

Draft for Public Review – Ph.D. Thesis Proposal

Personalized Goal Setting Support to Enhance Technology-Supported Learning

Conrad Borchers

Human-Computer Interaction Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213, USA
cborcher@cs.cmu.edu

August 24, 2025

Thesis Committee:

Vincent Aleven, Carnegie Mellon University, Co-Chair
Kenneth R. Koedinger, Carnegie Mellon University, Co-Chair
Geoff Kaufman, Carnegie Mellon University
Roger Azevedo, University of Central Florida
Emma Brunskill, Stanford University

*Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy.*

Copyright © 2025 Conrad Borchers

Abstract

Sustained student engagement remains a persistent challenge in realizing the full potential of AI-enhanced learning environments such as intelligent tutoring systems. Goal setting—defining clear, quantifiable objectives for effort and performance—has been shown to enhance motivation and performance. However, traditional goal-setting interventions in K-12 education, often reliant on teacher, parent, or tutor facilitation through paper-based contracts, face scalability and integration challenges in digital learning contexts.

This thesis addresses this gap by designing and studying technology-mediated, personalized, and data-driven approaches to facilitate and scale goal support, defined as structures that aid goal setting, feedback, and the distribution of rewards contingent on goal completion. Drawing inspiration from how traditional goal setting facilitates accountability between learners and a human tutor, the research explores how intelligent systems can similarly scaffold goal and effort regulation processes while preserving student agency in setting goals and learning associated metacognitive skills of effort regulation, calibration, and goal selection.

Specifically, this work presents the design, implementation, and evaluation of intelligent goal-setting support systems embedded within active learning platforms. Theoretical predictions center around estimating the utility of data-driven goal recommendations, performance feedback, and goal tracking mechanisms using learning system log data. Experimental studies in K-12 classroom settings, centered in grades 5-9, evaluate the effectiveness of adaptive goal-setting interventions on student engagement and learning outcomes. A secondary focus is placed on modeling how achievement events and goal adjustment moments influence student trajectories over time, investigating whether adaptive support can propel students into positive achievement cycles that may foster independent effort regulation.

Empirical findings demonstrate significant engagement and skill mastery benefits of goal-setting support in classrooms beyond tutoring without goal setting. Goal achievement was linked to larger intervention benefits, and adaptive goals, as well as data-driven feedback to guide goal adjustments, were found to enhance student goal completion rates significantly. Though broadly effective for over 80% of students, students who benefited less demonstrated higher levels of baseline engagement. It is conjectured that these students may already exhibit high degrees of intrinsic motivation, which are known to offset reward benefits for performance when rewards are clearly tied to performance or goals are perceived as coercive.

To test this account, the final proposed study will investigate the at-scale application of more frequent goal adjustments delivered through dashboards that enhance frequent student goal choice without researcher facilitation on-site. Explanatory variables assessed through surveys of intrinsic motivation, perceived choice, as well as goal orientation will further explicate and test this interpretation. Findings from this study are expected to guide future efforts in psychometric assessment that could guide finer-grained adaptivity in effort goal delivery through data.

Findings of this thesis advance SRL theories in adaptive learning systems by adding an adaptivity loop to the regulation of effort outside of practice, in addition to metacognitive support during learning. Practically, this work provides empirical evidence on the efficacy of adaptive goal recommendations and offers a scalable solution that alleviates teacher workload. Finally, this research advances scientific understanding of the longitudinal effects of incentivizing effort

regulation through extrinsic rewards across different learners, finding that beneficial goal-setting effects are generally sustained over time, but also can lead to differential effects if learners lack goal choice or scaffolds to set realistic goals that they can achieve. These insights can help make educational technologies for deliberate practice in STEM more effective for knowledge acquisition.

Keywords: Goal Setting, Effort Regulation, Intelligent Tutoring Systems, Learning Analytics, Educational Technology, Homework Support, Low-Cost Interventions, Motivational Scaffolding, Human-Computer Interaction, Hybrid Tutoring, Learning at Scale, Educational Data Mining.

Contents

1	Introduction	1
1.1	Research Objectives	2
1.2	Expected Contributions	2
1.3	Structure of the Thesis	3
2	Background and Related Work	4
2.1	Principles Guiding Data-Driven Goal Support	4
2.1.1	Principle 1: Provide Goal Achievement Feedback	4
2.1.2	Principle 2: Adapt Goal Recommendations to Past Performance	5
2.1.3	Principle 3: Retain Learner Autonomy for Extrinsic Goals	5
2.2	Theory of Change Overview	6
3	Goal Setting Contracts: A Low-Cost Intervention in Tutored Practice	9
3.1	Introduction and Related Work	9
3.1.1	Instructional vs. Motivational Caregiver Tools and Roles	10
3.2	Methods	11
3.2.1	Participants and Recruitment	11
3.2.2	Tutoring System Probe	12
3.2.3	Homework Contract Probe	12
3.2.4	Procedure	13
3.2.5	Data Analysis Methods	14
3.3	Results	15
3.3.1	RQ1 Engagement rates	15
3.3.2	RQ2: Obstacles to Caregiver Involvement	16
3.4	Discussion	17
3.4.1	RQ1: How Much Did Carergivers Engage With Each Probe?	17
3.4.2	RQ2: Opportunities to Overcome Involvement Obstacles	18
3.4.3	Limitations	18
3.5	Conclusion	19
4	Learning Benefits of Goal Setting with Rewards in Hybrid Tutoring	20
4.1	Introduction and Related Work	20
4.2	Methods	22
4.2.1	Sample and Recruitment	22
4.2.2	Materials	22

4.2.3	Procedures	23
4.2.4	Measures and Data Preprocessing	25
4.2.5	Data Analysis	25
4.3	Results	27
4.3.1	Descriptive Differences in Time Spent	27
4.3.2	Interrupted Time Series Modeling	27
4.3.3	Did Students Also Learn More During the Intervention?	28
4.4	Discussion	28
4.4.1	Engagement in Practice Improved Through Goal Setting (RQ1)	28
4.4.2	Changes in Practice Time Remained Stable Over Time (RQ2)	29
4.4.3	Skill Acquisition Benefits Exceed Engagement Benefits (RQ3)	29
4.4.4	Limitations and Future Work	29
4.5	Conclusion	30
5	Differential Effects of Adaptive Goal Setting and Achievement	32
5.1	Introduction and Related Work	33
5.1.1	The Role of Student Autonomy Under Extrinsic Rewards	34
5.1.2	Data-Driven Goal Support	35
5.1.3	The Present Study	35
5.2	Methods	36
5.2.1	Sample and Study Context	36
5.2.2	Experimental Design	36
5.2.3	Procedures and Intervention	37
5.2.4	Data Analysis Methods	38
5.3	Results	42
5.3.1	RQ1a: Do Adaptive, Self-Set Goals Improve Goal Achievement over Static, Teacher-Set Goals?	42
5.3.2	RQ1b: Do Adaptive, Self-Set Goals Enhance Goal Achievement Momentum?	43
5.3.3	RQ2: Do Adaptive, Self-Set Goals Improve Effort and Engagement?	44
5.3.4	RQ3: Do individual differences in prior effort and achievement explain differences goal-setting benefits?	46
5.4	Discussion	48
5.4.1	Benefits of Goal Adjustment and Selection for Goal Achievement	48
5.4.2	Low-Effort Students Benefited More from Goal Support	48
5.4.3	Intrinsic Motivation as a Potential Moderator of Intervention Differences	49
5.4.4	Explanatory Variables of Goal Intervention Benefit Beyond Intrinsic Motivation	49
5.4.5	Limitations and Future Work	50
5.5	Conclusion	51
6	Goal Support at Scale Beyond Effort and Proficiency	53
6.1	Design Rationale	53
6.2	Research Questions	54
6.3	Hypotheses	54
6.4	Study Design	55

6.5	Preliminary Interface Design	56
6.6	Plan B (Fall 2025 or Spring 2026): Custom Dashboard	57
6.7	Anticipated Contributions	57
6.8	Timeline of Completion	58
Bibliography		60

Chapter 1

Introduction

Active learning environments—such as intelligent tutoring systems and teachable agents—hold promise for improving educational outcomes by engaging students in sustained, effortful problem-solving activities [62]. However, these benefits are highly dependent on students’ ongoing motivation and commitment to practice. Although adaptive instructional systems effectively personalize learning experiences based on individual needs, their overall impact is constrained when students fail to invest sufficient time and effort [41].

Practice goal setting—the process of establishing clear, quantifiable objectives related to effort and performance, often linked to contingent rewards—has been shown to be a effective method for increasing student effort during learning [5]. Traditionally, such interventions have relied on paper-based contracts involving parents or teachers [5, 80]. However, these approaches face limitations in terms of scalability and integration with digital learning environments [82].

Thesis statement: *In this thesis, I contend that embedding evidence-based strategies for goal-setting may not only replicate, but exceed their known benefits [5], when effectively embedded into technology-enhanced learning environments that provide continuous feedback and data-driven personalization at scale.*

There are compelling reasons to conjecture that embedding practice goal setting within digital learning environments introduces unique affordances. First, these environments provide objective measures of student activity, allowing for continuous, automated feedback on effort and self-regulation—features that have proven critical in prior research [84]. Second, when well-designed, adaptive learning systems improvement across all student groups and thereby helping to close opportunity gaps in learning. This means that per unit of additional instruction learners engage with through goal setting, they may benefit more than from traditional instruction, and opportunity gaps may close [5, 62]. Third, such systems offer scalability, allowing human resources—such as teacher attention—to be distributed more efficiently, which we is known to improve or at least modulate in-tutor learning [20, 56, 59, 60].

Aligned with the principle of human-AI complementarity in educational technology design [52], the interventions examined in this dissertation extend beyond traditional technology support for self-regulated learning (SRL) by focusing on more foundational self-regulation processes, particularly effort regulation. In contrast, most AI-driven learning technologies designed to support SRL have concentrated narrowly on cognitive and metacognitive strategies during task execution [8, 68]. However, foundational SRL processes—such as managing effort and setting or calibrating goals—remain underexplored within intelligent learning environments, despite their

critical role in sustaining student engagement and achievement [70]. Notably, these processes often occur outside the immediate context of tutoring systems and learning sessions, taking place before learning begins and after it concludes [47]. The theoretical contribution of this work is to conceptualize and investigate how learning technologies can support these peripheral yet essential phases of self-regulation. By leveraging adaptive, data-driven insights, this research proposes interventions that guide students in managing their effort and goals beyond the confines of task-level support, thus expanding the scope and impact of intelligent learning environments and their adaptivity.

1.1 Research Objectives

This research seeks to embed adaptive goal-setting support into active learning systems. By leveraging learner performance data, the proposed research facilitates feedback delivery, personalized goal recommendations, and accountability structures that have traditionally relied on human facilitation [18]. Building on prior work in adaptive behavioral feedback [1] and the calibration of self-regulated effort through feedback [48], this research examines how AI-driven goal-setting interventions can promote sustained student engagement and improved learning outcomes. The final proposed study will aim at largely automating and scaling the process of goal support through a human-AI hybrid tutoring program or dashboard, including targeted support for students not meeting their goals. The central objective of this thesis is to explore the design, implementation, and evaluation of intelligent practice goal-setting support in active learning environments. Specifically, this research aims to:

- Evaluate the impact of practice goal-setting support on student effort and skill mastery in real-world educational settings using educational technologies based on active problem solving practice.
- Analyze how adaptive feedback loops influence student effort calibration and goal adjustment, and achievement, and learning.
- Develop scalable means that automates feedback and provides personalized goal recommendations.

1.2 Expected Contributions

This research contributes to both theoretical and practical advancements in AI-supported active learning. The anticipated contributions include:

- Frameworks for integrating goal-setting support into AI-driven learning environments, bridging research on goal attainment, self-regulation, and adaptive learning technologies.
- Empirical evidence on the effectiveness of adaptive, data-driven goal recommendations in improving student engagement and learning outcomes.
- Scalable methods for practice goal setting that reduces teacher workload while maintaining student autonomy and personalized learning pathways.
- Insights into the role of adaptive feedback in supporting foundational self-regulation activities necessary for effective goal pursuit in AI-enhanced learning systems.

1.3 Structure of the Thesis

The remainder of this thesis is organized as follows:

- Chapter 2 provides a comprehensive review of related work in goal setting, self-regulated learning, and active learning systems.
- Chapter 3 summarizes early design research and a pilot study with parents, traditionally stakeholders of goal contracts, paving the way for interventions that followed.
- Chapter 4 presents an empirical evaluation of goal setting with rewards in personalized learning through quasi-experimental methods.
- Chapter 5 focuses on differential outcomes of the intervention, including those predicated on adaptive goal setting.
- Chapter 6 describes the final proposed thesis study addressing scale and additional support for students with low goal achievement with the thesis completion timeline.

Chapter 2

Background and Related Work

This chapter was adapted from my published doctoral consortium paper:

Conrad Borchers, Kenneth R. Koedinger, and Vincent Aleven. 2025. Intelligent Support for Practice Goal Setting to Enhance Learning. In *Proceedings of the 26th International Conference on Artificial Intelligence in Education (AIED '25)*. Palermo, Italy.

Summary Statement in Relationship to Thesis

This chapter reviews existing research on goal setting, self-regulated learning (SRL), and their integration into AI-driven educational technologies. It identifies gaps in current approaches and motivates the need for intelligent, scalable goal-setting support in active learning environments. The chapter establishes a theory of change in adaptive practice goal setting motivated by empirical findings and theoretical considerations grounded in prior research.

2.1 Principles Guiding Data-Driven Goal Support

The chapter begins by outlining literature-derived principles that could make data-driven goal support effective. So below I outline each of these principles and relate them to this thesis research. Afterwards I then derive specific hypotheses and a theory of change building on top of the principles.

2.1.1 Principle 1: Provide Goal Achievement Feedback

Feedback is a cornerstone of effective learning, benefiting both domain-specific knowledge acquisition and broader self-regulation skills [84]. Regular, targeted feedback improves students' ability to monitor and adjust their effort toward achieving learning goals. In the context of AI-supported education, adaptive feedback—tailored in response to learner performance—has demonstrated effectiveness across various domains [1, 84]. However, most adaptive interventions in educational technology have focused narrowly on providing in-task support, such as

hints or scaffolding for problem-solving [8, 68], while neglecting foundational self-regulation activities like effort regulation and goal calibration [43, 70]. Similarly, seminal models of adaptivity in technology-enhanced learning systems, such as intelligent tutoring systems, do not consider adaptivity in relationship to student effort regulation beyond adaptivity to affect and engagement once learning is already underway [4].

What could feedback and adaptivity look like in the context of goal setting where students observe and potentially adjust goals over time? And how effective is that for learning? This question is a central one in this thesis. Recent studies highlight the potential of integrating adaptive feedback with data-driven goal recommendations to enhance students' capacity for self-regulation [18, 82]. Such interventions could not only guide students in adjusting their goals based on performance (as described next) but also sustain engagement by fostering reflective practice. Despite their promise, adaptive goal-setting mechanisms remain underexplored in AI-supported learning systems, presenting a critical opportunity for future research and design innovation.

2.1.2 Principle 2: Adapt Goal Recommendations to Past Performance

Goal-setting theory asserts that specific, challenging goals enhance performance by directing attention, mobilizing effort, and fostering persistence [70]. In educational contexts, this translates to practice goals framed around quantifiable objectives, such as problem completion or time allocation, which are proven to improve student engagement and outcomes [5, 57]. To maximize effectiveness, goal recommendations should be dynamically adapted to students' historical performance averages, based on evidence from behavioral health interventions [1]. This adaptive approach ensures that goals remain attainable yet sufficiently challenging, thereby optimizing the probability of success and amplifying engagement benefits [97].

Traditional implementations of goal-setting, including paper-based contracts involving parents or teachers, leverage extrinsic rewards and social accountability to sustain effort [5]. However, these methods face limitations in scalability and responsiveness, particularly in digital learning environments where real-time data can inform more nuanced adaptations. By automating goal-setting processes and embedding them within educational technologies, it becomes feasible to support large student populations without increasing the workload for teachers and caregivers [82].

Therefore, a central objective of this thesis is to devise interventions that can leverage learner data to adjust and recommend goals over time to improve longitudinal performance in addition to just providing feedback on performance as such. This could give learners the opportunity to gradually improve their learning achievement. Yet in the context of K-12 education such interventions have not yet been studied and methods for scaling such support for example through dashboards are also under explored.

2.1.3 Principle 3: Retain Learner Autonomy for Extrinsic Goals

Self-determination theory distinguishes between intrinsic motivation, driven by inherent interest, and extrinsic motivation, driven by external outcomes [35, 85]. Though often seen as oppositional, intrinsic and extrinsic motivators can jointly enhance performance, including in education [28]. Crucially, extrinsic motivation varies in autonomy—from externally imposed (e.g.,

avoiding punishment) to self-endorsed (e.g., pursuing meaningful goals) [83, 85]. This thesis defines autonomy as the degree of personal choice in motivation, consistent with SDT [35].

Autonomy has been shown to support goal progress by increasing effort, reducing conflict, and fostering readiness for change [64], with benefits confirmed in adolescents and young adults [65]. Yet, some students may not view math practice as worthwhile, perceiving classroom goals as imposed—beliefs that predict effort and achievement [46]. Whether student-involved goal setting improves outcomes remains an open question, particularly given the limited research on autonomy in naturalistic educational settings [67].

Given the notable empirical evidence to retain autonomy in goal setting with extrinsic rewards [28, 79] as well as my own design research demonstrating that students prefer final control over goals when they are recommended through AI [20], I build on these principles to study how effective learner-centered goal achievement in tandem with data driven recommendations is in authentic K-12 classrooms. This high level of learner oversight over the goal-setting process is also true to the original intervention I built on where parents or teachers jointly set and negotiate goals with the student [5, 57].

Yet, scaling such learner-centric interventions that retain autonomy within digital learning environments requires thoughtful integration. Embedding social accountability structures—such as tutor-mediated goal reviews or digital contracts—into AI-supported systems could sustain student motivation while preserving scalability.

2.2 Theory of Change Overview

In sum, data-driven goal support in AI-enhanced learning, as conceived in this thesis, is founded on the seamless integration of adaptive feedback, performance-based goal adjustments, and learner-mediated autonomy. By addressing these foundational principles, learning technologies could move beyond task-specific scaffolding toward fostering sustained student engagement and achievement. Addressing this gap is the primary contribution of the present thesis.

The presented research investigates how intelligent practice goal-setting support can improve student engagement and learning in active learning environments. Building on the principles outlined above, and grounded in foundational goal-setting and effort regulation theories [70], the study proposes and tests three central hypotheses that inform both the design and evaluation of the interventions.

The first hypothesis (H1) posits that *actively involving students in goal setting, while preserving their control over final goal selection, enhances effort regulation and engagement during online practice activities*. This claim builds on evidence from autonomy benefits in goal progress toward extrinsic rewards [64, 65] and past practices of goal-setting contracts in non-digital contexts, where learners are setting goals jointly with a teacher or parent [5, 80]. Does the same hold true when the students independently set goals with historical performance averages based on their data are used to negotiate practice goals with limited human oversight?

