International Society of the Learning Sciences

Designing Homework Support Tools for Middle School Mathematics Using Intelligent Tutoring Systems

Cindy Peng, University of California San Diego, cipeng@ucsd.edu
Conrad Borchers, Carnegie Mellon University, cborcher@cs.cmu.edu
Vincent Aleven, Carnegie Mellon University, aleven@cs.cmu.edu

Abstract: Prior studies identified effective, but mainly non-digital, homework aids. This research involved 18 middle school students in a lo-fi prototyping study to integrate traditional homework support tools with intelligent tutoring systems (ITS), leveraging rich log data for personalized learning. Feature investigations in standardized diaries, goal setting, homework graphing, and reminders revealed a preference for goal-setting, informing a refined prototype exploring students' attitudes towards self-set versus system-generated goals. Results indicate students prefer system-recommended goals, express goals by problem count and time, and value goal difficulty feedback to foster autonomy. Providing data-driven feedback and fading system-set goals in ITS could improve self-regulation during homework practices and equity.

Introduction and background

Homework is vital for academic achievement. Yet, not all students benefit equally, as effective homework practices require self-regulation. To address this, past work has devised strategies such as goal setting, learning strategy selection, progress monitoring, and outcome evaluation (Bembenutty, 2011). Specifically, goal setting boosts self-efficacy beliefs, as learners who set specific academic goals are more likely to engage in sustained effort and use self-regulatory strategies to achieve them (Zimmerman, 2002). While these homework support tools have been predominantly based on pen and paper, much of today's homework involves or is supported by technology, indicating the importance of effectively integrating educational technology into homework practices, rather than technology being a disruptive companion (Mrazek et al., 2021).

A gap exists in utilizing technologies to ensure homework completion, critical for academic success. Perhaps log data capturing student interactions could be leveraged to develop strategies that support homework – yet past work has not explored this opportunity. This scenario presents a dual-faceted challenge: the design dilemma of effectively adapting traditional homework support strategies into digital formats and the research imperative to evaluate these digital tools' impact and efficacy on students' existing digital practices and habits. Intelligent tutoring systems (ITS) could close the gap between data-driven support and homework tools, as their computer-assisted performance has a track record of benefitting student learning (Kulik & Fletcher, 2016). However, only *some* work has merged ITS with external support structures for deliberate practice. Long and Aleven (2013) showed ITS-data-enhanced skill diaries for goal setting and self-assessment significantly helped lower-performing students. Similarly, the ASSISTments ITS for homework, by leveraging detailed performance data to generate reports for tailored instruction, improved mathematics learning outcomes (Feng et al., 2023).

The present study develops design prototypes merging traditional homework support with ITS for deliberate practice. Our design prototypes integrate four analog strategies into digital ITS through log data on assignment progress, student performance, and estimated skill mastery levels. Further, we investigate students' perspectives—an area previously overlooked—to discover which support features resonate most with middle school students. In line with prior research's recommendations (Mrazek et al., 2021), the study also incorporates self-regulated learning strategies in mobile environments, designing mobile-friendly features for math practice on the go. We investigate two research questions (RQs): *RQ1: How do students perceive and prefer the integration of data-driven digital homework tools with traditional homework support strategies? RQ2: What are students' perceptions and preferences when interacting with an enhanced research prototype particularly focusing on goal-setting and data-driven recommendations derived from an intelligent tutoring system?*

Methods

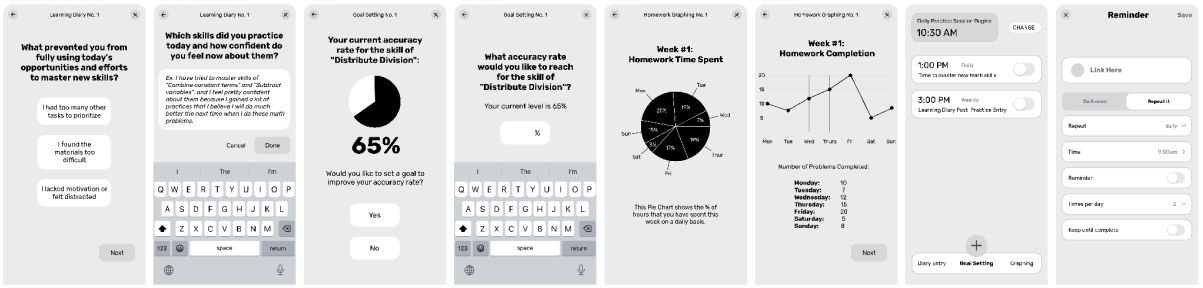
Eighteen middle school students in grades 6-10, predominantly from urban areas in the Mid-Atlantic and Pacific U.S. regions were recruited. Among them, 16.7% were in 6th grade, 33.3% in 7th, 38.9% in 8th, 5.6% in 9th, and 5.6% in 10th, comprising ten males and eight females, all Caucasian (10) or East Asian (8). Participants were divided into two phases with nine in each. Recruitment utilized multiple channels established in prior research partnerships with local schools and districts, including snowball sampling through university parent outreach centers. Interviews, conducted remotely via Zoom, lasted around 60 minutes. Transcripts were

