks, T.A., Raley, S.K., Gerasimova, D., Pace, J.R., Kwiatek, S.M., Fredrick, D., Stewart-Ginsburg, J.H., Chapman, R., et al.: The goal setting challenge app: Impact on transition goal attainment outcomes of students with disabilities. Career Development and Transition for Exceptional Individuals **47**(1), 4–18 (2024)
26. Simpson, M.A., Norberg, K.A., Fancsali, S.E.: Replicating an "astonishing regularity in student learning rate". In: Proceedings of the 17th International Conference on Educational Data Mining. pp. 420–425 (2024)
27. Thomas, D., Yang, X., Gupta, S., Adeniran, A., Mclaughlin, E., Koedinger, K.: When the tutor becomes the student: Design and evaluation of efficient scenario-based lessons for tutors. In: LAK23: 13th International Learning Analytics and Knowledge Conference. pp. 250–261 (2023)
28. Thomas, D.R., Lin, J., Gatz, E., Gurung, A., Gupta, S., Norberg, K., Fancsali, S.E., Aleven, V., Branstetter, L., Brunskill, E., et al.: Improving student learning with hybrid human-ai tutoring: A three-study quasi-experimental investigation. In: Proceedings of the 14th Learning Analytics and Knowledge Conference. pp. 404–415 (2024)
29. Wäschle, K., Allgaier, A., Lachner, A., Fink, S., Nückles, M.: Procrastination and self-efficacy: Tracing vicious and virtuous circles in self-regulated learning. Learning and instruction **29**, 103–114 (2014)