The second hypothesis (H2) focuses on the dynamic aspect of goal pursuit, hypothesizing that *regular feedback and adaptive goal recommendations based on students' historical performance improve the likelihood of goal achievement, thereby sustaining engagement and learning over time*. Feedback plays a dual role: it not only facilitates general domain learning but also strengthens students' self-regulation by improving their ability to monitor and adjust effort toward goal attain-

ment [84]. Furthermore, aligning goal recommendations with students' historical performance averages draws from past successful behavioral science interventions in adult populations [1]. These adaptive adjustments are expected to enhance the probability of goal achievement and, consequently, maintain student engagement, given the known motivational benefits of meeting goals, such as reduced procrastination [97].

The third hypothesis (H3) asserts that *achieving practice goals, especially when reinforced through rewards promotes further engagement and learning*. Empirical studies have demonstrated that goal-setting contracts with teachers, tutors, or caregivers can significantly boost student effort and learning outcomes [5, 57]. Such interventions leverage external accountability and contingent rewards to motivate students, yet, again, have seen little evidence when delivered at scale through data, AI, and limited human supervision (as is common in large classrooms or at-scale remote tutoring). Past research predicts that goal achievement boost motivational longitudinally, while lack of achievement can have the opposite effect [97]. To investigate this hypothesis, I model both the overall intervention effect as well as its relationship to these fine-grain, longitudinal achievement events.

Towards the proposed final thesis study, scalability of past goal interventions remains limited due to the need for continuous adult involvement. This research aims to address that limitation by automating goal recommendations and feedback mechanisms while preserving the motivational benefits of accountability provided by teachers or human tutors. By embedding goal-setting support within digital learning environments, data-driven goal setting could promote sustained engagement at scale, complementing human facilitation with scalable technological solutions [52].

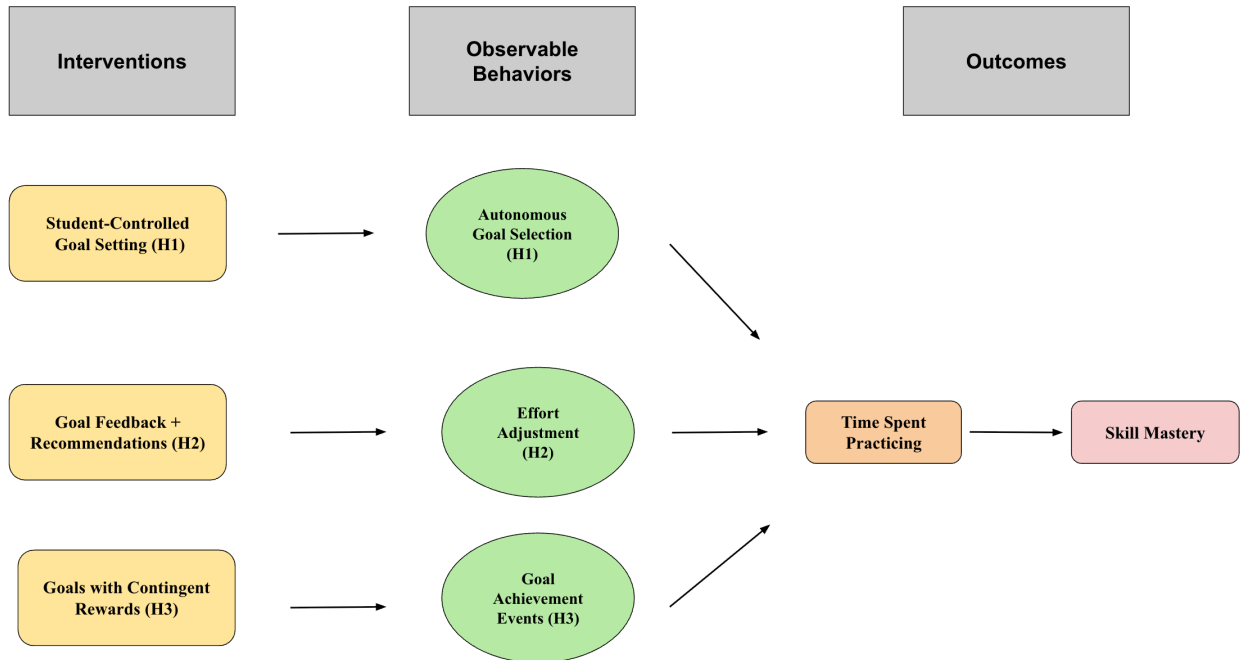


Figure 2.1: Theory of change diagram with three hypotheses: H1 (Active Goal Setting), H2 (Adaptive Goal Adjustments), and H3 (Impact of Goal Achievement). These hypotheses are examined in relation to practice engagement and skill mastery outcomes.

Together, these hypotheses form the basis of the thesis' **theory of change**, illustrated in Figure 2.1. The proposed model links student-controlled goal setting (H1), adaptive goal feedback and adjustments (H2), and the impact of goal achievement (H3) to key outcomes such as practice engagement and skill mastery. As I go on to discuss, I assume a general link whereby more practice engagement (most commonly measured in time across educational technologies) leads to more learning opportunities to practice cognitive skills and, hence, content mastery. The empirical relationship between opportunities and mastery in educational technologies providing practice opportunities with feedback and as-needed instruction through hints is backed up by large scale empirical evidence of tutoring system data [41, 62, 89]. While time might not always lead to productive learning opportunities, and the efficiency of time use may differ between students [11, 18], in general, I expect more time to lead more practice.

Each hypothesis is systematically examined across the studies presented in this dissertation. Chapter 5 tests **H1** by evaluating the effectiveness of student-led goal setting on engagement and learning outcomes through an experiment in contrast to teacher assigned goals. Chapter 5 also addresses **H2** by comparing static goals to adaptive, data-driven goal recommendations and investigating how goal achievement trajectories influence future success. The general effectiveness of goal setting with contingent rewards (H3) is both analyzed in Chapter 4 as well as through analyses of how rates of goal achievement relate to the general intervention effect (also in Chapter 5). Finally, Chapter 6 outlines proposed work that extends the findings related to **H3**, specifically targeting improved support for students who struggle to meet their goals.

To operationalize these hypotheses, goal-setting support interventions were designed to integrate with existing learning platforms data that track student activity. Leveraging real-time performance data, the interventions dynamically inform personalized goal recommendations while maintaining student autonomy in the final goal selection (see Chapter 5).

Chapter 3

Goal Setting Contracts: A Low-Cost Intervention in Tutored Practice

This chapter was adapted from my accepted conference paper:

Conrad Borchers, Ha Tien Nguyen, Paulo F. Carvalho, Kenneth R. Koedinger, and Vincent Aleven.
2025. Goal Setting Engages More Caregivers Than Instructional Support in Online Math Homework.
Accepted to the European Conference on Technology Enhanced Learning (EC-TEL '25).

Summary Statement in Relationship to Thesis

In this chapter, I report on a pilot study of a lightweight goal-setting contract implemented in a middle school classroom and integrated into a math tutoring system. The contract was compared to a more involved intervention that had been co-designed with caregivers and required hands-on support via in-system collaboration. Despite its simplicity, the goal-setting contract achieved substantially higher uptake among students, suggesting that low-friction goal support may offer a more scalable alternative to resource-intensive designs. These findings motivate the broader integration of classroom-based goal contracts into intelligent learning systems to support self-driven and personalized student practice, as is explored in the ensuing chapters of this thesis.

3.1 Introduction and Related Work

Caregiver (i.e., a parent or member of a child's caring community) involvement can be critical for academic achievement, contributing motivational and instructional homework support. Caregiver involvement has been linked to higher grades and better test performance [55]. However, many caregivers face barriers, such as limited time, resources, and content knowledge, hindering homework support [81], and exacerbating opportunity gaps.

What motivates caregivers to support their students with homework? Seminal work [53] identifies three key factors: (a) perceptions of responsibility, (b) beliefs that their help is effective, and (c) perceptions of whether their student and school expect their involvement. However,

since this framework was established nearly three decades ago, the homework environment has become increasingly digital [75], potentially altering caregiver roles. Recent research highlights new opportunities for caregiver involvement brought by the increasing availability of educational technologies in K-12 education, such as improving teacher-student communication through instant messaging and supporting student self-regulation during online homework through learning analytics [45, 78, 82].

Yet, these studies do not fully explore how novel technologies and interventions can encourage caregiver involvement in ways aligned with [53]’s model. For instance, technology could enhance caregivers’ perceived efficacy by helping them better understand their student’s homework, particularly in challenging subjects like mathematics [78]. Alternatively, it could strengthen motivational support as another key caregiver homework role [82]. While promising, the adoption rates of such interventions remain unclear. This study addresses this gap by comparing two novel interventions designed to support caregiver involvement in digital math homework and their adoption rates, including for known-to-be effective interventions such as goal setting [5, 18]. As a secondary contribution, we study the obstacles to caregiver support in students’ tutoring system practice that technology may introduce or solve.

3.1.1 Instructional vs. Motivational Caregiver Tools and Roles

The current study examines the feasibility and affordances of two distinct approaches to caregiver involvement in helping students engage in tutoring system practice. The first approach, a “hands-on” design, encourages caregivers to collaborate with students and bridge knowledge gaps, argued to address curricular challenges caregivers often report [78, 81]. The second approach, a “hands-off” goal-setting model, leverages caregiver-student goal setting, which has been found effective in past research but has not been studied in tutoring systems [5].

As an instructional, hands-on tool, we study intelligent tutoring systems (ITS) that enhance student learning across K–12 and early college settings [66]. While ITS have shown promise in digital homework [39], their potential to help caregivers support their students in ITS practice remains underexplored. Caregivers often excel at providing emotional and motivational encouragement but often lack confidence in instructional roles [50]. This raises the following key questions: How can an ITS, suitably enhanced, complement caregivers’ strengths? Can these systems enhance caregivers’ instructional capabilities [53]? Research has highlighted caregivers’ need for tools that help bridge content knowledge gaps [78]. Nguyen et al. [78] found that caregivers favor conversational support systems that offer step-level guidance during homework when integrated with ITS, underscoring the importance of integrating instructional aids into ITS.

In contrast, as a motivational, hands-off tool, we study a goal-setting approach that draws from research on homework behavioral contracts [80]. Cooper [30] defines contingency contracts as written agreements between two parties that specify roles, assigned tasks, designated rewards for the student, and the conditions required to earn rewards. Lab studies indicate that student practice time increases when parents and students jointly establish goals [57]. More recent literature reviews further support the effectiveness of contingency contracting in fostering cognitive and non-cognitive skill development in students [5].

While tutoring systems can help address instructional barriers faced by caregivers, goal-setting contracts may be more accessible due to requiring less active engagement. We conducted technology probe studies in two middle schools with distinct student populations. This setting

allowed us to examine differences in the feasibility, adoption, and perceived effectiveness of these tools. We ask:

RQ1: To what extent do caregivers engage with (hands-on) intelligent tutoring and (hands-off) goal-setting probes for homework support?

RQ2: What are strategies to overcome common obstacles to caregiver engagement in computer-supported homework?

3.2 Methods

This study investigates the feasibility and desirability of novel designs for middle school caregiver homework support in mathematics ITS. Technology probes allow us to (1) collect information about real-world use and user, (2) field-test the novel caregiver tool, and (3) generate design insights into homework support tools in middle school math supported by learning technologies [54].

3.2.1 Participants and Recruitment

The study involved two American middle schools: East School, a suburban school in the Northeastern US, and West School, a suburban school in the Pacific Northwest. Both schools were recruited through previously established research partnerships with schools and school districts. Permission to conduct research was obtained from both schools following the approved IRB protocol.

We initiated contact with participating mathematics teachers by obtaining research permissions from school administrators. Collaborative discussions with administrators and teachers helped determine the classes involved in the study and set adequate data collection time frames and mathematics content.

Students of the East School participated in the study across two eighth-grade classes, taught by the same teacher and totalling 44 students. One class, Class 1, consisting of 24 students, was taught a standard eighth-grade math curriculum, while the other class, Class 2, with 20 students, followed an advanced math track based on state test scores. The student population of East School is predominantly Caucasian (90%) and 45% of the student population was classified as low-income based on Free/Reduced Lunch (FRL) statistics available. Less than 5% of the students are English Language Learners (ELL). Overall, 35% of the total school-wide student population met state math standards, representing curricular proficiency expectations for students of that grade level.

At West School, a classroom of 31 students participated. The classroom comprised standard eighth-grade math students, special education students, and English language learners (ELL). One teacher acted as the primary instructor. A second teacher provided additional support, particularly for the special education students. Based on national statistics, the student population at West School is 20% Caucasian and 75% of the students were classified as low-income based on Free/Reduced Lunch (FRL) and official state government resources, and 25% were ELL. Overall, 30% of the total school-wide student population met math standards assessed by the Smarter Balanced Assessments. All school demographics are rounded to 5% points to preserve anonymity.

3.2.2 Tutoring System Probe

3.2.2.1 Recruitment Process

Student-caregiver pairs in both schools were recruited to receive access to the caregiver support probe, which involved caregivers using the tutoring system “hands-on.” First, a study strategy was determined with the participating teachers. After securing approval from the school and the teacher, the participating mathematics teacher distributed an email to all caregivers containing information about the study and its objectives. These details were provided as an email attachment containing an official letter about ten days before the study was to start. This letter explained that caregivers had the opportunity to participate in a study trialing a novel homework support tool designed for caregivers, emphasizing that participation was voluntary and detailing the data collection process. Based on teacher discussions, translated study materials were provided to multilingual households at West School so that no household was excluded from participating in the study.

In addition to the letter, caregivers completed an informal, 30-minute onboarding Zoom session, which walked them through the tool’s functionalities. One research team member walked participants through the system using a slideshow, as well as live demos of the tool. Participants were given an opportunity to have any questions answered and to test the system live through test accounts during the session.

Across both schools, students and caregivers could sign up for the study and continue to do so until the last day. Caregiver consent and student assent were obtained for tool access. One caregiver per household could participate. Caregivers signed up for the study by filling out an online consent form in Qualtrics, followed by a brief survey on their household’s demographics. Finally, all caregivers received a third letter that allowed them to opt out of having their student’s practice log data retained for research purposes. No caregiver, however, opted out of that option.

3.2.2.2 Tutoring System Technology Probe Design

Details on the hands-on caregiver system design and its implementation can be found in past research and is not described in detail in this thesis as it is not the focus of its inquiry [19, 78, 96].

3.2.3 Homework Contract Probe

A second design probe was introduced at West School only, focusing on caregiver accountability beyond problem-solving support. This probe, and the decision to only trial it at West School, emerged from preliminary thematic analyses of interviews at East School, which revealed caregivers’ inability to provide accountability support through the tool. The probe aimed to assess caregivers’ perceived usefulness and response rates, aligning with the broader research goal of developing an intervention accessible to diverse caregivers. To that end, this probe did not require a formal sign-up process but was handed out to students similar to a homework assignment requiring caregiver sign-off at home, as detailed below.

The decision to implement a goal-setting contract was also informed by design research indicating that students prefer goal-setting support over other evidence-based, offline homework interventions when using tutoring systems [82]. Adapted from recommendations in [80], the

homework contract was performed on paper and encouraged caregivers and students to negotiate practice goals and, optionally, rewards. Caregivers could also commit to providing tutoring or other forms of support. Fig. 3.1 shows a contract example. To ensure accessibility, the contract was translated into multiple languages for non-English-speaking caregivers.

Homework Contract

Purpose: This contract will support you (the student) to practice math at home. You will practice math with tutoring software on your laptop or phone. This contract will also support your caregiver/parent in helping you reach your goals. This contract asks you to set goals, expectations, and rewards related to your math practice.

What to do? 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 the student and the parts marked by red should be filled by the caregiver/parent.

As part of the contract, the student agrees to:

- ☐ Spend 15 minutes per day practicing math at home with the tutoring software
If the promise is met, as a parent, I reward the student by:
Rewarding him with 5 dollars
- ☐ Spend 1:30 hour(s) per week practicing math at home with the tutoring software
If the promise is met, as a parent, I reward the student by:
paying 20 dollars if he gets the full 1:30
- ☐ When practicing math at home, I want to master 1 skills every 2 [days/weeks]
[Note: the yellow bars in your tutoring software turn green when you have mastered a skill]
If the promise is met, as a parent, I reward the student by:
taking him to a store of his choosing

[Optional: Fill out the line below with any other promises]

☐ _____

As part of the contract, the caregiver/parent agrees to:

- ☒ Provide the rewards above upon goal completion.
- ☒ Spend around 1:30 hour(s) per week helping my child in their homework.

[Optional: Fill out the line below with any other promises]

☐ _____

Figure 3.1: Example redacted homework contract.

On the first day of the study at West School, a researcher introduced and passed out homework contracts for students to complete at home. A participating teacher reminded the students to complete and submit the contracts to the researcher in class daily. Twelve students returned homework contracts in the first week, and five students returned homework contracts in the second week, with 17 out of 31 (54.8%) homework contracts returned by the end of the study.

3.2.4 Procedure

The study took place over four-week periods. The first two weeks involved classroom practice without caregiver involvement, followed by two weeks with the intelligent caregiver support

module probe to support caregiver engagement. Students practiced linear equations using the Lynnette ITS [72] for about 30 minutes daily in class. Lynnette offers immediate feedback and step-level hints. Daily assignments included 12 problems, with unfinished problems completed at home. For students who finished early, Lynnette integrated additional units on graph interpretation. Log data captured student practice, caregiver engagement, and tool interactions, including system usage time, hint requests, problem completion rates, and caregiver notifications. These metrics were used to assess engagement and identify usage patterns, logged in the standard DataShop format for tutoring systems [61]. A researcher was present in the classroom throughout the study to address technical issues, observe student engagement, and document factors influencing tool use. Teachers maintained their regular classroom routines, offering occasional support during student practice.

3.2.5 Data Analysis Methods

3.2.5.1 Engagement Rates (RQ1)

were defined by the extent of caregiver and student interaction with the probes. Caregiver engagement was measured through: **Caregiver Tool Signup and Usage:** Caregivers were considered engaged if they signed up and recorded at least one session in the tutoring system’s support module based on log data. **Caregiver Tool Signup but No Usage:** Dropout rates reflected caregivers who signed up but did not use the system, based on the absence of logs. **Homework Contract Return:** Engagement with the homework contract probe was measured by the return of a signed contract, whether signed by a caregiver or another household member acting in a caregiver role.

For an exploratory analysis featuring 31 West School students, we measured engagement through average weekly math practice time based on session start and end timestamps in log data. To compare engagement between students who completed the contract and those who did not, we conducted a *t*-test, with students as the analysis unit. We replicated the *t*-test for in-class and out-of-class usage as separate outcomes, as we predicted that out-of-school student engagement might be more sensitive to caregiver support differences. For robustness, we also report Wilcoxon rank-sum tests as we observed skew in our data.