analyzed qualitatively using open coding, consolidating derived themes through team discussions among three researchers. Extracted quotes were grouped into themes based on the four design features' perceived utility (addressing RQ1) and perceptions of participants' interactions with a refined goal-setting tool (addressing RQ2).

In the first phase, we gathered participants' self-reported backgrounds and preferences for four design features (Figure 1), inspired by evidence-based analog homework support tools and integrated with personalized ITS data to enhance student learning. We evaluated initial receptiveness, utility, and favored features for enhancing mathematics homework practices on mobile platforms. Structured questions focused on interface elements, content displays, user interactions, and design feedback. Participants ranked the features based on perceived effectiveness at the end, noting standout aspects, improvement areas, and less effective elements.

Figure 1

Prototypes of the four features investigated during Phase One, designed to augment Mathtutor (1).



Standardized Diaries guide students in self-reflection through question prompts on learning activities, enhancing self-regulation, motivation, and performance (Eckerlein et al., 2019). Our prototype uses ITS data alongside daily diaries with pre- and post-practice reflections, referencing ITS-tracked student activities adaptively. Multiple-choice and text-input fields prompt for self-assessment on learning approaches, effort, and motivation. ***Goal Setting*** strengthens self-regulation by setting practice objectives, implementing strategies, and monitoring progress (Zimmerman, 2002). Our prototype uses real-time data on student skill mastery, allowing students to select math skills to improve, setting accuracy and learning opportunity goals. It provides feedback on accuracy and error rates, tracks problem completion and time, and recommends personalized goals based on actual performance. ***Homework Graphing*** lets students visually monitor assignments, encouraging homework completion and student independence (Bryan & Burstein, 1998). Our prototype uses assignment-specific student log data for tracking, featuring pie charts, bar graphs, and line graphs, contrasting system-generated charts against student-inputted data in bar and line graphs. Metrics included problem completion count and minutes spent on homework, allowing students to compare achievements against goals. ***Reminders,*** typically sent via email or apps, boost student homework engagement (Zavaleta Bernuy et al., 2021). Our prototype incorporates ITS learning metrics into reminder curation for contextual relevance, offering customizable reminders, with options for frequency, date, time, and duration linked to diary entries, goals, and graphs, including pop-up alerts.

Phase Two, informed by Phase One, further explored students' goal-setting preferences in interviews, including goal types, inclination towards self-set or system-suggested goals, and goal-setting routines with an enhanced prototype that showed performance metrics, offered system-generated goals, and displayed visual data on skill accuracy and time spent. We examined students' perceptions of goal recommendations based on perceived efficacy, preferences for skill- versus time-based goals, and reactions and subsequent approaches upon either achieving or not achieving goals. These investigations were motivated by prior findings that low self-regulators often set non-specific goals and are dissatisfied with their performance (Bembenutty, 2011). Follow-up questions delved into reasons for accepting or rejecting recommended goals, opinions on feedback about goal achievement and difficulty, and views on autonomy and flexibility in goal setting and modification.

Results and findings

In the final ranking activity of Phase One, goal setting emerged as the top choice among six out of nine participants. Four participants independently advocated for integrating goal setting with homework graphing to compare goals with actual achievements and with reminders for goal-related tasks.