3.2.5.2 Observation Notes (RQ2)

Analyses related to RQ2 derived strategies to overcome common obstacles to caregiver involvement in digital homework. Specifically, researchers observed students during homework practice sessions in classrooms, noting factors that influenced engagement or lack of participation—such as barriers, challenges, or personal preferences—through brief informal discussions with students (and, at times, with participating teachers). Some student responses reflected caregivers’ perspectives on tutoring system support and reasons for non-participation or probe signups. These observations were triangulated with pre- and post-interview data and log data analyses to provide a comprehensive understanding of caregiver involvement challenges.

We used an open-ended thematic approach and consensus-based analysis [31, 49]. Two researchers reviewed class observation notes and separately developed themes related to RQ2. Themes were grouped according to forms of caregiver engagement, involvement styles, and

household obstacles. The researchers then discussed the generated themes iteratively until they reached a consensus.

3.3 Results

3.3.1 RQ1 Engagement rates

We analyzed system log data to gain quantitative insight into caregiver signup rates. Signup rates were generally low (12%). The dropout rate, defined as signing up to use the tool but not using it, was 20% in East School and 50% in West School. While caregiver engagement with

Table 3.1: Comparison of Caregiver Signup Rates for the Instructional, Hands-On Tool between East and West Schools.

	East School	West School	Overall
# Students (Periods)	44 (2)	31 (1)	75 (3)
# Signups (Dropouts)	5 (1)	4 (2)	9 (3)
Signup Rate	11.4%	12.9%	12.0%
Dropout Rate	20.0%	50.0%	33.3%

the caregiver support module was comparatively low at West School, households were relatively receptive to the homework goal-setting contract probe. All households were sent the contract, and 54.8% returned a signed homework contract. Among those, 34.5% agreed to provide help with homework, and 27.6% agreed to combine help with offering rewards. Overall, adoption rates of the hands-on instructional tool (12.0%, $N=75$) and hands-off, goal-setting tool (54.8%, $N=31$) were significantly different based on a two-sample test for equality of proportions, $\chi^2(1) = 19.49, p < .001$.

3.3.1.1 Engagement with Tutoring System after Goal Contract Completion

An exploratory, associational analysis using a sample of 31 West School students was conducted. Completing the homework contract was strongly associated with increased practice in and out of school settings. System log data indicated that students who returned the homework contract practiced more than twice as much per week with the computer-based tutoring system ($M = 67.65$ mins, $SD = 62.83$ mins) compared to those who did not ($M = 34.33$ mins, $SD = 28.07$ mins), $t(29) = 2.03, p = .052$. Due to apparent skew in the data, we confirmed that this difference was also marginally significant using a Wilcoxon rank-sum test, $U = 162.00, p = .054$. The difference was even more pronounced when filtering data outside of regular classroom hours, where students who returned a contract practiced three times as much ($M = 34.32$ mins, $SD = 95.14$ mins) compared to those who did not ($M = 11.54$ mins, $SD = 57.46$ mins), $t(26) = 1.91, p = .067$. A similar trend was observed for in-class practice ($M = 13.76$ mins, $SD = 4.81$ mins for contract returners vs. $M = 10.25$ mins, $SD = 6.45$ mins for non-returners), $t(29) = 1.62, p = .117$. Both differences were similarly marginal when tested using a Wilcoxon rank-sum test (p 's $> .054$).

3.3.2 RQ2: Obstacles to Caregiver Involvement

3.3.2.1 Student Perspectives on Non-Participation

Based on informal interviews with students during classroom practice at West School (Section 3.2.5.2), several students (25.8%) mentioned language barriers as a significant obstacle; five students (16.1%) reported that their non-English-speaking caregivers could not understand the study requirements or interact with the tool. Caregiver availability and resources also played a role, as three students (9.7%) cited busy caregivers who “worked both day and night shifts.” Notably, three West School students independently came forward to say that their caregiver was not available to sign contracts; two of them then asked if another family member could sign the contract and one later returned a contract signed by another family member.

At East School, students might have declined to participate because of the prospect of getting notified about their struggling with a skill and asking their caregiver for help (which the participating teacher communicated to students before the study). This mechanism was changed between Studies I and II so that students could always ask for caregiver help through the system, independent of struggle. The East School teacher phrased it the following way:

TeacherEast: *“[Other students] were a little turned off about their parent being notified when they got work wrong/what content they struggle with while doing the parent study. I told them not to fear that!”*

This perspective was also echoed by a East School student, sharing a preference for requesting help from their caregiver at any point in time, instead of that option being open only after the system detected that they were struggling:

Stu3East: *“I guess I was a little I mean, cause once you like you do it, cause you have to do it like a certain amount of times to get the option. So it’s kind of So at that point it was a little frustrating, I guess, because you’d have to like, you know, you’ve already gotten it wrong so many times.”* **Stu3East:** *“I wish it would have been like an option with like the hint thing”*

3.3.2.2 Teacher Perspectives on Non-Participation

Some caregivers perceived the signup process and probe as “extra work,” with the East School teacher remarking that broad caregiver engagement has been challenging for the district:

TeacherEast: *“I had a few students tell me that their parents didn’t want them doing the study due to it “being extra work”.” and “Parent involvement at a middle school level is quite difficult. Our district especially has this struggle as it has a wide range of socioeconomic status.”*

A similar insight emerged from one of the West School teachers, who highlighted that caregiver engagement for additional education-related activities is challenging for them, in part due to language barriers. Even though translated documents explaining the study procedure and goal contract translations were provided, they were likely insufficient for participation:

Teacher1West: *“[West School] is a unique place...I think doing it earlier would have gotten more people because students check out more after Christmas in 8th grade. So if they aren’t engaged, then the parents are much less likely to be engaged. The diversity we had brings issues with trust of unknowns and the translation doesn’t always get the message through in the way we want. And it being a core 8 special education and Multi-language learner class usually has more apathy towards school and extra work. Like I said, in my [geometry] class we would have probably had almost 100% involvement just for bonus points. But that isn’t the reality of school.”*

In summary, language barriers and caregiver availability impacted participation, particularly at West School, where students struggled to engage. At East School, while students were more engaged, some were initially hesitant about notifying caregivers when struggling with content. Teacher feedback from both schools highlighted that caregiver involvement was an ongoing challenge.

3.4 Discussion

Caregiver involvement is crucial for student success, yet factors like time constraints and limited content knowledge often hinder support [81]. Tutoring systems present opportunities to enhance caregiver homework support, but effective technology designs [45, 78]. We examined two caregiver engagement modes—a real-time intelligent support module and a goal-setting contract—in middle school math classrooms. Findings illustrate how both approaches relate to participation and engagement, offering insights for future designs.

3.4.1 RQ1: How Much Did Carergivers Engage With Each Probe?

3.4.1.1 Lowering Barriers to Active Caregiver Participation

Across schools, the proportion of caregivers who signed up for the tutoring system support was similar (12.9% at West vs. 11.4% at East School), but dropouts were notably higher at West School (50% vs. 20%). Observations and informal student comments suggest that some families found the formal signup and login process burdensome; dropout after enrollment could signify that some students were able to navigate the tutoring system well on their own (or because no opportune moment to help live through the system arose). In contrast, the caregiver homework contract provided less friction, as it was completed on paper and handed out to students, similar to regular homework assignments.

Future research could systematically explore technology access and competing family responsibilities in contexts where caregivers speak multiple languages or hold multiple jobs [78,81]. Automated and multimodal translation assistance may be explored to improve inclusivity in future tool design iterations [98].

3.4.1.2 Aligning Support Tools with Household Routines

Caregivers often engage in motivational support through encouragement and accountability, yet many lack confidence in tutoring [53, 78]. Accordingly, significantly more households engaged with the goal-setting contract (55%) compared to the tutoring system involvement tool (12%). Caregivers were more willing to engage in accountability mechanisms than real-time instructional guidance. Such accountability benefits students based on past research on contingency contracts in education, which have been shown to improve student persistence and practice [5,80]. Integrating such motivational support into ITS—potentially through digital contracts or automated progress reminders—could enhance student engagement without overburdening caregivers [21]. A limitation of our study is that we cannot gauge how much caregivers engaged with the homework contract beyond completing or signing off on it, which is subject to future research. Our log data analysis showed that students who completed the contract with

their caregiver or household member practiced significantly more out of class than those who did not. This associational analysis, based on 31 West School students, could be explained by a selection effect (whereby more motivated students are more likely to complete a practice contract with their caregiver) or by the previously documented effectiveness of goal setting in non-digital homework [5, 80]. Yet, from response rates in this study alone, it is evident that these contracts are likely to be adopted by a broader range of caregivers and households than directly involving caregivers into tutoring system instruction. Recent evidence offers quasi-experimental support for the effectiveness of practice contracts on learning and engagement in educational technology [18].

3.4.2 RQ2: Opportunities to Overcome Involvement Obstacles

3.4.2.1 Recognizing the Broader Household Support Network

Observations at West School revealed a need to design for peer and broad household support beyond the student-caregiver dyad. Three students at West School asked if family members other than their caregivers could complete the homework contract with them. These students requested to involve their older siblings. Bronfenbrenner’s Ecological Systems Theory emphasizes that multiple layers of environmental resources (e.g., families and schools) influence a student’s development [25]. The theory has been successfully used to understand student developmental outcomes [77]. Accordingly, tool design should integrate the student’s broader family context and other surrounding mesosystems, such as extended school and community, that connect the family microsystem to other student experiences.

3.4.2.2 Integrating Goal Setting Into Tutoring Systems

Our current goal-setting intervention presents a low-friction, low-resource tool for caregiver involvement in digital homework. It is possible, though subject to future research, that better integration of the goal-setting contract with the tutoring system could lead to better student learning. For instance, systems could integrate lightweight progress notifications in relationship to goals, which has shown promise in past research [7]. Further, goal-setting tools for tutoring systems [82] could be merged with data-driven insights for caregivers and students on goal completion based on log data. Such analytics have been previously successfully used in teacher-facing dashboards [95]. Integration with other household rituals and media for exercising accountability (e.g., through calendar applications or reminders) may further increase the ease with which caregivers can exercise accountable support.

3.4.3 Limitations

The design recommendations we offered assume that caregivers and students will have access to learning technologies, digital devices running them, and the internet, which is not always the case [58]. It is important to note that fundamental needs in caregiver-student relationships are prerequisites for effective relationships between educational institutions and caregivers, such as trust [86]. The technology designed here is not intended to displace these needs.

Our study was conducted in two schools with unique institutional contexts, demographics, and resources. Although these sites provided valuable contrasts in socioeconomic and linguistic

diversity, the findings may not be generalizable to all K–12 settings, given differences in technology adoption and school-home communication, including rural, charter, or private schools.

Moreover, while we captured student practice time based on log data and qualitative indicators of caregiver involvement, we did not measure actual academic improvements. Our focus was on feasibility and adoption rather than learning gains. Log data and our interviews are also limited in telling the story of how exactly students and caregivers engaged with goal setting and technology at home. Analyses of different caregiver use strategies, which could be studied through at-home observations or richer diary logs and experience sampling, could further enhance the design and effectiveness of the tools studied here.

3.5 Conclusion

We studied two design probes that allow caregivers to support their students in practicing with tutoring systems. We contrasted motivational and instructional aids that aim to overcome time, knowledge, and resource constraints. While caregiver access to a conversational support module in a tutoring system for live interactions saw low signup and usage rates (12% overall), significantly more households chose to engage with a goal-setting probe (55%).

These findings contribute novel evidence that tools enhancing motivational caregiver roles may be more widely adopted than tools focused on enhancing instructional support that caregivers provide to students. Even when caregivers lack content knowledge or have limited availability, goal-setting, and flexible accountability supports may foster meaningful increases in student practice. Observational data further highlight the broader “caring community” that surrounds the student—siblings, other relatives, and trusted adults—who may step in where a traditional parent–child homework partnership is difficult to establish. Future research may focus on improving the integration between goal-setting contracts and tutoring systems, which can track goal progress through log data.

Acknowledgments

This research was funded by the Institute of Education Sciences (IES) of the U.S. Department of Education (Award #R305A220386). We thank Sina Shahmoradi for contributions to early data analyses and literature review.

Chapter 4

Learning Benefits of Goal Setting with Rewards in Hybrid Tutoring

This chapter was adapted from my published conference paper:

Conrad Borchers, Alex Houk, Vincent Aleven, and Kenneth R. Koedinger. 2025. Engagement and Learning Benefits of Goal Setting with Rewards in Human-AI Tutoring. In *Proceedings of the 26th International Conference on Artificial Intelligence in Education (AIED '25)*. Palermo, Italy.

Summary Statement in Relationship to Thesis

In this chapter, I present a quasi-experimental study examining the effects of goal-setting contracts with rewards on student engagement and learning in a hybrid human-AI tutoring context. In hybrid tutoring classrooms, learners are supported by technology and human tutors through remote video conferencing software [92]. The intervention was implemented in collaboration with school partners and integrated into weekly classroom routines, requiring minimal teacher involvement. Results showed that students who participated in goal setting spent more time on math practice (about 25%) and demonstrated substantially higher skill mastery (about 40%). The findings suggest that simple, reward-linked goal-setting can be effectively embedded into hybrid tutoring programs to boost both engagement and learning (H3 of the thesis' theory of change; see Chapter 2). This work supports the broader thesis aim of developing scalable, intelligent goal-support tools that motivate sustained practice in AI-enhanced learning environments.

4.1 Introduction and Related Work

Goals are “*object or aim of an action, for example, to attain a specific standard of proficiency, usually within a specified time limit*” [69]. Setting goals as a motivational and performance-enhancing strategy has been extensively studied. Seminal work by Locke and Latham [70] 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 [57]. *Homework behavioral contracts*, or contingency contracts, are one common method for implementing goal setting in classroom and homework settings [80]. Cooper [30] 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 [57]. 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 [5].

Integrating goal setting with personalized learning remains underexplored [36, 88]. Goal-setting contracts are typically performed on physical paper, limiting integration with the feedback and data-driven capabilities of learning systems [82]. 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 [48, 76]. 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 [41, 57, 62]. 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 [29].

Although past AIED systems have supported learners in setting goals (e.g., [37]), these studies focused on process goals and improving metacognition *during* learning rather than performance goals typical for goal-setting contracts [80]. Additionally, prior research on SRL goal support has usually been limited to short-term instructional interventions (e.g., of a few hours [37]). 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 [74], we estimated the influence of our intervention on practice time. Further, to validate whether students learn more, or merely engage more [41, 62, 89], 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 [52]. 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.

4.2 Methods

4.2.1 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).

4.2.2 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 [14]. IXL is often used by teachers in the United States [73] 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.

4.2.3 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. 4.1). 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. [80]. 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.

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

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:

☒ 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).**

☒ 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]

☒ 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

Figure 4.1: Illustration of the goal-setting process with an example contract.

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.

To support student goal achievement, tutors could view whether a student had met their goal in the previous week via a dashboard (Fig. 4.2). 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 [91].

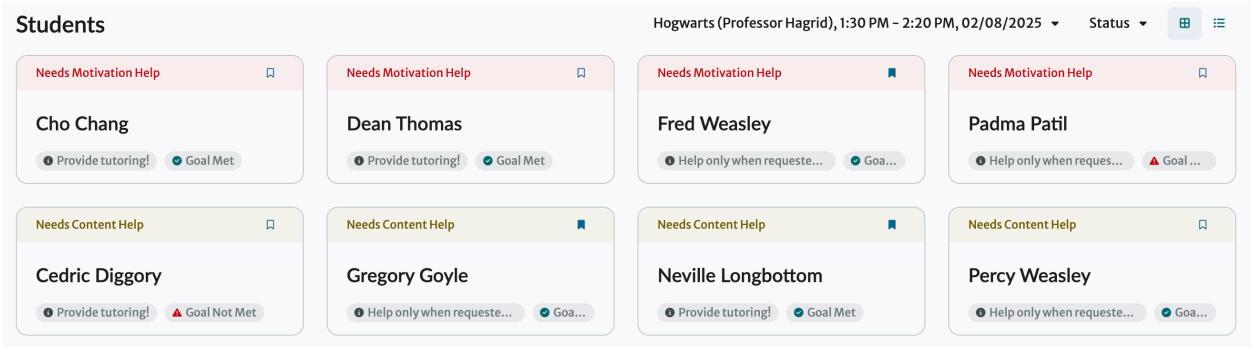


Figure 4.2: Example student data dashboard as seen by remote tutors.

4.2.4 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 4.2.5).

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).

4.2.5 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 [74]. 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. 4.3, 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.

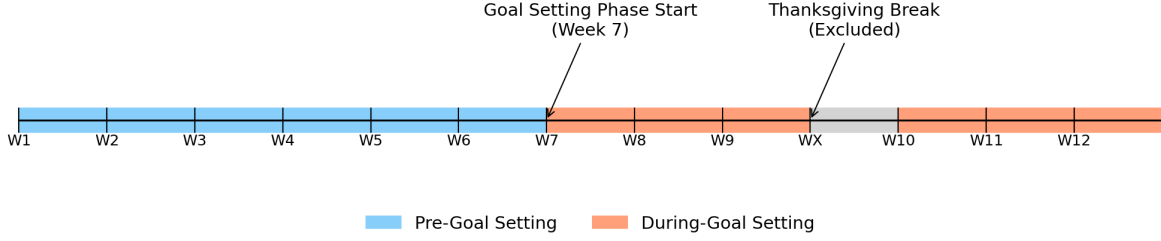


Figure 4.3: Schematic Study Timeline

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 [74]. 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 [17]. 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 4.2.4). The model of weekly student engagement (Y_{ij}) was specified as follows:

$$Y_{ij} = \beta_0 + \beta_1(\text{Week}_j) + \beta_2(\text{Week Goal}_j) + \beta_3(\text{Goal (Yes/No)}_j) + u_i + \varepsilon_{ij} \quad (4.1)$$

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 4.2.4) to examine patterns in skill acquisition before and during goal setting.

4.3 Results

4.3.1 Descriptive Differences in Time Spent

Fig. 4.4 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.

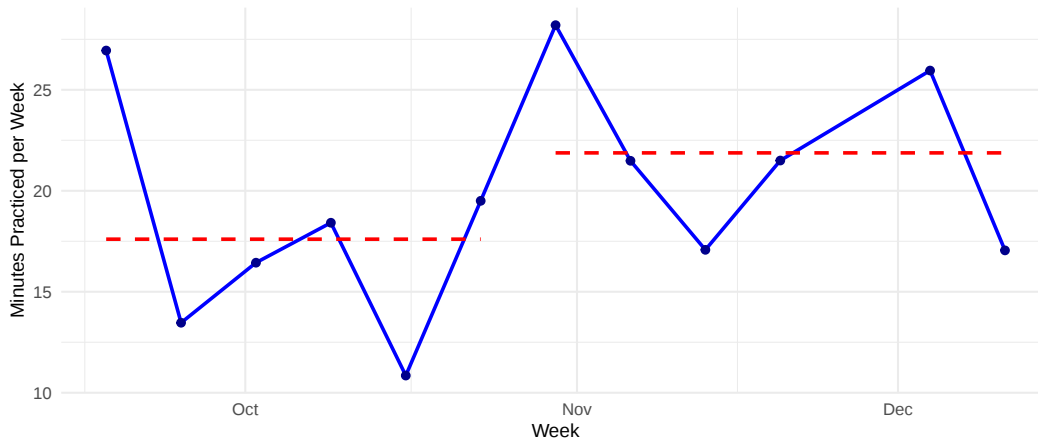


Figure 4.4: Average number of practice minutes across weeks with dashed reference lines showing the average outcomes pre-and-during goal setting.

Table 4.1: 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

4.3.2 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 4.1 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.

4.3.3 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 [10]. Therefore, we investigated skill mastery estimated from the IXL learning software before and during goal setting (RQ3). Table 4.2 summarizes the descriptive statistics for skills practiced, proficient, and mastered across two 6-week periods: pre-goal setting and during-goal setting.

Table 4.2: 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).

4.4 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.

4.4.1 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 [5, 57]. 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, 5]. 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 [57]. 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.

4.4.2 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 [97]). 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.

4.4.3 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 [10].

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 [82]. 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 [20].

4.4.4 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) simulta-

neously influence performance, rather than necessarily undermining each other, including in education [28]. 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 [97]. 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.

4.5 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).

Chapter 5

Differential Effects of Adaptive Goal Setting and Achievement

This chapter was adapted from my work-in-progress paper:

Conrad Borchers, Kenneth R. Koedinger, and Vincent Aleven. In Preparation. Student-Centered, Adaptive Goal-Setting Leads to Better Goal Achievement than Teacher-Assigned, Static Goals in Technology-Enhanced Learning. *Manuscript in Preparation*.

Summary Statement in Relationship to Thesis

This chapter presents evidence that adaptive, student-centered goal setting with data-driven calibration can improve goal achievement in hybrid tutoring environments (particularly for students with lower prior effort), replicating patterns of results seen in the previous chapter. Drawing on data from two schools in a randomized crossover design, we find that students in the Adaptive Skills condition achieved their weekly goals more frequently than in the static, teacher-assigned condition.

These benefits were most pronounced after mid-intervention goal adjustments based on students' personalized goal-achievement feedback, suggesting that personalized performance feedback fosters better calibration and sustained success. Momentum effects were stronger in the adaptive condition: students who met their goals in one week were substantially more likely to meet them again in the following week. Furthermore, intervention effects were significantly stronger for students who met their goals more frequently, after adjusting for prior effort. This underscores the dual advantage of helping students adjust their goals using data: practice and calibration benefits.

Both goal-setting conditions significantly increased time-on-task, but only adaptive, self-set goals significantly improved skill proficiency. While students with adaptive goals also demonstrated greater gains in proficiency and time on task than under static goals, these differences did not reach statistical significance. Notably, students with higher baseline practice levels (i.e., time on task) benefited less from either intervention. Still, as much as 84% of all students showed

positive practice-related improvement during goal setting. Goal difficulty was uncorrelated with intervention benefits.

Given that students with high prior effort tended to benefit less, differences in motivational profiles may play a critical role. Students with high prior effort may already possess stronger intrinsic motivation, whereas those who benefited most from goal setting may have begun with lower levels of intrinsic motivation. In such cases, extrinsic motivators (e.g., rewards) can serve as an entry point for fostering intrinsic motivation. For highly motivated students, more effective support may involve engaging with their goal hierarchy—clarifying both the reasons for pursuing a goal and the strategies for achieving it. Across all students, providing more frequent, data-driven feedback is expected to enhance the intervention’s impact. The final study will assess whether intrinsic motivation is maintained by the intervention through validated surveys and will measure goal hierarchies using established instruments such as goal orientation scales.

In relationship to this thesis’ theory of change (see Chapter 2), these findings support H1 and H2 regarding the benefits of student-centered, adaptive goals, and H3 by showing that goal achievement is a predictor of intervention benefit. The results inform the final chapter’s focus on enhancing calibration support, increasing feedback frequency, and integrating automated scaffolds at larger scale—while systematically measuring intrinsic motivation and goal hierarchies—to produce more explanatory accounts of intervention effectiveness and interindividual differences that can guide future interventions.

5.1 Introduction and Related Work

Active learning—central to AIED systems such as intelligent tutoring systems, teachable agents, and inquiry-based platforms—requires students to engage in problem-solving rather than passive instruction [63]. When instructional support is sufficient and not excessive, active learning yields benefits across all levels of prior knowledge [62]. However, its effectiveness hinges on students’ motivation to persist through effortful practice—a finding consistently replicated across demographic groups and real-world classrooms using cognitive modeling methods [41, 62, 89]. This raises a critical question: how can research-based interventions help reduce effort-related disparities so that all students can benefit from personalized learning technologies?

Goal setting, especially when paired with extrinsic rewards, has emerged as a promising strategy to promote student engagement in adaptive learning environments [5, 18]. However, an important open question concerns the heterogeneous effectiveness of such interventions. Effort is shaped by multiple factors—including students’ expectancy-value beliefs and motivational orientations [93]—and the impact of extrinsic motivators is known to depend on perceived autonomy, that is, voluntarily choosing to pursue self-set goal as opposed being compelled to pursue and externally-set goal [64, 65]. This suggests a need for adaptive, data-driven goal-setting supports that not only personalize learning but also scaffold and maintain student autonomy. However, such adaptive goals, thus far, have only been investigated in health behavior interventions, with some success [1]. Further, it is known that not all students are well-calibrated to set realistic goals for themselves [48]. Thus, we see a crucial gap (and opportunity) in the literature: leveraging educational technology log data to (a) guide student effort regulation via personalized goal recommendations while (b) preserving student autonomy, (c) potentially improving their goal calibration via data-driven feedback [84]. Notably, effort regulation outside of practice is a

dimension of adaptivity typically not considered in personalized learning systems, such as intelligent tutoring systems [4].

Studying the differential effects of adaptive goal-setting support—where learners are guided by system-generated recommendations to adjust their self-set goals—rather than passively accepting static, teacher-assigned goals, holds both theoretical and practical relevance. Theoretically, it contributes to a growing body of research on how autonomy-supportive environments can enhance motivation and self-regulated learning, particularly under conditions involving extrinsic incentives. It also advances our understanding of how learners’ agency in setting and calibrating goals can influence learning benefits of goal setting [5, 18]. Practically, this research informs the design of future, scalable interventions that provide personalized scaffolds for goal calibration, particularly for students who struggle to set appropriate goals on their own. More broadly, this work matters because it offers a promising pathway toward reducing differences in educational outcomes predicated on student effort by tailoring motivational supports to individual learners’ needs—helping all students benefit from the opportunities afforded by AI-enhanced learning environments.

To investigate the effectiveness of adaptive, data-driven goal support compared to static goals, we conducted a large-scale, randomized crossover experiment with middle school students participating in a hybrid tutoring program. Hybrid tutoring programs integrate both technology-based and human tutoring support [92]. The study compares different goal-setting conditions: teacher-assigned static goals versus student-selected goals with data-driven goal recommendations. Both conditions were accompanied by weekly, extrinsic rewards. The goal recommendations leverage log data from the learning system to help students adjust their goals in response to their progress, aiming to foster autonomy and reduce disparities in effort and achievement.

5.1.1 The Role of Student Autonomy Under Extrinsic Rewards

Self-determination theory (SDT) distinguishes between intrinsic motivation, which refers to doing something because it is inherently interesting or enjoyable, and extrinsic motivation, which refers to doing something because it leads to a separable outcome [35, 85]. While these forms of motivation are often viewed as opposing forces, meta-analytic evidence suggests that intrinsic and extrinsic motivators can work in tandem to influence performance, particularly in educational contexts [28]. Importantly, the benefit extrinsic rewards varies in the degree of autonomy it allows [28, 79]. For example, a student might complete homework to avoid punishment (low autonomy) or choose to do it in pursuit of a valued outcome, such as a desirable career or academic success (high autonomy) [85]. In this study, we use the term autonomy to refer to this degree of personal choice, consistent with the framework established by SDT [35].

Prior research indicates that autonomy supports goal-directed behavior and achievement when extrinsic rewards are present [28, 79]. These meta-analytic findings explain why some studies have found no loss of intrinsic motivation through rewards and praise [16], while others have argued the contrary [33]. Autonomy contributes to greater goal progress by enabling individuals to exert more effort, experience less internal conflict, and feel more prepared to change their behavior [64]. These benefits have also been confirmed in meta-analytic studies involving young adults and high school students [65]. However, not all students may perceive mathematics practice as personally meaningful. In classroom settings, goals may be viewed as externally imposed rather than self-endorsed. Such perceptions and beliefs are known predictors of effort and

achievement in mathematics [46]. This raises an important empirical question: does engaging students in actively setting their own goals—as opposed to having goals assigned—lead to better learning outcomes? Despite the theoretical importance of autonomy, empirical research on how autonomy affects daily goal pursuit in naturalistic educational settings remains limited [67].

5.1.2 Data-Driven Goal Support

While prior work has explored open learner models and dashboards using learner performance data to support self-regulated learning [20,26], few studies have examined how these tools can be leveraged to scaffold students’ goal-setting processes—particularly in the pre-actional and post-actional phases of learning [43]. One exception is Long and Alevan [71], who noted learning benefits of students who were able to choose practice problems in a math tutoring system based on their estimated knowledge proficiency. This form of goal setting, however, did not include tangible goal performance standards and rewards.

Indeed, most personalized learning systems support students *during* practice, focusing on monitoring progress or delivering help in response to errors [3,9], rather than helping students plan and adjust goals before and after practice. This leaves a critical gap in the development of students’ metacognitive regulation. Recent design work has begun to address this need by developing systems that not only track students’ performance but also generate personalized goal recommendations based on log data [20,22]. For instance, adaptive dashboards have been designed to display mastery progress, visualize effort, and provide dynamic goal suggestions that reflect a student’s prior achievement level. In a prototyping study with middle school students, Borchers et al. [22] found that students responded positively to such recommendations, especially when they retained final control over their goals. Students perceived the guidance as helpful for pacing and calibrating effort, aligning with broader findings in the behavioral sciences showing the effectiveness of adaptive goal-setting in other domains such as health [1].

Building on these findings, the current study evaluates the effects of such data-driven goal-setting supports in a longitudinal experiment with hundreds of students, testing whether adaptive recommendations informed by learning analytics can improve student effort, calibration, and ultimately goal achievement—beyond what static, teacher-assigned goals can offer. This approach positions data-driven goal setting as a promising addition to the broader toolkit of adaptive learning technologies [4].

5.1.3 The Present Study

The present study experimentally tests the effectiveness of personalized goal setting in authentic classroom practice with educational technology. Specifically, we conducted a randomized crossover experiment with over 184 middle school students across two public schools, examining the effects of adaptive, student-centered goal-setting on effort, goal achievement, and skill mastery during digital math practice. We address the following research questions:

- **RQ1a:** Does setting adaptive, student-centered weekly goals increase students’ **likelihood of achieving their goals** compared to static, teacher-assigned goals?
- **RQ1b:** Does receiving a reward for goal achievement in a given week **predict higher success** in the following week, and does this momentum effect differ by goal condition?

- **RQ2:** Does setting adaptive, student-centered weekly goals improve students' **effort and engagement** (as measured by skill proficiency and practice time) compared to static, teacher-assigned goals?
- **RQ3:** To what extent do **individual differences in prior effort and goal achievement** explain the benefits of goal setting?

The study contributes novel empirical insights into the effects of adaptive goal setting support, offering evidence about which students benefit most, under what conditions, and why. We also conduct exploratory analyses of prior effort, goal achievement, and goal difficulty to better understand whether and how personalized goal recommendations can mitigate differences in student engagement and achievement.

5.2 Methods

5.2.1 Sample and Study Context

We analyzed data from a hybrid tutoring program running for 11 weeks between January and April 2025. Students ($N=184$) were from two charter schools in the Mid-Atlantic United States, serving grades 5-9. All students enrolled at the schools were invited to participate in the tutoring program, and those who provided consent were included in the sample. The study followed an institutionally approved IRB protocol and school permissions were obtained. In both schools, nearly all students were African American and from low-income backgrounds. In School 1 ($N=101$), all students were male. In School 2 ($N=83$), genders were approximately equally represented.

As part of the hybrid tutoring program, online human tutoring is available to students during one (School 1) or two (School 2) class session per week, during which they engage in math practice with the IXL software. A researcher and the school's two math teachers (four total) facilitated the goal-setting activities in person. Each classroom session was about 43 minutes, though occasionally varied based on instructor need. Twelve classes were served. At School 1, this included two six grade classes and two fifth grade classes. At School 2, this included two 6th-grade and two 7th-grade classes, as well as three 8th and one 9th-grade classes.

All learners used IXL Math, an adaptive online platform for personalized math practice aligned with Common Core standards. Prior research shows that IXL use leads to significant learning gains compared to non-IXL schools over a three-year period in grades 3–8 [14]. Widely used in U.S. classrooms [73], IXL adapts problem difficulty through real-time analytics and mastery-based algorithms. Students receive immediate feedback and motivational messages after correct answers, with explanations or walkthroughs provided for errors. The platform also offers occasional worked examples and tracks skill progress, enabling teachers to monitor performance and support learning. Achievements and milestones help maintain student engagement.

5.2.2 Experimental Design

At the outset, all students completed a three-week baseline period during which no experimental manipulation occurred. This baseline period served as a quasi-experimental contrast to students

achievement levels during the two goal-setting interventions. Students were then randomly assigned to one of two experimental groups. In the first four weeks of the intervention, Group 1 received *adaptive, student-set goals* supported by algorithmically generated recommendations, while Group 2 received *static, teacher-assigned goals*. In both conditions, students received weekly extrinsic rewards (i.e., fruit snacks; in discussion with participating teachers and school administrators) for meeting their goals, allowing us to isolate the effects of goal-setting adaptivity and student autonomy.

A condition crossover occurred after Week 4. Group 1 transitioned to static goals, and Group 2 began receiving adaptive goals with data-informed recommendations. In both conditions, students received a single goal adjustment opportunity—at Week 3 for Group 1 and Week 7 for Group 2—allowing us to examine the benefits of mid-intervention calibration and its interaction with goal-setting autonomy.

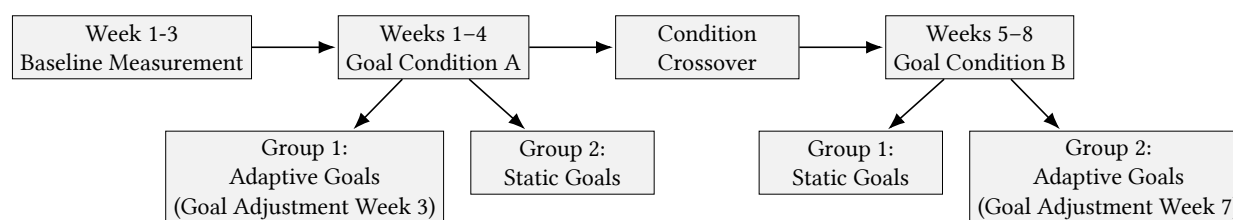


Figure 5.1: Study design schematic. Students were assigned to either adaptive or static goal-setting conditions, with a crossover after four weeks. All students received midpoint goal calibration in the adaptive goals condition.

5.2.3 Procedures and Intervention

All students participated in a hybrid tutoring model integrated into their regular math class (see Section 5.2.1). During these sessions, students logged into the IXL adaptive math platform and the Pencil video conferencing platform on their assigned Chromebooks. Upon login, they were greeted by a remote tutor and placed into individual breakout rooms. Tutors observed students’ IXL practice via screen sharing, provided content support as needed, and offered motivational encouragement. Tutoring groups were staffed by consistent tutor teams introduced to students at the start of the academic year to promote familiarity and rapport. Classroom teachers and research assistants supported students on-site by resolving technical issues and helping in distributing goal setting materials as well as goal contracts and reports.

During the eight-week intervention phase, students were assigned to goal-setting conditions (adaptive or static; see Figure 5.1). In the first week of each condition, students completed a paper-based math practice contract, distributed by a research assistant. Adapted from recommendations by Peacock et al. [80], the contract prompted students to commit to weekly practice goals and offered clear reward contingencies. Students could accept a teacher-assigned goal for practice time or a self-set goal for skills mastered per week, depending on their condition. As mentioned, in consultation with teachers. All students who met their weekly goals received a fruit snack the following week at the beginning of the period, with two rewards per week for sustained streaks of goal achievement (of at least three weeks). Rewards were typically distributed by research assistants on site, and occasionally by teachers who had access to a shared weekly goal report.

Static versus Adaptive Goals. Students in the static condition were assigned a teacher-set goal, typically defined as staying logged into IXL and Pencil for the duration of the class period and practicing (typically 30–40 minutes per period depending on the class). In contrast, students in the adaptive condition selected their own skill-based goals (e.g., mastering one or more grade-level skills per week). Contracts in the adaptive goal condition mentioned an initial recommendation of one skill per week to anchor student choice, with one skill per week being based on past school averages in the program. To support goal calibration, these students also received data-informed goal recommendations generated from their past practice data at the midpoint of the intervention segment after two weeks. These reports showed their average skills mastered and percentage of goal completion in previous weeks. The algorithm followed a tiered logic: students who underachieved their goal by a margin of 33% or more received encouragement to reduce their goal or reflect on challenges; students who met their goal were praised and encouraged to sustain it; and students who overachieved by more than 33% on average were invited to increase their goal. Goal recommendations were based on the midway point between students' past goal and their historical performance (e.g., 3 if students chose 2 but achieved 4 skills per week) and rounded to whole numbers. This rather simple algorithm was based on the fact that established recommendation algorithms typically require data from nine measurement points or more [1], which our design did not allow for. Yet, it followed a similar logic: take a goal close to students historical average (and round it up as needed to encourage growth). Further, the explicit contrast between students initial goal and actual performance was designed to improve their goal calibration [48]. Students in the static goal condition did not receive a midway report to isolate the effect of goal achievement feedback.

An illustrative excerpt of the paper-based goal setting and recommendation process is shown in Figure 5.2. Students were always given the option to accept, reject, or modify the recommended goal, thus maintaining autonomy.

Goal Monitoring and Feedback. During the intervention, tutors used a centralized dashboard (Figure 5.3) to monitor all students' progress toward their weekly goals. Tutors were trained to reference these goals during sessions and encourage students as-needed (Figure 5.3). Tutors reinforced goal achievement with verbal praise and brief motivational interactions. For students who failed to meet their goals, tutors provided scaffolded guidance based on students' self-reported reasons, drawing on strategies from prior research on responsive hybrid tutoring [91]. This support, however, is not central to the present study's focus, though it ensured that both types of goals had minimal perceived differences in the classroom to ensure a fair comparison.