Goal Setting. Students valued tangible metrics for self-set goals, *[P8] "Combining accuracy with attempts or time helps me understand what this percentage means for my progress."* Notably, all nine students preferred system-generated goal recommendations over self-formulated goals, seeing them as reliable and attainable. *[P3] "I'd love to have recommendations when I don't want to push myself too hard." [P7] "I'm not*

creative with goals and would actually prefer system recommendations for better, more detailed goal planning." ***Standardized Diaries.*** Six out of nine participants were satisfied with the assessment statements and diary questions. While some appreciated the reflective aspect, others raised practical concerns, indicating a need for balancing depth of reflection with practicality for on-the-go use: *[P7] "Reflections aren't as effective on mobile, especially for math."* ***Homework Graphing.*** Three participants preferred system auto-generated graphs for efficiency, four favored graphs derived from data input by students, and two chose to draw their own. *[P8] "I find self-graphing challenging, but don't want the system to do everything. I prefer giving inputs for graph generation."* All but one participant desired graphs that compare their actual achievements and progress with set goals. ***Reminders.*** Most students appreciated prompts for initiating homework practices, exam preparation, and unmastered skills. Seven out of nine participants preferred the system to auto-set reminders based on tracking their past practices but valued the ability to make adjustments: *[P3] "The system will be better at suggesting since I am not really good at finding reminder times."* However, four participants expressed skepticism, noting that reminders could be easily ignored, indicating the need to integrate them into students' practice routines.

Phase Two generated three major insights. **Insight 1: Students naturally articulate homework goals by problem count and study time duration.** While five out of nine participants favored skill-based goals over time-based goals, students generally required clarity on the number of problems needed to gauge the effort and reach mastery of skills: *[P13] "I like to see how many more questions I need to do to improve on that certain skill;"* and time spent for orienting toward skill mastery: *[P12] "prefer goal 3 as it tells me clearly that I have 12.2% now, and I can do 20 minutes a day to get higher accuracy."* When asked to relate these practices to skill improvement, there was a unanimous preference for focusing on one skill at a time: *[P12] "I prefer sticking with one skill as the system often provides better time estimates for mastering it, and the system can suggest the top three essential skills to focus on."* **Insight 2: Students prefer system-set goals, particularly when self-set goals are perceived as inadequate.** All nine students were positive about system-generated recommendations, citing three distinct reasons. First, insecurities about interpreting system feedback on learning progress led only four students to attempt setting their own goals after seeing their performance metrics (e.g., % correct answers, time spent, and opportunity missed). [*P11] "After seeing performance, I don't really know how to set up my goals but want to see system recommendations."* Second, eight students perceived system-recommended goals to be more effective than self-set goals, attributing this effectiveness to the system's ability to track progress and make accurate predictions for areas of improvement. [*P14] noted: "The system knows how to calculate your levels, but if you are making your own goals, you need to experiment multiple times"* while [*P15] noted: "It has a good idea of the amount of time I spend on practicing math without making the goal too easy or too hard."* Third, seven students indicated they would switch to using system recommendations if they did not meet their own goals. The rationale, shared by *P14*, *P16*, *P17*, and *P18*, was similar: *"If I did not meet my own goal, it may be because I set an unreasonable goal. The system can probably make a better calculation for me."* Notably, all students described misalignment between system-recommended goals and their perceived mastery or schedule as potential reasons for not following recommendations, highlighting a need to adapt recommendations to students' existing homework routines. **Insight 3: Students appreciated feedback on self-set goal difficulty level which could foster autonomy.** All nine participants preferred designs featuring system alerts about goal difficulty levels and progress tracking. They valued system-set goals for reliability, especially when personal goals fall short. *[P14] "I would like the system to give me an option to generate new goals if current ones don't work."* Participants also sought comparisons between personal and system-set goals to calibrate their adequacy. *[P14] "Even if I set up my own goals... want system feedback to see if I am behind or ahead now."* P16 emphasized the need for the system to be responsive to their achievements: *"The system should respond fast for my needs, if I finish the goal the system should know how to verify it and give other effective goals in time."* Finally, students appreciated difficulty feedback on self-set goals to fine-tune their time management and effort.