5.2.4 Data Analysis Methods

Weekly goal achievement and IXL practice data were logged in a centralized spreadsheet maintained by the tutor program and research team. For each student, we recorded goal type, number of skills practiced and mastered, minutes spent in IXL, and whether the weekly goal was achieved. These logs were based on with IXL-generated performance reports and aggregated using custom Python scripts.

We used the IXL "skill proficient" flag to count the number of skills completed each week. This lenient threshold (typically marked at 80% proficiency in IXL) was chosen based on observations and teacher input during Week 1 of the intervention. Many students treated this level as the point

You have been knocking your goal out of the park! Do you want to ramp up your goal? This feedback will help you become better at setting and achieving your goals. *Please fill the parts in blue below.*

Past goal: 20 minutes of IXL practice per week

Your achievement (11/13–11/26): 30 minutes per week (150% of goal)

Updated goal recommendation: 25 minutes of IXL practice per week

I, the student, respond:

☐ YES, I AGREE to my updated goal

☒ NO, I DO NOT AGREE to my updated goal. Instead, I will practice 28 minutes per week

☐ NO, I DO NOT AGREE to my updated goal. Instead, I will keep my old goal

Figure 5.2: Illustration of student interaction with an updated goal recommendation based on historical performance averages. Students retained full control over whether to adopt, reject, or modify the suggestion.

at which they considered a skill “done.” To avoid undermining motivation and ensure alignment with classroom interpretations of progress, we adopted this operationalization for all skill-based outcome variables.

To examine the effects of adaptive versus static goal setting on student behavior and performance, we conducted a series of mixed-effects models aligned with each research question. We used both logistic and linear mixed-effects modeling approaches, accounting for the nested structure of the data (weekly observations nested within students). All models were implemented in R using the lme4 package [15], and model assumptions (e.g., homoscedasticity) were checked using standard diagnostic procedures.

RQ1a: Do adaptive goals improve goal achievement? To estimate whether adaptive, student-set goals increased students’ likelihood of achieving weekly goals, we modeled the binary outcome of reward receipt using a mixed-effects logistic regression. Fixed effects included goal condition (skills vs. focus), segment (pre- vs. post-adjustment), and their interaction. A random intercept for each student was included to account for repeated measures. The interaction term allowed us to test whether adaptive goals became more effective over time, particularly after the midpoint goal calibration.

RQ1b: Do past rewards predict future success (momentum effects)? To test whether goal achievement in one week predicted higher rates of success in the subsequent week, we grouped the data by goal condition and prior-week reward status, then calculated the proportion of students who achieved their goal in the following week. We then conducted a χ^2 -test on the contingency table of goal achievement counts by prior reward status and goal condition to assess

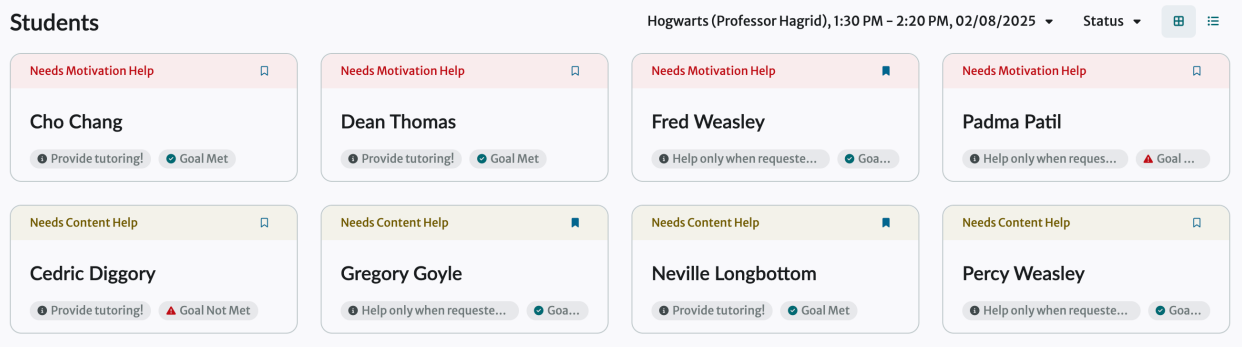


Figure 5.3: Example student data dashboard as seen by remote tutors. The dashboard enabled tutors to monitor weekly goal completion and tailor support accordingly.

whether there associations between the type of goal (self-set vs. assigned) and the rate of continued success. For each group, for descriptive plots, we also computed binomial 95% confidence intervals using the Agresti-Coull method.

RQ2: Do adaptive goals improve effort and learning (skill proficiency)? To assess whether goal condition influenced student engagement and productivity, we employed mixed-effects interrupted time-series models [74]. The unit of analysis was the student-week, with repeated measures for each participant across a 3-week baseline period and an 8-week intervention phase.

Model specification. For each outcome, we included: (a) a fixed effect for the overall time trend across the study (*Week Number*), (b) binary indicators for whether the *Adaptive Skills Goal* or *Adaptive Focus Goal* intervention was active in a given week, and (c) post-intervention time variables counting weeks since the start of each goal condition (*Weeks Since Skills Goal Started*, *Weeks Since Focus Goal Started*).

This structure allowed estimation of the baseline slope before any intervention, the immediate level effect of each goal condition when active, and the change in slope for each goal condition relative to baseline (i.e., interrupted time series trends). All time variables were centered such that Week 1 of the 8-week intervention was coded as $t = 1$, and baseline weeks were coded as $t = -2, -1, 0$. The general interrupted time-series model for outcome Y_{ij} (student i in week j) was:

$$\begin{aligned}
 Y_{ij} = & \beta_0 + \beta_1(\text{Week Number}_j) \\
 & + \beta_2(\text{Adaptive Skills Goal Active}_j) \\
 & + \beta_3(\text{Adaptive Focus Goal Active}_j) \\
 & + \beta_4(\text{Weeks Since Skills Goal Started}_j) \\
 & + \beta_5(\text{Weeks Since Focus Goal Started}_j) \\
 & + u_{0i}^{(\text{student})} + u_{0i}^{(\text{teacher})} + u_{0i}^{(\text{school})} + u_{0i}^{(\text{grade})} \\
 & + u_{1i}^{(\text{student})}(\text{Week Number}_j) + \varepsilon_{ij},
 \end{aligned} \tag{5.1}$$

where β_0 is the baseline intercept, β_1 is the overall time trend, β_2 – β_3 are level differences when each goal condition is active, and β_4 – β_5 are slope changes when each goal condition is active. Random intercepts $u_{0i}^{(\text{student})}$, $u_{0i}^{(\text{teacher})}$, $u_{0i}^{(\text{school})}$, and $u_{0i}^{(\text{grade})}$ capture clustering at the student, teacher, school, and grade levels, respectively. A random slope $u_{1i}^{(\text{student})}$ for *Week Number* accounts for within-student variation in time trends, following recommendations to assume the highest possible complexity in random effects structure in hierarchical linear models [12]. Residual error is represented by ϵ_{ij} .

Effort outcome. Weekly skill proficiency (count of skills marked proficient) was modeled using a mixed-effects Poisson regression:

$$\begin{aligned}
 \text{Skills Proficient} \sim & \text{Week Number} + \text{Adaptive Skills Goal Active} \\
 & + \text{Adaptive Focus Goal Active} \\
 & + \text{Weeks Since Skills Goal Started} \\
 & + \text{Weeks Since Focus Goal Started} \\
 & + (1 + \text{Week Number} \mid \text{Student}) \\
 & + (1 \mid \text{School}) + (1 \mid \text{Teacher}) + (1 \mid \text{Grade}).
 \end{aligned} \tag{5.2}$$

Note on adjustment period. We did not estimate a separate double-interrupted time-series for the midpoint goal adjustment, as adding an additional intervention indicator and time trend rendered the model non-identifiable for our sample size. Therefore, we did not focus on this period in the primary analysis. However, we note that interrupted time trends observed during the adaptive skill intervention implicitly capture changes due to goal adjustments, which happen during the intervention.

Secondary outcome. To assess time-on-task (minutes practiced), we fit a linear mixed-effects model on the log-transformed outcome using the $\log(1+x)$ transformation to accommodate zero values. This preserves all observations (including weeks with no practice) and yields multiplicative effects on the original scale.

$$\begin{aligned}
 \log(1 + \text{Minutes Practiced}) \sim & \text{Week Number} + \text{Adaptive Skills Goal Active} \\
 & + \text{Adaptive Focus Goal Active} \\
 & + \text{Weeks Since Skills Goal Started} \\
 & + \text{Weeks Since Focus Goal Started} \\
 & + (1 + \text{Week Number} \mid \text{Student}) \\
 & + (1 \mid \text{School}) + (1 \mid \text{Teacher}) + (1 \mid \text{Grade}).
 \end{aligned} \tag{5.3}$$

All models were fit in lme4 in R [15]. Parameter estimates were reported with 95% confidence intervals. Fixed-effect significance was assessed using Wald z -tests (generalized models) or t -tests (linear models).

RQ3: Do individual differences in prior effort and achievement explain differences in goal-setting benefits? To examine whether prior effort and achievement explained student-level differences in the benefits of the goal-setting interventions, we first derived a student-level estimate of the two intervention effects (adaptive and static goals) from the interrupted time-series models described above. For each outcome—weekly skill proficiency and weekly minutes practiced—we refit the models with added random slopes for the goal condition indicators at the student level. This specification allowed the magnitude of the goal-setting effects to vary across students, rather than assuming a constant effect for all participants.

From each model, we extracted the student-specific random slopes for the adaptive and static skills goal condition and added it to the corresponding fixed effects to obtain the total goal-setting effect for that student and model. For the skill proficiency model (Poisson), these estimates were expressed as log-rate ratios (i.e., incidence rate ratios).

To form a single composite metric of the goal-setting effect per student across outcomes, we combined the skill proficiency and practice time effects by first standardizing each to a z-score and then averaging them. This approach placed both outcome measures on a common scale while giving them equal weight in the composite. The resulting composite score thus reflects each student's overall responsiveness to goal setting (be it adaptive or static) across both productivity and effort outcomes. The rationale behind averaging multiple latent intervention effects per student across outcomes is that more measurements are expected to yield a more reliable estimate. With that said, we also separately explored an average of student-level goal intervention effects of adaptive and static goals, to observe any difference by goal type.

To answer RQ3, we then used this composite as the dependent variable in separate analyses, entering prior effort (mean weekly minutes practiced during the baseline period before any goal setting occurred) and goal achievement (% of goals completed during the interventions) as predictors, including their interaction. In an exploratory analysis, we separately fit two additional models predicting the student-level adaptive or static goal effect, respectively. All moderation models were fit using standard multiple linear regression with student as the unit of analysis.

Throughout all analyses model assumptions were checked through residual diagnostics. These diagnostics included Q–Q plots to assess residual normality, residuals-versus-fitted plots to check homoscedasticity and linearity, scale–location plots to detect heteroscedasticity, and residuals-versus-leverage plots (with Cook's distance) to identify influential observations.

5.3 Results

5.3.1 RQ1a: Do Adaptive, Self-Set Goals Improve Goal Achievement over Static, Teacher-Set Goals?

Across both study phases, students achieved a higher percentage of goals when they set them themselves with *Adaptive Skills* guidance (29–36%) than when following *Static Minutes* goals set by teachers (20–21%). In addition, students in the *Adaptive Skills* condition showed an increase in goal achievement from Phase 1 to Phase 2 (29% to 36%), consistent with the provision of a goal adjustment opportunity based on performance data (see Table 5.1). Answering RQ1, these patterns indicate that both the source of the goal (self-set vs. teacher-set) and the presence of data-driven goal adjustment opportunities contributed to higher rates of goal achievement.

Table 5.1: Goal achievement rates (%) by goal-setting approach across study phases. Students were more likely to reach goals when setting them themselves with data feedback (*Adaptive Skills*) compared to teacher-set goals (*Static Minutes*).

Goal Type	Phase 1	Phase 2
Static Minutes	19.7% [15.6%, 23.8%]	20.8% [16.6%, 25.0%]
Adaptive Skills	29.4% [24.6%, 34.1%]	36.4% [31.4%, 41.4%]

Based on the mixed-effects logistic regression, these differences were statistically significant. Students in the *Adaptive Skills* condition had higher odds of achieving their weekly goals than those in the *Static Minutes* condition ($OR = 1.85, p < .001$). In addition, a significant interaction between goal condition and phase ($OR = 1.36, p < .001$) indicated that the advantage of adaptive, student-set goals increased from pre- to post-adjustment, aligning with the introduction of data-driven goal calibration.

As a manipulation check, we confirmed that these improvements in goal achievement coincided with students, on average, adjusting their goal upward when overachieving, and adjusting their goal downward when underachieving (see Figure 5.4). Students who were not given a recommendation retained their goal, with slight variation introduced by absentees who completed their contract late, that is, only in the second half of the intervention.

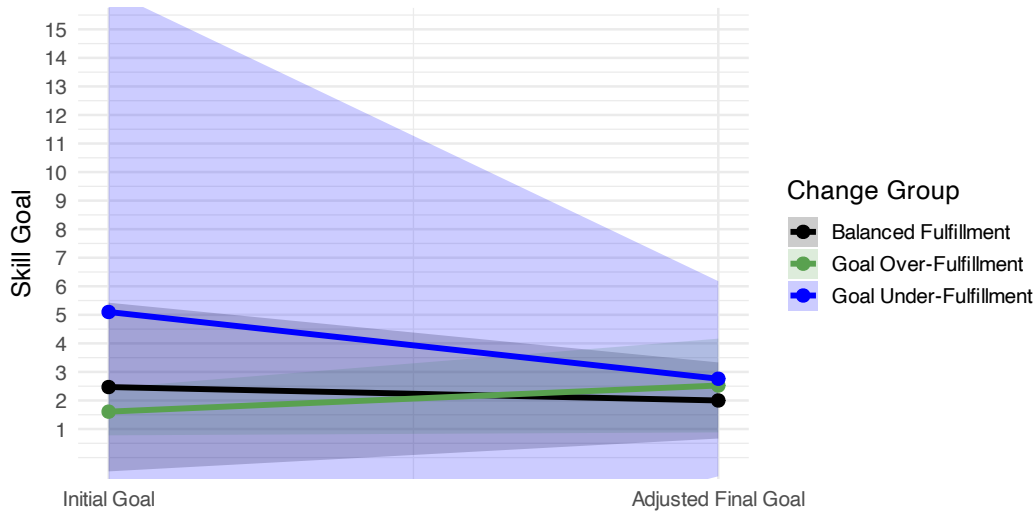


Figure 5.4: Students' self-set goals in the adaptive goals condition before and after data-driven feedback based on type of recommendation, which was linked to performance.

5.3.2 RQ1b: Do Adaptive, Self-Set Goals Enhance Goal Achievement Momentum?

We examined whether achieving a goal in one week predicted higher rates of success in the subsequent week, testing for a momentum effect in goal achievement. Descriptively, Figure 5.5 shows that students who received a reward in the previous week were more likely to meet their goal in the following week, consistent with a momentum or reinforcement effect. This pattern

held for both goal types, but the increase was more pronounced for students in the *skills goal* condition. Students with *Adaptive Skills* goals who earned a reward in the prior week achieved their goal the following week 57% of the time, compared to 26% after not earning a reward—an increase of nearly 30 percentage points (an increase of factor 2.23). For *Static Minutes* goals, the increase was smaller, rising from 18% to 36% (about 18 points; an increase of factor 2.05).

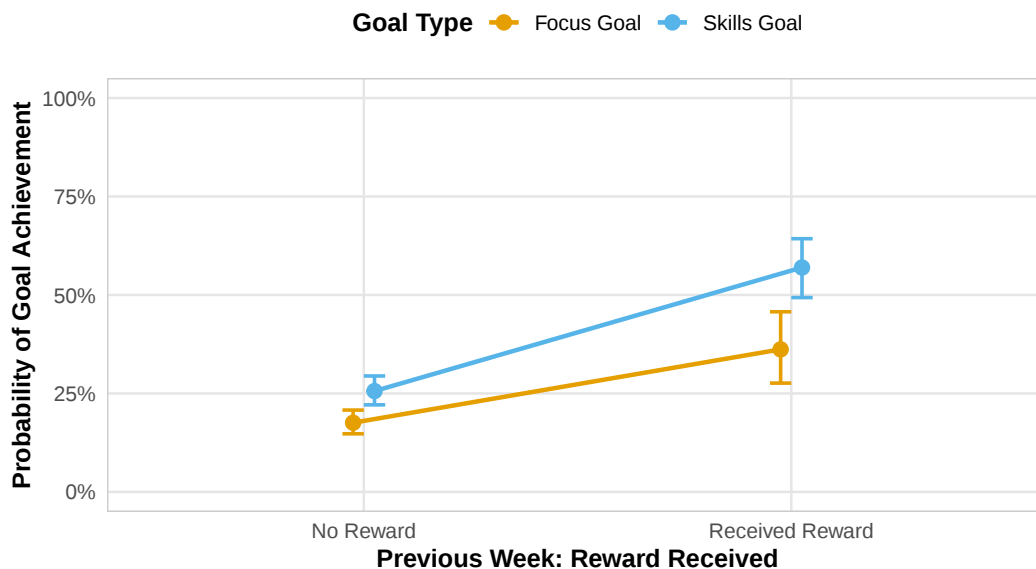


Figure 5.5: Probability of meeting goal again after meeting goal or not in the previous week (momentum) across conditions.

This condition difference was statistically reliable. Specifically, the null hypothesis that momentum (i.e., previous week’s achievement) is independent of next-week achievement is independent of condition was rejected, $\chi^2(1, N = 379) = 7.49, p = .006$.

5.3.3 RQ2: Do Adaptive, Self-Set Goals Improve Effort and Engagement?

Reporting descriptive differences first, the *No Goal* baseline condition showed the lowest practice time ($M = 23.6$ min) and proficiency ($M = 1.03$ skills), with an average of 22.8 minutes required per skill. Relative to baseline, the *Static Minutes* condition was associated with a 15.7% increase in practice time and a 34.0% increase in skills gained, while the *Adaptive Skills* condition showed a 21.2% increase in practice time and a 41.7% increase in skills gained. Both goal-setting conditions also showed greater efficiency (fewer minutes per skill). The *Adaptive Skills* condition had the highest proficiency ($M = 1.46$) and the lowest minutes per skill (19.6), slightly outperforming the *Static Minutes* condition ($M = 1.38$, 19.9 min/skill; see Table 5.2). Next, we turn to whether these differences are significant based on interrupted time-series models.

Table 5.2: Descriptive statistics for practice time, proficiency, and efficiency by goal condition. Values are presented as $M \pm SD$.

Goal Condition	Practice (min)	Proficient	Min./Skill
No Goal	23.6 $\pm$ 30.1	1.03 $\pm$ 2.09	22.8
Static Minutes	27.3 $\pm$ 23.5	1.38 $\pm$ 2.10	19.9
Adaptive Skills	28.6 $\pm$ 24.7	1.46 $\pm$ 2.18	19.6

Skills proficient per week. Using the mixed-effects interrupted time-series model specified above, students demonstrated a positive overall week-to-week trend in skills proficient (Incidence Rate Ratio, IRR = 1.05, 95% CI [1.02, 1.08], $p < .001$), corresponding to a 5% weekly increase. Weeks under the self-set *skill* goal condition were associated with a 24% higher proficiency rate relative to non-intervention weeks in terms of skills proficient per week (IRR = 1.24, 95% CI [1.03, 1.50], $p = .023$). In contrast, the teacher-set *minutes* (focus) goal condition showed no reliable difference in proficiency (IRR = 1.08, 95% CI [0.90, 1.31], $p = .416$). Post-start slopes for both goal conditions were not statistically different from zero (adaptive goals: IRR = 0.96, $p = .141$; static goals: IRR = 0.98, $p = .505$), indicating no changes in trend. Random effects showed that most variance in skill proficiency per week was attributable to differences between students (ICC = 0.758), with smaller contributions from teachers (ICC = 0.053) and schools (ICC = 0.018). Grade-level differences were negligible beyond these factors (ICC $\approx$ 0).

Minutes practiced per week. In the linear mixed-effects model on $\log(1 + \text{minutes})$, students showed an overall significant week-to-week increase (estimate = 0.05 on the log scale, 95% CI [0.02, 0.09], $p = .003$). Weeks under the self-set adaptive *skill* goal condition were associated with a significant increase in minutes practiced (estimate = 0.52, 95% CI [0.26, 0.79], $p < .001$). Weeks under the *minutes* (focus) goal were also associated with a significant increase (estimate = 0.37, 95% CI [0.10, 0.63], $p = .007$). Post-start slopes for skills and focus were not statistically different from zero (adaptive goals: -0.07 , $p = .143$; static goals: -0.04 , $p = .358$). Variance partitioning indicated modest clustering at the student level (ICC = .109), with smaller components for school (ICC = .054) and teacher (ICC = .042); grade-level variance was too small (i.e., near-zero) to be reliably estimated.

Direct comparison of intervention coefficients. Although adaptive, self-set *skill* goals showed larger point estimates than teacher-set *minutes* (focus) goals across models, the difference between a significant and a non-significant coefficient is not, by itself, necessarily statistically significant. We therefore tested the null hypothesis that the two intervention coefficients are equal ($\beta_{\text{skills}} - \beta_{\text{focus}} = 0$) using planned Wald contrasts with single-step multiplicity adjustment.

In the skills-per-week model (Poisson), the skills-vs-focus contrast was 0.138 (SE = 0.110), $z = 1.257$, $p = .209$; on the incidence-rate scale this corresponds to a ratio of IRRs of $e^{0.138} = 1.15$ with a 95% CI [0.93, 1.42], which trended toward a adaptive goal benefit but yielded no significant difference between the two goal conditions in their effect on proficiency. In the minutes-per-week model (linear), the contrast was -0.19 minutes (SE = 2.93), $z = -0.065$, $p = .948$, with a 95% CI $[-5.94, 5.56]$ minutes, likewise showing no significant difference between the intervention effects. Taken together, while adaptive *skill* goals significantly increased proficiency and

both goal conditions increased time-on-task, we did not find evidence that the adaptive *skill* goal effect was statistically larger than the *minutes* (focus) goal effect.

5.3.4 RQ3: Do individual differences in prior effort and achievement explain differences goal-setting benefits?

Descriptively, Figure 5.6 summarizes the student-level composite goal effect across all outcomes (skills vs. minutes practice) and goal types (static vs. adaptive). The top panel shows the regression coefficients (standardized outcome) with 95% CIs: higher *Prior effort* (z) predicts smaller benefits after adjusting for goal achievement (negative estimate, CI excluding zero), whereas *Goal achievement* (z) is positively associated with the composite effect after adjustment for prior effort. The middle panel provides descriptive scatterplots with linear *univariate* fits showing a flat univariate association for goal achievement but a pronounced negative association for prior effort. The bottom panel displays the distribution of the *raw* composite intervention effect, with a dashed line at zero; most students out of the 182 students for which student-level intervention effects could be computed (84%) exhibit positive goal setting intervention effects, while 30 of 182 students (16%) fall below zero, indicating a negative composite response to goal setting with rewards.

Across all models ($N = 182$ students for which intervention effects were estimable), *prior effort during baseline* was a strong, negative predictor of subsequent goal-setting benefits: students who practiced more minutes before any goal-setting showed smaller estimated intervention gains (combined composite: $b = -0.77$, 95% CI $[-0.87, -0.68]$, $p < .001$; static-goal effect: $b = -0.75$, 95% CI $[-0.85, -0.64]$, $p < .001$; adaptive-goal effect: $b = -0.79$, 95% CI $[-0.89, -0.69]$, $p < .001$).

Goal achievement during the intervention was positively associated with the composite measure of student-level intervention benefit, but its strength varied by goal type. For the combined composite outcome, higher goal achievement predicted larger benefits ($b = 0.14$, 95% CI $[0.04, 0.24]$, $p = .007$). Disaggregating by goal type, the association was significant for the *static (minutes/focus)* goal effect ($b = 0.19$, 95% CI $[0.08, 0.29]$, $p < .001$) and positive but not statistically reliable for the *adaptive (skills)* goal effect ($b = 0.09$, 95% CI $[-0.01, 0.18]$, $p = .074$).

Across all models, there was no significant interaction between prior effort and goal achievement when predicting the student-level intervention effect (p 's $> .416$). Another potential confounder for the intervention effect could be goal difficulty, as highly challenging goals are sometimes known to be very motivating [69]. However, Pearson's product-moment correlation between goal difficulty (expressed in the average, student-level number of skills per week set as their target) and the composite student-level intervention effect was non-significant, $r = -0.06$, 95% CI $[-0.21, 0.09]$, $p = .458$, indicating no evidence that goal difficulty was associated with overall effectiveness of goal setting with rewards. Model fit was moderate to strong: $R^2_{\text{adj}} = .568$ for the combined composite, $.527$ for the static-goal effect, and $.594$ for the adaptive-goal effect.

Taken together, students who more consistently *met their weekly goals* derived greater overall benefit from goal setting—especially for teacher-set minutes targets—while students with *higher baseline practice effort* tended to show smaller incremental gains under both adaptive (skills) and static (minutes) goal conditions.

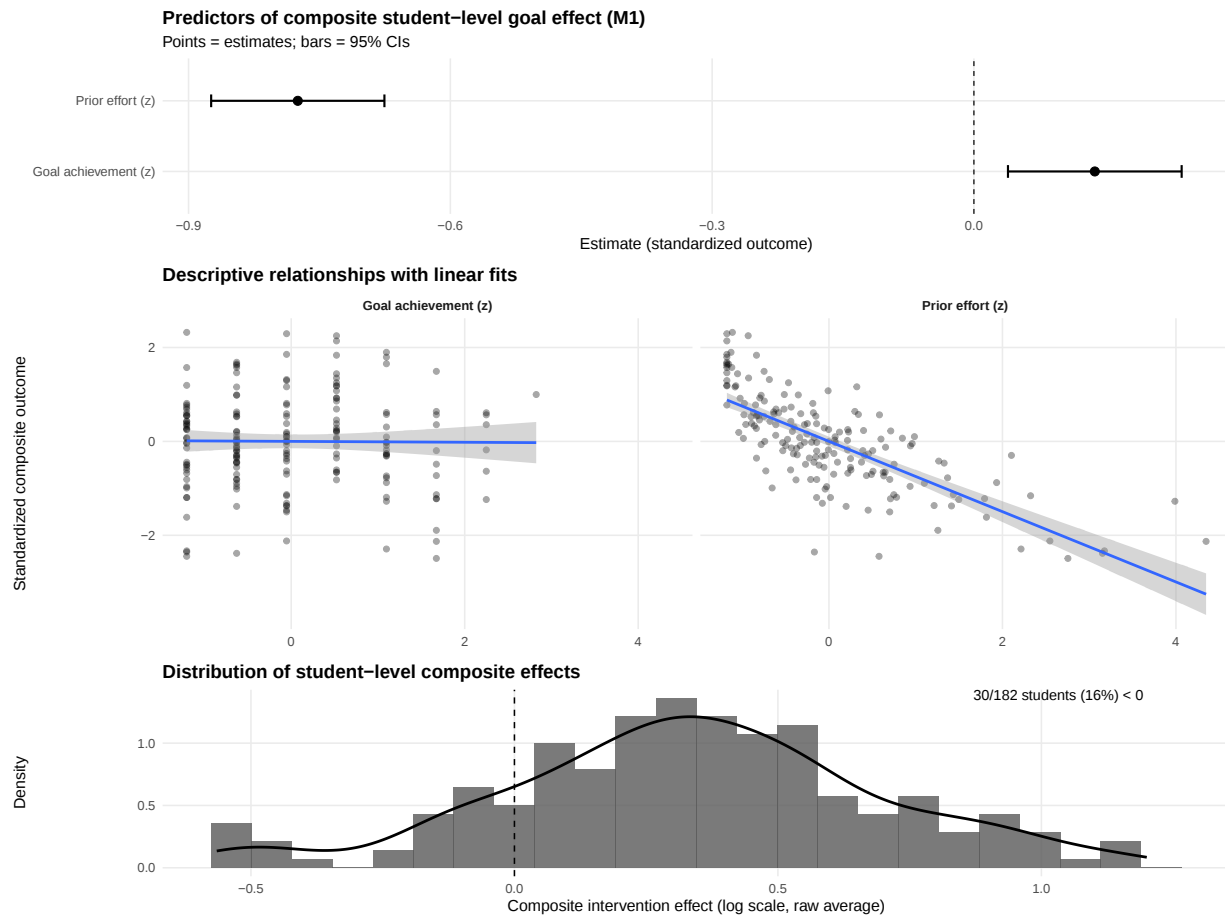


Figure 5.6: Model summary figure predicting student-level intervention benefits averaged and standardized across time and skill outcomes as well as goal types (static vs. adaptive): (left) coefficient estimates with 95% CIs; (middle) descriptive scatterplots with linear fits for standardized predictors; (right) histogram of the *raw* composite intervention effect (dashed line at 0; panel annotation reports the share < 0).

5.4 Discussion

This study provides further evidence that adaptive, student-centered goal setting with data-driven calibration can improve goal achievement and learning outcomes in hybrid tutoring environments.

5.4.1 Benefits of Goal Adjustment and Selection for Goal Achievement

Extending the findings from Chapter 4 to a second site with a larger number of classrooms, we found that self-set goals supported by performance-based recommendations produced higher rates of weekly goal completion than static, teacher-assigned goals. This advantage became more pronounced following the midpoint calibration, accentuating the potential of timely, data-driven feedback to help students align their goals with actual performance. Students tended to adjust their goals downward after underachievement and upward after overachievement, consistent with our recommendations and with meta-analytic evidence on post-success and post-failure goal adjustment [90]. In that meta-analysis, self-efficacy emerged as a key mediator of upward adjustment after success, suggesting that assessing self-efficacy in the final study may provide further explanatory insight.

The momentum analysis further revealed that meeting a goal in one week more than doubled the likelihood of meeting it again the following week, with a stronger effect under adaptive goals. This pattern aligns with theories of positive reinforcement, in which goal achievement and self-efficacy can mutually reinforce one another to sustain or improve performance [97]. The finding that students who more frequently met their goals also derived larger benefits from the intervention suggests that earlier and more frequent opportunities for goal adjustment and selection could strengthen these effects.

Finally, based on these findings, providing more frequent, data-driven feedback—beyond a single mid-intervention calibration—is likely to amplify benefits for all students. The final study in this thesis will enable refinement of automated, frequent, and adaptive goal-setting models to maximize both effectiveness and student autonomy, which is known to moderate the relationship between extrinsic rewards and performance for intrinsically motivated students [28, 79], as discussed next.

5.4.2 Low-Effort Students Benefited More from Goal Support

Across both conditions, goal setting was associated with increased practice time; however, only adaptive goals produced a statistically significant improvement in skill proficiency. While the adaptive condition showed descriptively larger gains than the static condition, the difference between interventions was not statistically significant, suggesting that a larger sample may be required to reliably detect such differences.

Moderation analyses revealed a robust pattern: students with higher prior effort during the baseline period derived significantly smaller benefits from either intervention. In contrast, students with lower baseline effort showed the largest relative gains. Overall, 84% of students demonstrated some improvement in practice-related outcomes during goal setting, whereas the remaining 16%—those typically already exerting high effort without formal goals—showed no im-

provement. In the next subsection, we discuss more into reasons why and contrast these findings with past work.

These findings contribute novel evidence to the literature on adaptive, autonomy-supportive goal setting in three ways. First, on a theoretical level, they highlight the importance of initial effort level as a moderator of intervention effectiveness, offering an empirical basis for tailoring goal-setting supports to student profiles. Second, on a practical level, they extend prior work by showing that data-informed goal setting activities and recommendations can meaningfully benefit students in authentic classroom contexts. Third, they suggest that the greatest returns from adaptive goal setting with rewards may come from engaging students who might otherwise remain under-involved in technology-enhanced practice due to low intrinsic motivation, providing evidence for hypotheses from prior research framing extrinsic motivators as an entry point for engaging the academically unmotivated [6, 51]. Collectively, these insights highlight the potential of adaptive goal-setting systems to reduce effort-related student difference in learning, thereby enabling more broadly effective access to the benefits of personalized, technology-enhanced instruction.

5.4.3 Intrinsic Motivation as a Potential Moderator of Intervention Differences

One explanation for that finding that high prior effort-students benefited less from the goal setting intervention is that students with high prior effort were already more intrinsically motivated, while those who benefited most from goal setting may have started with lower intrinsic motivation. In such cases, extrinsic motivators (e.g., rewards) can serve as an entry point for building interest and persistence [6, 51, 83]. The conjecture that more highly-motivated students benefited less from extrinsic rewards also aligns with the long-standing debate on whether extrinsic rewards undermine intrinsic motivation [28]. The integrative meta-analysis by Cerasoli et al. [28] meta-analysis found that incentives tied to performance can erode the positive effects of intrinsic motivation on achievement when rewards are clearly tied to performance. Similarly, another meta-analysis found that perceived choice can increase performance and intrinsic motivation [79]. These relationships could have been potentially offsetting intervention benefits for high-performing, intrinsically motivated students. However, as mentioned, whether such undermining occurs depends on perceived autonomy, something this study's design sought to protect by allowing students to choose their goals during both initial and mid-intervention goal setting and by framing rewards as part of a voluntary contract rather than as externally imposed. Monitoring perceived choice and control (i.e., autonomy) alongside intrinsic motivation could help further probe this conjecture in future work.

5.4.4 Explanatory Variables of Goal Intervention Benefit Beyond Intrinsic Motivation

Beyond motivation levels, goal orientation may further explain differential effects of this intervention seen for high-effort students. Cerafoli and Ford [27] show that mastery goal orientation, compared to performance orientation, mediates the relationship between intrinsic motivation and achievement, as mastery orientation promotes genuine skill development. Students with

performance-oriented goals may be more responsive to extrinsic rewards, potentially explaining some performance gains under goal setting. In future work, assessing students' goal hierarchies—clarifying both the “why” and “how” behind their goals—may reveal which orientations interact most productively with different goal-setting strategies. Explanatory accounts of these student-level differences could help narrow intervention effect differences in the present study.

A further possibility is that some students experienced the goal and reward structure as threatening, triggering performance-avoidance tendencies. According to Atkinson's framework, fear of failure can reduce motivation when goals are framed in ways that emphasize external control and evaluation [32]. This mechanism could be an alternative explanation for why a subset of students exhibited reduced benefits, especially if they perceived a high risk of “failing” to meet their goal. Therefore, another measure of student differences in future work to produce explanatory accounts of student intervention differences is performance and avoidance tendencies.

5.4.5 Limitations and Future Work

Given the limitation of unobservable motivational moderators mentioned above (e.g., intrinsic motivation, perceived control, failure avoidance, and goal orientation), future adaptive interventions could incorporate relevant measures to better identify which students are most likely to benefit from different forms of goal setting, goal choice, and reward tracking. Such personalization would enable a data-driven, psychometrically informed delivery of goal recommendations and incentive structures. Pairing these interventions with implementation intention interventions to improve performance [42, 44] could maintain or strengthen intrinsic motivation, as both are correlated in past research [24].

Second, this study was situated exclusively within the IXL mathematics environment. While IXL provides a rich, skill-based practice structure suitable for adaptive goal calibration, not all educational technologies offer similar granularity of skill-level data or allow students to exercise meaningful choice in practice selection. It is possible that technologies with less student choice in practice can further diminish perceived control in goal setting contexts, which, as mentioned, could limit intervention benefits of goals with contingent rewards. Future work should examine whether adaptive goal-setting frameworks generalize to other domains and platforms—particularly those with less fine-grained progress tracking or more constrained task selection—while maintaining the core elements of autonomy, calibration, and feedback.

Third, although the crossover design strengthens causal inference compared to past work on goal setting benefits in educational technology [18], the relatively short intervention periods (four weeks per condition) may have limited the opportunity to observe longer-term effects on self-regulation, goal calibration, and transfer of goal-setting strategies beyond the immediate tutoring context. Sustained implementation over a full academic year could yield deeper insights into whether adaptive goal setting produces durable changes in effort regulation and learning outcomes, or whether effects diminish once novelty or extrinsic rewards fade. While research on higher education samples predict that performance and effort benefits would vanish once extrinsic rewards are taken away [13], it is possible that (both cognitive and metacognitive) learning benefits acquired during goal setting periods aid learners beyond the period of reward provision.

5.5 Conclusion

The present findings provide empirical support for the thesis' theory of change. Consistent with H1 and H2, student-centered, adaptive goals—supported by data-driven calibration—led to higher rates of goal achievement and, though not significantly so (which is inconsistent with the theory of change), greater gains in learning compared to static, teacher-assigned goals. In line with H3, we also found that students' goal achievement itself was a significant predictor of the overall benefit students derived from the intervention, underscoring the importance of mechanisms that promote both the setting and attainment of appropriately challenging goals.

Taken together, these results point toward a dual benefit of adaptive goal setting in technology-enhanced learning: (1) immediate improvements in productivity, reflected in increased practice time and higher proficiency gains, and (2) enhanced goal calibration, as students adjusted their targets in response to performance-based feedback and were subsequently more likely to achieve them. These benefits were particularly pronounced for students with lower baseline effort, suggesting that adaptive, autonomy-supportive goal setting can help close effort-related gaps in engagement and achievement for students with low initial (intrinsic) motivation.

At the same time, the findings reveal an important boundary condition: students with high prior effort were less responsive to the intervention, raising questions about the interplay between intrinsic motivation, autonomy, and the perceived value of extrinsic incentives, all known moderators of such incentives' effect on performance in other domains. Addressing this heterogeneity will require deeper investigation into motivational profiles, goal orientations, and perceptions of control to produce explanatory accounts of student-level intervention benefits.