Discussion and conclusion

Our study explores an underexplored design space: Data-driven digital homework tools that integrate intelligent tutoring systems and evidence-based homework support structures in middle school mathematics. Through interactive student interviews, we identified goal-setting as the most preferred direction out of four explored options. Student perceptions of a goal-setting prototype that balances proactively set and data-driven, recommended goals highlighted a high level of initial reliance and comfort with system-set goals while appreciating the opportunity to adjust goal difficulty. This balance presents an opportunity for fostering self-regulated learning competencies by fading system-level support in goal setting, encouraging students to independently set and adjust practice goals. Independent goals are desirable as they have been linked to self-regulated learning in ITSs (Long & Aleven, 2013). Real-time feedback on goal achievement and difficulty could present learning opportunities for improved self-regulated learning. Students valued such feedback to

regulate their practice effort and time management. Overall, real-time feedback on goal difficulty while allowing learners to adjust goals could foster self-regulatory skills crucial for homework success. Fine-grain learner data from ITSs could inform such feedback while providing a platform for learning-oriented goal selection mastery.

Key design recommendations for mobile-centered learning applications include students appreciating tangible metrics related to the expected number of problems and study duration to achieve skill-related mastery goals during deliberate practice (Zimmerman, 2002). An open question is whether students can, or should, eventually graduate from these tangible metrics to skill mastery-related goals only. Evidence suggests that praising and expressing effort-focused goals (e.g., time spent) may already strengthen mastery goal orientation (Droe, 2013). Furthermore, students preferred practicing one skill at a time, seemingly in contrast to literature that suggests interleaving practice is beneficial for learning (Taylor & Rohrer, 2010). To honor students' preferences while retaining the potential benefits of interleaving, future goal-setting designs could deliver skill switch recommendations during practice or multi-skill practice problem sets that students can overwrite or customize, similar to the goal difficulty recommendation regime students favored.

Considering limitations, participants engaged with mobile interfaces via computer screen shares rather than on mobile devices, preserving crucial tactile elements that may have affected perceptions of the prototypes. Moreover, the studied sample lacked racial and ethnic diversity, being predominantly Caucasian and East Asian. Future design iterations should consider diverse students' learning experiences. Lastly, students' expressed needs and actual use could differ in practice. Still, the current study has surfaced important considerations for digital homework support tool design to be further studied through naturalistic observations.

Endnotes

(1) Prototypes modeled after Lynnette, an equation-solving tutor, part of the Mathtutor ITS: https://mathtutor.web.cmu.edu

References

Bembenutty, H. (2011). Meaningful and maladaptive homework practices: The role of self-efficacy and self-regulation. *Journal of Advanced Academics*, *22*(3), 448-473.

Bryan, T., & Burstein, K. S. (1998). Teacher-selected strategies for improving homework completion. *Remedial and Special Education*, *19*(5), 263-275.

Droe, K. L. (2013). Effect of verbal praise on achievement goal orientation, motivation, and performance attribution. *Journal of Music Teacher Education*, *23*(1), 63-78.

Eckerlein, N., Roth, A., Engelschalk, T., Steuer, G., Schmitz, B., & Dresel, M. (2019). The role of motivational regulation in exam preparation: Results from a standardized diary study. *Frontiers in Psychology*, *10*, 81.

Feng, M., Heffernan, N., Collins, K., Heffernan, C., & Murphy, R. F. (2023). Implementing and evaluating ASSISTments online math homework support at large scale over two years: Findings and lessons learned. *International Conference on Artificial Intelligence in Education* (pp. 28-40).

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research, 86*(1), 42–78.

Long, Y., Aleven, V. (2013) Skill diaries: Improve student learning in an intelligent tutoring system with periodic self-assessment. *International Conference on Artificial Intelligence in Education* (pp. 249-258).

Mrazek, A. J., Mrazek, M. D., Ortega, J. R., Ji, R. R., Karimi, S. S., Brown, C. S., Alexander, C. A., Khan, M., Panahi, R., Sadoff, M., Scott, A., Tyszka, J. E., & Schooler, J. W. (2021). Teenagers' smartphone use during homework: An analysis of beliefs and behaviors around digital multitasking. *Education Sciences*, 11(11), 713.

Taylor, K., & Rohrer, D. (2010). The effects of interleaved practice. *Applied cognitive psychology*, *24*(6), 837-848.

Zavaleta Bernuy, A., Zheng, Q. Y., Shaikh, H., Petersen, A., & Williams, J. J. (2021). Investigating the impact of online homework reminders using randomized A/B comparisons. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (pp. 921-927).

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64-72.

Acknowledgements

This research was funded by the Institute of Education Sciences (IES) of the U.S. Department of Education (Award #R305A220386).