From a practical standpoint, the study demonstrates that even relatively simple, paper-based goal-setting processes—when paired with timely, data-informed recommendations—can be feasibly integrated into classroom tutoring environments. The design elements that appeared most influential—maintaining student autonomy, using clear performance feedback for calibration, and coupling goals with tangible but low-cost rewards—are readily transferable to other contexts and subject areas, provided the underlying learning platform can generate sufficiently granular performance data on a weekly level.

As theoretical contributions, these results extend work on self-regulated learning and self-determination theory by showing how autonomy-supportive structures can coexist with extrinsic incentives without necessarily undermining engagement for the great majority of learners, provided that choice and meaningful feedback are preserved. They also align with broader educational aims of personalizing technology-enhanced instruction not only to learners' knowledge states but also to their motivational and self-regulatory needs.

Ultimately, the evidence presented here strengthens the argument that adaptive goal setting with rewards, when implemented thoughtfully, can boost persistence, improve learning outcomes, and narrow engagement gaps across student populations, ultimately making educational technology more broadly effective.

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).

Chapter 6

Proposed Work: Intelligent Goal Support at Scale and Going Beyond Effort and Proficiency Outcomes

The prior chapters of this thesis established that adaptive, student-driven goal setting with data-driven calibration and goal adjustments can improve goal achievement and learning outcomes in hybrid tutoring environments—particularly for students with lower baseline effort. These gains were observed alongside improvements in goal calibration, as students adjusted targets in response to performance-based feedback. However, the evidence also revealed a boundary condition: students with higher prior effort benefited significantly less from either adaptive or static goal-setting interventions, suggesting that a “one-size-fits-all” approach may not optimize outcomes across the full spectrum of learner profiles.

This final phase of the thesis proposes a scaled, automated intervention that explicitly studies this heterogeneity through richer explanatory variables assessed through surveys and other qualitative data. The ultimate aim is to deliver goal-setting support that is both adaptive to student performance history and mindful of motivational student differences, while remaining scalable through automation.

6.1 Design Rationale

The findings in Chapter 5 point to three key design priorities for the final study:

1. Increase the frequency and adaptivity of goal calibration – In earlier studies, calibration opportunities were limited to a single midpoint adjustment, yet that single adjustment significantly increased goal completion (and higher rates of goal completion were correlated with larger intervention benefits). Providing more frequent, automated calibration could amplify benefits for students who are under- or overachieving and sustain momentum for those meeting or exceeding their goals.
2. Incorporate motivational and self-regulatory measures – The lack of direct measures for intrinsic motivation, self-efficacy, perceived control, and goal orientation limited the explanatory power of prior analyses. Including these constructs will enable better targeting

of supports and illuminate the psychological mechanisms underlying student-level benefits of goal setting with extrinsic rewards.

3. Studying goal hierarchies and value: As a secondary research goal, implementation intentions will be studied as a manipulated goal support variable that enables students to (a) increase their effort and performance toward goals and (b) verbalize their intrinsic “why” behind their goal through reflection, which will add more qualitative data beyond survey measures of goal orientation.
4. Automate scaffolds to enable scale and generalizability – Prior interventions relied on manual facilitation for goal feedback and adjustment in classrooms, which all used the same educational technology (IXL Math). Embedding these processes within into more classrooms via online delivery of goal setting tools and dashboards will reduce reliance on on-site personnel, increasing feasibility and generalizability for large-scale deployment while minimizing observer effects of on-site facilitators.

6.2 Research Questions

- RQ1: Calibration & Achievement:** Does increasing the *frequency and adaptivity* of goal calibration (with autonomy preserved) improve *weekly goal achievement* and *skill proficiency* relative to standard goal contracts with lower frequency?
- RQ2: Differential Effects:** Do *implementation intentions with goal reflection prompts* reduce the *effort–benefit gap*, i.e., increase benefits for *low-effort* students without diminishing outcomes (or intrinsic motivation) for *high-effort* students?
- RQ3: Mechanisms:** To what extent do *intrinsic motivation, self-efficacy, perceived control, failure avoidance, and goal orientation* moderate and explain the effects of adaptive goal support on achievement and learning?

6.3 Hypotheses

- H1: Main Effect—Implementation Intention:** Students receiving implementation intentions will experience larger intervention benefits on effort and skill mastery [44] and larger intrinsic motivation [24], which in turn will make the intervention more effective for students with high prior effort (as opposed to not effective as observed in Chapter 5).
- H2: Main Effect—Calibration:** Students receiving *more frequent, adaptive calibration* will show higher *goal achievement, practice effort, goal calibration* and *skill proficiency* than those with less regular goal feedback and adjustment in line with findings in Chapter 5.
- H3: Main Effect—Motivation and Choice:** Students receiving *more frequent, adaptive calibration* will show higher *perceived choice and intrinsic motivation* consistent with findings in [79].

H4: Autonomy-Preserving Rewards: Adaptive supports with regular goal check ins preserve *student choice* when extrinsic rewards are present and this perceived choice will be sufficiently high such that there will be *no significant decrease* in *intrinsic motivation* post-intervention consistent with [28].

H5: Mechanism—Motivational Profiles: Effects of adaptive goal supports will be *stronger* when students have higher *perceived control* [28, 79], *mastery goal orientation* [27], and weaker *performance-avoidance* [32].

6.4 Study Design

Plan A (Fall 2025): Hybrid Tutoring

Design: Similar to previous studies of this thesis, the final study will most likely follow an interrupted time series design spanning approximately 12 weeks; baseline (3-5 weeks) followed by intervention (7-9 weeks) across at least three schools in a hybrid tutoring program, leading to a sample of approximately 300 or more students. Educational technologies may include IXL, i-Ready, and MobyMax, improving generalizability of effects beyond previous studies which were limited to IXL only.

Conditions: All participating students will receive the adaptive goal-setting treatment, consisting of continuous, system-delivered calibration opportunities and automated feedback on goal progress. The only experimental manipulation will be the presence or absence of *implementation intentions with goal reflection prompt* at the point of goal setting, which will be randomly assigned at the student level.

- **Implementation Intentions Condition:** In addition to setting or adjusting their weekly goal, students will complete a short, structured “if-then” plan linking a specific anticipated challenge (e.g., distractions, time constraints) to a planned strategy for overcoming it. The prompt will also include a reflection item asking students to articulate the personal value or “why” behind their goal, which will be prompted either in-system or by the human tutor.
- **No Implementation Intentions Condition:** Students will set or adjust their weekly goal using the same adaptive calibration interface but will not receive the implementation intentions or value-reflection prompt.

This structure aims at students benefiting from the frequent adaptive calibration treatment while isolating the added effect of implementation intentions on effort and motivational outcomes.

Survey Measures: Pre- and post-intervention in the Fall semester, all students will complete brief validated survey scales to capture motivational and self-regulatory constructs for testing **RQ3** as well as **H3+H5** described above. These may be administered via online forms or with pen and paper. Acknowledging the high breadth of these measures and potential practical constraints in administering them given available student and classroom time, they are listed in order of priority in which they will administered if a subset must be chosen:

- **Motivation Questionnaire (MMQ) for secondary school students - Intrinsic value subscale to measure intrinsic motivation** [40]. Three items, 5-point Likert scale.
- **Intrinsic Motivation Inventory (IMI) – Perceived Choice** for autonomy [87]. Seven items, 7-point Likert scale.
- **Achievement Goal Questionnaire–Revised (AGQ-R)** [38] with self-performed adaptations to middle school mathematics to measure performance and mastery goal orientation as well as achievement approach and failure avoidance. Four times three items, 5-point Likert scale.

Goal Cadence Variation Across Sites: Hybrid tutoring partner sites vary in their goal-setting cadences and corresponding data update/feedback schedules: Some sites review and adjust goals weekly, others follow a biweekly cadence, certain sites set goals monthly or once per semester. This natural variation will be leveraged to test **H1** by examining whether more frequent adaptive calibration opportunities lead to larger gains in goal achievement, practice effort, and skill proficiency. The variation also allows exploratory dose–response analyses to determine the optimal calibration frequency for different student profiles.

Primary outcomes modeled through interrupted time series models, similar to previous studies:

- Weekly goal achievement (0/1).
- Practice effort (minutes practiced per week).
- Skill proficiency gains (skills mastered per week).

Secondary outcomes analyzed through regression models by aggregating measures on the student level, similar to previous studies:

- Calibration accuracy: $|\text{goal} - \text{achieved}|$ divided by goal.
- Changes in survey-measured motivational constructs (listed above).
- Moderation of survey-measured motivational constructs at pre with goal intervention benefits on the student level (calculated analogous to Chapter 5).
- Stretch goal: Manual or semi-automated content analysis of tutor-student transcripts during goal purpose prompts (e.g., the “why” behind student goals).

6.5 Preliminary Interface Design

To build on existing infrastructure of the hybrid tutoring project, every student will be assigned an effort goal (e.g., number of minutes practiced) and progress goal (e.g., number of skills mastered) within a pre-determined and standardized timeframe for each partner school. Once a cycle of that timeframe is completed, goal data will be uploaded into the web application to populate relevant goal analytics and recommendations, which hybrid tutors will peruse with the student via screen share as part of the regular tutoring activities. As the study’s manipulation, about 50% of students will be randomly assigned to a condition where they are asked to generate implementation intentions when setting and adjusting goals (as outlined above).

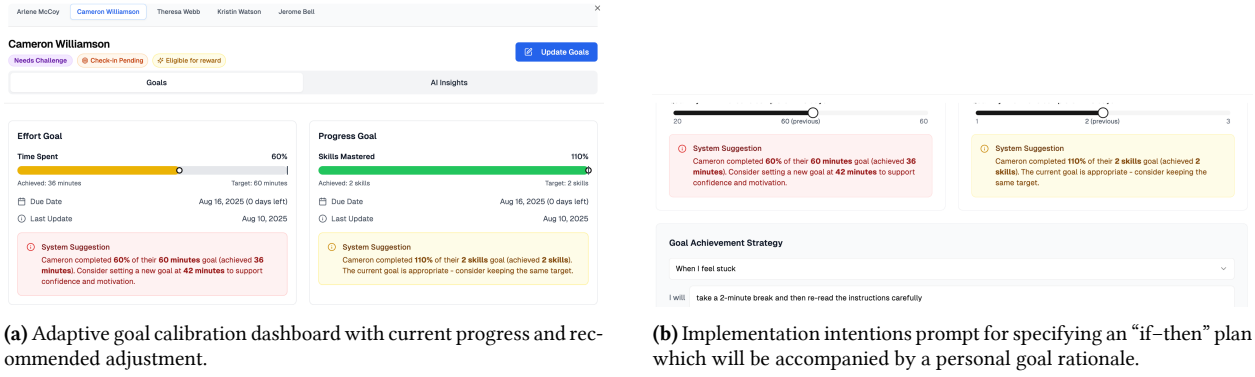


Figure 6.1: Preliminary interface designs for the adaptive goal-setting intervention. The left panel shows the goal calibration dashboard; the right panel displays the implementation intentions.

6.6 Plan B (Fall 2025 or Spring 2026): Custom Dashboard

Plan B would replace or supplement the hybrid tutoring study with a large-scale, fine-grained log data intervention using the MathTutor intelligent tutoring system integrated with TutorShop [2], augmented by a student-facing dashboard to support goal setting and monitoring [23]. Scheduled for the 2025/26 academic year, this version would leverage development resources (led by the doctoral candidate) to optimize tutoring system integration and streamline data capture. School recruitment would aim for classroom research without the (at present) arguably rare hybrid tutoring setup, enabling stronger generalizability compared to prior work. The intervention design would closely align with the grant-funded requirements of comparing the intervention to a paper-based baseline control condition without goal support, while providing richer, time-stamped interaction logs for analyzing goal-setting effects at a more granular process level but with a weaker control condition.

6.7 Anticipated Contributions

The proposed work will make both theoretical and practical contributions to the literature on goal setting, motivation, and scalable learning analytics in technology-enhanced education.

1. Advancing Theory on Adaptive Goal Setting and Motivation. By explicitly examining how calibration frequency, motivational profiles, and implementation intentions interact, this study will refine existing models of adaptive goal setting within the framework of self-determination theory [34, 94]. Prior chapters demonstrated that adaptive calibration benefits low-effort students more than high-effort students but left open questions about psychological mechanisms. The proposed design integrates direct measures of intrinsic motivation, perceived control, failure avoidance, and goal orientation, enabling more precise theorizing about why and for whom adaptive supports work, which can inform goal delivery at more fine-grain levels of personalization in future work.

2. Introducing Goal Reflection and Implementation Intentions as a Scalable Support Mechanism. While implementation intentions have been widely studied in laboratory and

short-term interventions, their integration into an ongoing, data-driven goal-setting process in authentic classroom contexts remains underexplored. This work will evaluate whether pairing adaptive calibration with reflection on implementation intentions, sometimes shown to correlate with higher intrinsic motivation, strengthens persistence and achievement, particularly for low-effort students, without undermining intrinsic motivation for high-effort students through extrinsic rewards.

3. Identifying Optimal Calibration Frequency for Different Learner Profiles. Natural variation in goal-setting cadences across partner sites will enable dose–response analyses of calibration frequency. This will yield empirically grounded recommendations for the optimal number and timing of calibration opportunities, balancing the need for timely feedback with the workload of frequent goal adjustments.

4. Scaling Goal Support Through Platform-Embedded Automation. The proposed work moves beyond facilitator-dependent implementations by embedding goal calibration, feedback, and adjustment prompts directly into online platforms and dashboards. Both Plan A and Plan B emphasize automation, allowing interventions to be deployed without dedicated on-site personnel, thereby reducing costs, minimizing observer effects, and increasing scalability to multiple educational contexts and sites.

5. Generating High-Resolution Process Data for Learning Analytics. Through integration with multiple educational technology platforms (Plan A) or intelligent tutoring systems with TutorShop (Plan B), the study will produce fine-grained, time-stamped interaction logs that link goal-setting behaviors to subsequent learning actions and outcomes. These datasets will support process-level analyses of momentum effects, calibration accuracy trajectories, and the micro-patterns of goal pursuit in authentic classroom environments.

6. Strengthening External Validity and Generalizability. By extending beyond a single platform or delivery model to include multiple educational technologies (Plan A) or an ITS with broader classroom adoption potential (Plan B), the study will assess whether adaptive goal supports generalize across content domains, instructional modalities, and school contexts. This test of generalizability increases the robustness and policy relevance of this thesis’ findings for large-scale deployment.

6.8 Timeline of Completion

The planned timeline for completing this work is presented in Table 6.1.

Date	Expected Activity
September 12th 2025	Proposal Defense
September 2025	Finalization of Study Preparations
October - December 2025	Running of 6-8 week Field Study
October - December 2025	Job Market Applications (Materials Completed)
December 2025-February 2026	Data Analysis and Interpretation
March 2026-April 2026	Job Market and Pending Funding Decisions
Plan A: February-May 2026	Dissertation Writing and Defense Preparation for June Defense
Plan B: February-May 2026	Running and Inclusion of Second Study into Thesis; Later Defense

Table 6.1: Proposed timeline for addressing feedback and goal achievement challenges in hybrid tutoring.

Bibliography

- [1] Marc A Adams, Jane C Hurley, Michael Todd, Nishat Bhuiyan, Catherine L Jarrett, Wesley J Tucker, Kevin E Hollingshead, and Siddhartha S Angadi. 2017. Adaptive goal setting and financial incentives: a 2×2 factorial randomized controlled trial to increase adults' physical activity. *BMC public health* 17 (2017), 1–16.
- [2] Vincent Aleven, Conrad Borchers, Yun Huang, Tomohiro Nagashima, Bruce McLaren, Paulo Carvalho, Octav Popescu, Jonathan Sewall, and Kenneth Koedinger. 2025. An integrated platform for studying learning with intelligent tutoring systems: CTAT+ TutorShop. *arXiv preprint arXiv:2502.10395* (2025).
- [3] Vincent Aleven, Bruce McLaren, Ido Roll, and Kenneth Koedinger. 2006. Toward meta-cognitive tutoring: A model of help seeking with a Cognitive Tutor. *International Journal of Artificial Intelligence in Education* 16, 2 (2006), 101–128.
- [4] Vincent Aleven, Elizabeth A McLaughlin, R Amos Glenn, and Kenneth R Koedinger. 2016. Instruction based on adaptive learning technologies. *Handbook of research on learning and instruction* 2 (2016), 522–560.
- [5] Abdullah Alwahbi. 2020. The Use of Contingency Contracting in Educational Settings: A Review of the Literature. *Educational Research and Reviews* 15, 6 (2020), 327–335.
- [6] Michael W Asher and Judith M Harackiewicz. 2024. Using choice and utility value to promote interest: Stimulating situational interest in a lesson and fostering the development of interest in statistics. *Journal of Educational Psychology* (2024).
- [7] Elisabetta Aurino, Edward Tsinigo, and Sharon Wolf. 2022. *Nudges to improve learning and gender parity: preliminary findings on supporting parent-child educational engagement during Covid-19 using mobile phones*. Technical Report. EdTech Hub.
- [8] Roger Azevedo, François Bouchet, Melissa Duffy, Jason Harley, Michelle Taub, Gregory Trevors, Elizabeth Cloude, Daryn Dever, Megan Wiedbusch, Franz Wortha, and others. 2022. Lessons learned and future directions of MetaTutor: Leveraging multichannel data to scaffold self-regulated learning with an intelligent tutoring system. *Frontiers in Psychology* 13 (2022), 813632.
- [9] Roger Azevedo, Amy M Witherspoon, Amber Chauncey, Candice Burkett, and Ashley Fike. 2009. MetaTutor: A MetaCognitive Tool for Enhancing Self-Regulated Learning.. In *AAAI fall symposium: Cognitive and metacognitive educational systems*. 14–19.
- [10] Ryan Baker, Jason Walonoski, Neil Heffernan, Ido Roll, Albert Corbett, and Kenneth Koedinger. 2008. Why students engage in “gaming the system” behavior in interactive learning environments. *Journal of Interactive Learning Research* 19, 2 (2008), 185–224.
- [11] Ryan Sjd Baker. 2007. Modeling and understanding students' off-task behavior in intelligent tutoring systems. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 1059–1068.
- [12] Dale J Barr, Roger Levy, Christoph Scheepers, and Harry J Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language* 68, 3 (2013), 255–278.
- [13] Lisa Barrow and Cecilia Elena Rouse. 2018. Financial incentives and educational investment: The impact of performance-based scholarships on student time use. *Education Finance and Policy* 13, 4 (2018), 419–448.
- [14] Bozhidar M Bashkov. 2021. Assessing the impact of IXL Math over three years: A quasi-experimental study. *ESSA Research Report. Online Submission* (2021).

-
- [15] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of statistical software* 67 (2015), 1–48.
 - [16] George G Bear, Jessica C Slaughter, Lindsey S Mantz, and Elizabeth Farley-Ripple. 2017. Rewards, praise, and punitive consequences: Relations with intrinsic and extrinsic motivation. *Teaching and Teacher Education* 65 (2017), 10–20.
 - [17] Benjamin M Bolker. 2015. Linear and generalized linear mixed models. *Ecological statistics: contemporary theory and application* (2015), 309–333.
 - [18] Conrad Borchers, Alex Houk, Vincent Aleven, and Kenneth R. Koedinger. 2025a. Engagement and Learning Benefits of Goal Setting with Rewards in Human-AI Tutoring. In *Proceedings of the 26th International Conference on Artificial Intelligence in Education*.
 - [19] Conrad Borchers, Ha Tien Nguyen, Paulo F. Carvalho, Kenneth R. Koedinger, and Vincent Aleven. 2025b. Involving Parents in Tutoring Systems to Increase Content Confidence: A Design Probe Study. In *Proceedings of the 26th International Conference on Artificial Intelligence in Education (AIED 2025)*. Palermo, Italy.
 - [20] Conrad Borchers, Jeroen Ooge, Cindy Peng, and Vincent Aleven. 2025c. 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*. 810–816.
 - [21] Conrad Borchers, Cindy Peng, Qianru Lyu, Paulo F. Carvalho, Kenneth R. Koedinger, and Vincent Aleven. 2025d. Student Perceptions of Adaptive Goal Setting Recommendations: A Design Prototyping Study. In *Proceedings of the 26th International Conference on Artificial Intelligence in Education*.
 - [22] Conrad Borchers, Cindy Peng, Qianru Lyu, Paulo F. Carvalho, Kenneth R. Koedinger, and Vincent Aleven. 2025e. Student Perceptions of Adaptive Goal Setting Recommendations: A Design Prototyping Study. In *Proceedings of the 26th International Conference on Artificial Intelligence in Education (AIED 2025) (Lecture Notes in Computer Science)*. Springer, Palermo, Italy.
 - [23] Conrad Borchers, Cindy Peng, Qianru Lyu, Paulo F Carvalho, Kenneth R Koedinger, and Vincent Aleven. 2025f. Student perceptions of adaptive goal setting recommendations: a design prototyping study. In *International Conference on Artificial Intelligence in Education*. Springer, 244–251.
 - [24] Tracey A Brickell and Nikos LD Chatzisarantis. 2007. Using self-determination theory to examine the motivational correlates and predictive utility of spontaneous exercise implementation intentions. *Psychology of Sport and exercise* 8, 5 (2007), 758–770.
 - [25] Urie Bronfenbrenner. 2000. *Ecological systems theory*. American Psychological Association.
 - [26] Susan Bull. 2020. There are open learner models about! *IEEE Transactions on Learning Technologies* 13, 2 (2020), 425–448.
 - [27] Christopher P Cerasoli and Michael T Ford. 2014. Intrinsic motivation, performance, and the mediating role of mastery goal orientation: A test of self-determination theory. *The Journal of psychology* 148, 3 (2014), 267–286.
 - [28] Christopher P Cerasoli, Jessica M Nicklin, and Michael T Ford. 2014. Intrinsic motivation and extrinsic incentives jointly predict performance: a 40-year meta-analysis. *Psychological bulletin* 140, 4 (2014), 980.
 - [29] Daniel H Chang, Michael Pin-Chuan Lin, Shiva Hajian, and Quincy Q Wang. 2023. Educational design principles of using AI chatbot that supports self-regulated learning in education: Goal setting, feedback, and personalization. *Sustainability* 15, 17 (2023), 12921.
 - [30] JO Cooper. 2007. *Applied behavior analysis*. Pearson/Merrill-Prentice Hall.
 - [31] Juliet M Corbin and Anselm Strauss. 1990. Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology* 13, 1 (1990), 3–21.
 - [32] Martin V Covington and Kimberly J Müeller. 2001. Intrinsic versus extrinsic motivation: An approach/avoidance reformulation. *Educational psychology review* 13, 2 (2001), 157–176.
 - [33] Edward L Deci, Richard Koestner, and Richard M Ryan. 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin* 125, 6 (1999), 627.

- [34] Edward L Deci and Richard M Ryan. 1980. Self-determination theory: When mind mediates behavior. *The Journal of mind and Behavior* (1980), 33–43.
- [35] Edward L Deci and Richard M Ryan. 1985. The general causality orientations scale: Self-determination in personality. *Journal of research in personality* 19, 2 (1985), 109–134.
- [36] Jessica DeMink-Carthew, Mark W Olofson, Life LeGeros, Steven Netcoh, and Susan Hennessey. 2017. An analysis of approaches to goal setting in middle grades personalized learning environments. *RMLE Online* 40, 10 (2017), 1–11.
- [37] Melissa C Duffy and Roger Azevedo. 2015. Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior* 52 (2015), 338–348.
- [38] Andrew J Elliot and Kou Murayama. 2008. On the measurement of achievement goals: critique, illustration, and application. *Journal of educational psychology* 100, 3 (2008), 613.
- [39] Mingyu Feng, Chunwei Huang, and Kelly Collins. 2023. Promising Long Term Effects of ASSISTments Online Math Homework Support. In *International Conference on Artificial Intelligence in Education*. Springer, 212–217.
- [40] Logan Fiorella, So Yoon Yoon, Kinnari Atit, Jason R Power, Grace Panther, Sheryl Sorby, David H Uttal, and Norma Veurink. 2021. Validation of the Mathematics Motivation Questionnaire (MMQ) for secondary school students. *International Journal of STEM Education* 8, 1 (2021), 52.
- [41] Gillian Gold, Conrad Borchers, and Paulo F. Carvalho. 2024. Students’ Academic Performance and Goal Orientation Relate to Initial Knowledge but Not Learning Rate. In *14th International Conference on Learning Analytics & Knowledge (LAK)*. Kyoto, Japan.
- [42] Peter M Gollwitzer. 1999. Implementation intentions: strong effects of simple plans. *American psychologist* 54, 7 (1999), 493.
- [43] Peter M Gollwitzer. 2012. Mindset theory of action phases. *Handbook of theories of social psychology* 1 (2012), 526–545.
- [44] Peter M Gollwitzer and Paschal Sheeran. 2006. Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in experimental social psychology* 38 (2006), 69–119.
- [45] Alyssa R Gonzalez-DeHass and Patricia P Willems. 2024. Middle-school students and digital homework: The evolving role of family engagement. *Middle School Journal* 55, 4 (2024), 25–34.
- [46] Barbara A Greene, Teresa K DeBacker, Bhuvaneswari Ravindran, and A Jean Krows. 1999. Goals, values, and beliefs as predictors of achievement and effort in high school mathematics classes. *Sex roles* 40, 5 (1999), 421–458.
- [47] Ashish Gurung, Jionghao Lin, Zhongtian Huang, Conrad Borchers, Ryan S Baker, Vincent Aleven, and Kenneth R Koedinger. 2025. Starting Seatwork Earlier as a Valid Measure of Student Engagement. *arXiv preprint arXiv:2505.13341* (2025).
- [48] Allyson F Hadwin and Elizabeth A Webster. 2013. Calibration in goal setting: Examining the nature of judgments of confidence. *Learning and instruction* 24 (2013), 37–47.
- [49] David Hammer and Leema K Berland. 2014. Confusing claims for data: A critique of common practices for presenting qualitative research on learning. *Journal of the Learning Sciences* 23, 1 (2014), 37–46.
- [50] Muhammad Asif Hasan, Nurul Fazmidar Mohd Noor, Siti Soraya Binti Abdul Rahman, and Mohammad Mustaneer Rahman. 2020. The transition from intelligent to affective tutoring system: a review and open issues. *IEEE Access* 8 (2020), 204612–204638.
- [51] Suzanne Hidi and Judith M Harackiewicz. 2000. Motivating the academically unmotivated: A critical issue for the 21st century. *Review of educational research* 70, 2 (2000), 151–179.
- [52] Kenneth Holstein, Vincent Aleven, and Nikol Rummel. 2020. 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. Springer, 240–254.

-
- [53] Kathleen V Hoover-Dempsey and Howard M Sandler. 1997. Why do parents become involved in their children’s education? *Review of educational research* 67, 1 (1997), 3–42.
 - [54] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, and others. 2003. Technology probes: inspiring design for and with families. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 17–24.
 - [55] William H Jeynes. 2007. The relationship between parental involvement and urban secondary school student academic achievement: A meta-analysis. *Urban education* 42, 1 (2007), 82–110.
 - [56] Qiao Jin, Conrad Borchers, Stephen Fancsali, and Vincent Aleven. 2025. Who to Help? A Time-Slice Analysis of K-12 Teachers’ Decisions in Classes with AI-Supported Tutoring. In *Proceedings of the 18th International Conference on Educational Data Mining (EDM 2025)*. Palermo, Italy.
 - [57] Alice L Kahle and Mary Lou Kelley. 1994. Children’s homework problems: A comparison of goal setting and parent training. *Behavior therapy* 25, 2 (1994), 275–290.
 - [58] Anya Kamenetz. 2019. It’s a smartphone life: More than half of US children now have one. *National Public Radio* 31 (2019).
 - [59] Shamy Karumbaiah, Conrad Borchers, Ann-Christin Falhs, Kenneth Holstein, Nikol Rummel, and Vincent Aleven. 2023a. Teacher noticing and student learning in human-AI partnered classrooms: a multimodal analysis. In *Proceedings of the 17th International Conference of the Learning Sciences-ICLS 2023*, pp. 1042–1045. International Society of the Learning Sciences.
 - [60] Shamy Karumbaiah, Conrad Borchers, Tianze Shou, Ann-Christin Falhs, Pinyang Liu, Tomohiro Nagashima, Nikol Rummel, and Vincent Aleven. 2023b. A spatiotemporal analysis of teacher practices in supporting student learning and engagement in an AI-enabled classroom. In *International Conference on Artificial Intelligence in Education*. Springer, 450–462.
 - [61] Kenneth R Koedinger, Ryan Sjd Baker, Kyle Cunningham, Alida Skogsholm, Brett Leber, and John Stamper. 2010. A data repository for the EDM community: The PSLC DataShop. *Handbook of educational data mining* 43 (2010), 43–56.
 - [62] Kenneth R Koedinger, Paulo F Carvalho, Ran Liu, and Elizabeth A McLaughlin. 2023. An astonishing regularity in student learning rate. *Proceedings of the National Academy of Sciences* 120, 13 (2023), e2221311120.
 - [63] Kenneth R Koedinger, Jihee Kim, Julianna Zhuxin Jia, Elizabeth A McLaughlin, and Norman L Bier. 2015. Learning is not a spectator sport: Doing is better than watching for learning from a MOOC. In *Proceedings of the second (2015) ACM Learning@Scale*. 111–120.
 - [64] Richard Koestner. 2008. Reaching one’s personal goals: A motivational perspective focused on autonomy. *Canadian Psychology/Psychologie Canadienne* 49, 1 (2008), 60.
 - [65] Richard Koestner, Nancy Otis, Theodore A Powers, Luc Pelletier, and Hugo Gagnon. 2008. Autonomous motivation, controlled motivation, and goal progress. *Journal of personality* 76, 5 (2008), 1201–1230.
 - [66] James A Kulik and John D Fletcher. 2016. Effectiveness of intelligent tutoring systems: a meta-analytic review. *Review of educational research* 86, 1 (2016), 42–78.
 - [67] Roberto Legaspi, Wenzhen Xu, Tatsuya Konishi, Shinya Wada, and Yuichi Ishikawa. 2022. Multidimensional analysis of sense of agency during goal pursuit. In *Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*. 34–47.
 - [68] Xinyu Li, Yizhou Fan, Tongguang Li, Mladen Raković, Shaveen Singh, Joep van der Graaf, Lyn Lim, Johanna Moore, Inge Molenaar, Maria Bannert, and others. 2025. FLORA engine: Using Analytics to Measure and Facilitate Learners’ Own Regulation Activities. *Journal of Learning Analytics* (2025), 1–24.
 - [69] Edwin A Locke and Gary P Latham. 2002. Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American psychologist* 57, 9 (2002), 705.

BIBLIOGRAPHY

- [70] Edwin A Locke and Gary P Latham. 2019. The development of goal setting theory: A half century retrospective. *Motivation Science* 5, 2 (2019), 93.
- [71] Yanjin Long and Vincent Aleven. 2017. Enhancing learning outcomes through self-regulated learning support with an open learner model. *User Modeling and User-Adapted Interaction* 27 (2017), 55–88.
- [72] Yanjin Long, Kenneth Holstein, and Vincent Aleven. 2018. What exactly do students learn when they practice equation solving? refining knowledge components with the additive factors model. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*. 399–408.
- [73] Crystal Derika-Mae Matthews. 2025. *Middle School Mathematics Teachers' Use of Diverse Mathematics Educational Technology Tools: IXL, i-Ready, and/or MATHia*. Ph.D. Dissertation. Walden University.
- [74] David McDowall, Richard McCleary, and Bradley J Bartos. 2019. *Interrupted time series analysis*. Oxford University Press.
- [75] Alissa J Mrazek, Michael D Mrazek, Joshua R Ortega, Rosie R Ji, Sana S Karimi, Chelsea S Brown, Chelsie A Alexander, Maliha Khan, Rhoshen Panahi, Madeline Sadoff, and others. 2021. Teenagers' smartphone use during homework: an analysis of beliefs and behaviors around digital multitasking. *Education Sciences* 11, 11 (2021), 713.
- [76] Krista R Muis, Philip H Winne, and John Ranellucci. 2016. The Role of Calibration Bias and Performance Feedback in Achievement Goal Regulation. *International Education Research* 4, 1 (2016), 14–36.
- [77] Lisa A Newland. 2015. Family well-being, parenting, and child well-being: Pathways to healthy adjustment. *Clinical psychologist* 19, 1 (2015), 3–14.
- [78] Ha Tien Nguyen, Conrad Borchers, Meng Xia, and Vincent Aleven. 2024. Designing Tools for Caregiver Involvement in Intelligent Tutoring Systems for Middle School Mathematics. In *Proceedings of the 18th International Conference of the Learning Sciences-ICLS 2024*, pp. 865-872. International Society of the Learning Sciences.
- [79] Erika A Patall, Harris Cooper, and Jorgianne Civey Robinson. 2008. The effects of choice on intrinsic motivation and related outcomes: a meta-analysis of research findings. *Psychological bulletin* 134, 2 (2008), 270.
- [80] Gretchen Gimpel Peacock, Ruth A Ervin, Edward J Daly, and Kenneth W Merrell. 2009. *Practical handbook of school psychology: Effective practices for the 21st century*. Guilford Press.
- [81] Praise-El K Pelemo. 2022. *Parental Involvement: Investigating the Roles, Barriers, and Strategies of Effective Parent-Teacher Relationships in Middle School Math*. Ph.D. Dissertation. Southeastern University.
- [82] Cindy Peng, Conrad Borchers, and Vincent Aleven. 2024. 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.
- [83] Stacy J Priniski, Cameron A Hecht, and Judith M Harackiewicz. 2018. Making learning personally meaningful: A new framework for relevance research. *The Journal of Experimental Education* 86, 1 (2018), 11–29.
- [84] Ido Roll, Eliane Stampfer Wiese, Yanjin Long, Vincent Aleven, and Kenneth R Koedinger. 2014. Tutoring self-and co-regulation with intelligent tutoring systems to help students acquire better learning skills. *Design recommendations for intelligent tutoring systems* 2 (2014), 169–182.
- [85] Richard M Ryan and Edward L Deci. 2000. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology* 25, 1 (2000), 54–67.
- [86] Annika Schweizer, Sebastian Niedlich, Judith Adamczyk, and Inka Bormann. 2017. Approaching trust and control in parental relationships with educational institutions. *Studia paedagogica* 22, 2 (2017), 97–115.
- [87] Self-Determination Theory. 2022. *Intrinsic Motivation Inventory (IMI) – Complete*. https://selfdeterminationtheory.org/wp-content/uploads/2022/02/IMI_Complete.pdf Scale description and full items.

-
- [88] Karrie A Shogren, Valerie L Mazzotti, Tyler A Hicks, Sheida K Raley, Daria Gerasimova, Jesse R Pace, Stephen M Kwiatek, Darcy Fredrick, Jared H Stewart-Ginsburg, Richard Chapman, and others. 2024. The Goal Setting Challenge App: Impact on transition goal attainment outcomes of students with disabilities. *Career Development and Transition for Exceptional Individuals* 47, 1 (2024), 4–18.
 - [89] Mary Ann Simpson, Kole A Norberg, and Stephen E Fancsali. 2024. Replicating an “Astonishing Regularity in Student Learning Rate”. In *Proceedings of the 17th International Conference on Educational Data Mining*. 420–425.
 - [90] Maria Theobald, Wy Ming Lin, Michiko Sakaki, Kou Murayama, and Garvin Brod. 2025. How (much) do people revise their goals after success and failure? The Triple-A model of goal revision. *Educational Psychologist* (2025), 1–21.
 - [91] Danielle Thomas, Xinyu Yang, Shivang Gupta, Adetunji Adeniran, Elizabeth Mclaughlin, and Kenneth Koedinger. 2023. 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*. 250–261.
 - [92] Danielle R Thomas, Jionghao Lin, Erin Gatz, Ashish Gurung, Shivang Gupta, Kole Norberg, Stephen E Fancsali, Vincent Aleven, Lee Branstetter, Emma Brunskill, and others. 2024. Improving student learning with hybrid human-AI tutoring: A three-study quasi-experimental investigation. In *Proceedings of the 14th Learning Analytics and Knowledge Conference*. 404–415.
 - [93] Ulrich Trautwein, Oliver Lüdtke, Inge Schnyder, and Alois Niggli. 2006. Predicting homework effort: support for a domain-specific, multilevel homework model. *Journal of educational psychology* 98, 2 (2006), 438.
 - [94] Robert J Vallerand. 2000. Deci and Ryan’s self-determination theory: A view from the hierarchical model of intrinsic and extrinsic motivation. *Psychological inquiry* 11, 4 (2000), 312–318.
 - [95] Anouschka van Leeuwen, Carolien AN Knoop-van Campen, Inge Molenaar, and Nikol Rummel. 2021. How teacher characteristics relate to how teachers use dashboards: Results from two case studies in K-12. *Journal of Learning Analytics* 8, 2 (2021), 6–21.
 - [96] Devika Venugopalan, Ziwen Yan, Conrad Borchers, Jionghao Lin, and Vincent Aleven. 2025. Combining large language models with tutoring system intelligence: A case study in caregiver homework support. In *Proceedings of the 15th International Learning Analytics and Knowledge Conference*. 373–383.
 - [97] Kristin Wäschle, Anne Allgaier, Andreas Lachner, Siegfried Fink, and Matthias Nückles. 2014. Procrastination and self-efficacy: Tracing vicious and virtuous circles in self-regulated learning. *Learning and instruction* 29 (2014), 103–114.
 - [98] Jian Yang, Hongcheng Guo, Yuwei Yin, Jiaqi Bai, Bing Wang, Jiaheng Liu, Xinnian Liang, Linzheng Cahi, Liquan Yang, and Zhoujun Li. 2024. m3P: Towards Multimodal Multilingual Translation with Multimodal Prompt. *arXiv preprint arXiv:2403.17556* (2024).