

The Challenge of Challenges: Examining the Impact of Difficulty Dynamics on Mastery Learning in Math

Seiyon M. Lee
University of Florida
leeseiyon@ufl.edu

Anthony F. Botelho
University of Florida
abotelho@coe.ufl.edu

ABSTRACT

Adaptive learning systems promote mastery learning by allowing students to hone and harness the foundational skills through a number of problem-solving opportunities. To properly scaffold this learning process, it is instrumental that the problems are structured and sequenced in ways that are not too difficult to be discouraging but also not too easy for students to unchallenging. While considerable work focused on improving the cognitive models to inform the design of adaptive scaffolding and other interventions, it remains unclear how the differences in problem-level difficulty, whether they are negligible or dramatic, shape students' learning and their engagement within the adaptive learning systems. In this work, we aim to explore the effect of the problem sequence on students' learning processes and outcomes. We concluded with a discussion on the opportunities and challenges of our work to inform the design and development of adaptive learning systems in ways that are more responsive to individual problem-solving experience.

Keywords

Adaptive Learning Systems, Mastery Learning, Problem Sequence, Engagement

1. INTRODUCTION

To achieve mastery in a skill, it is critical for a student to engage with the right amount of both opportunities and challenges. While students need opportunities to apply what they learned in earlier problems to solidify their conceptual understanding, they also need challenges to continuously refine and broaden their zone of proximal development [10]. In recent years, a growing number of adaptive learning systems (i.e., ASSISTments, Mathia, and others) promote mastery learning by providing students with a large pool of practice problems until they have mastered the skill. Coupled with adaptive support, such as hints and feedback, these systems have shown considerable success in improving learning in classrooms as well as other informal settings [1, 5].

S. M. Lee and A. F. Botelho. The challenge of challenges: Examining the impact of difficulty dynamics on mastery learning in math. In B. Paaßen and C. D. Epp, editors, *Proceedings of the 17th International Conference on Educational Data Mining*, pages 994–997, Atlanta, Georgia, USA, July 2024. International Educational Data Mining Society.

© 2024 Copyright is held by the author(s). This work is distributed under the Creative Commons Attribution NonCommercial NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license.
<https://doi.org/10.5281/zenodo.12730027>

An important design feature for successful mastery learning is that students' problem-solving processes are scaffolded. Some systems allow students to adjust the problem-level difficulty by accessing on-demand hints and other forms of scaffolding as they deem necessary [8]. While this affordance empowers students to exercise agency in their learning, it is not without limitations. First, students' use of hints often informs the system designers whether or not they have used it too quickly or too fastidiously, as evidenced in the development of gaming detectors [7]. It also puts a strain on students with the assumptions that (a) they have the capacity to use the optional hints only when it is timely and appropriate, and (b) their hint usage can impact on their self-efficacy. In light of the challenges associated with self-selected scaffolding, it is critical for adaptive learning systems to have the capacity to strategically sequence the problem difficulty.

In this study, we explore the effect of difficulty dynamics on learning in the context of a web-based learning environment to promote mastery in math. Toward this goal, we examine the different types of difficulty dynamics and conduct a regression analysis to examine the effect of varying sequences of problem difficulty on learning outcomes. As students were randomly assigned problems with varying levels of difficulty, there was considerable variance in terms of difficulty dynamics. Therefore, we focused our attention to the first three problems students encountered, whether they gradually escalated in difficulty or not. We discuss the preliminary findings, and share some opportunities and challenges in our work to open the floor to any well-informed suggestions with the larger goal of better understanding how difficulty dynamics shape learning processes and outcomes and how it can inform the design of adaptive learning systems so that the mastery learning experience is made optimal for all learners.

2. CONTEXT

This exploratory study uses the log data of ASSISTments, an online math homework platform for Grades 3 through 12 [4]. Launched in the early 2000s, ASSISTments blends the notion of formative assessment with assistance (e.g. hint). With the goal of promoting mastery, ASSISTments continually assigns students with new skill builders problems until they demonstrate mastery. The randomized problems lend themselves to natural randomized experiments at the problem-level that we can leverage to study effects of difficulty dynamics. It is also worth noting that ASSISTments does not involve active recruitment of its users but

students are self-selected by choosing to use the system for their needs. Due to this non-random selection method, it is possible that the dataset may consist of students who are inherently motivated to seek out and use online resources. Despite the limitation, the dataset holds significance within educational research due to its substantial user base and the scale of data available. This is evidenced by the extensive work around ASSISTments for randomized controlled experiments, providing robust evidence of its efficacy in contributing to educational research [9, 6, 3].

3. DATA

The dataset includes data from a sample of 759 students in the U.S. who used the ASSISTments during the 2016-2019 academic years. These students are primarily from the northeastern part of the U.S., where ASSISTments was initially developed and distributed. We focus specifically on students' interactions with the Skill Builder assignments where they tackle a number of different problems on a targeted skill until they demonstrate mastery (or get a streak of three correct answers). During this process, students not only get immediate feedback on their answers, but they can also request hints on a limited number of available problems to break down their own problem-solving processes into smaller steps.

While the entire dataset consists of several different action-level features (e.g. the number of hints and attempts, and correctness), we focus on difficulty dynamics and learning outcome. Specifically, we extract the problem difficulty of the first set of three math problems that students encounter, which is informed by the empirical evidence that significant number of students quit and stop out within the first three opportunities [2]. The predictors are the problem difficulties of the first, second, and third problems, as well as whether the level of difficulty has increased or decreased in the next problem. In the model, we also include a measure of a student's ability (or prior knowledge) with the consideration that students who come with greater knowledge and experience on a specific skill would be less impacted by getting a series of challenge problems. The problem difficulty and student ability parameters have been computed using the Rasch model, which is widely used for item response theory (IRT).

For the measure of mastery learning, we use a delayed learning outcome collected through the Automatic Reassessment and Relearning System (ARRS) [11] where teachers reassess their students with a demonstrated mastery on a skill about a week after they learned it in the ASSISTments. Some students engaged with more than one reassessment problem. One student received as many as 38 problems for reassessment. However, we limited the outcome variable to whether or not they got the correct answer at the very first problem, considering that most students ($N=250$) received only one reassessment problem. This would also be the fairer approach than measuring the average correctness, which can bias against students who receive more problems.

4. METHODS

To explore the relationship between patterns of problem difficulty and learning, we identify the different dynamics of problem difficulty. Out of 7,410 problems, we first com-

puted the problem-level difficulty as the percentage of students who found the correct answer to the problem, where the higher value of difficulty indicates that the problem was successfully solved by a larger number of students. Within our sample, we labeled three groups of problems based on their levels of difficulty as Easy, Medium, and Hard. If the percentage of the students getting it correct was higher than 90%, then the problem was considered to be easy. When the percentage was lower than 60%, it was considered as a problem that is hard for students to solve.

We then define the sequence of the problem difficulty for each student. In our sample, some students were able to achieve mastery quicker than others by succeeding in getting a streak of three correct answers. A total of 179 students demonstrated mastery with flying colors through this fast track. In a more dynamic case, some students engaged with more than 100 problems until they could master the skill. As students were randomly assigned to the problems with varying levels of difficulty, there was significant variation in how problems were presented to the students in terms of difficulty. In this study, we focused on how the first three problems were sequenced, whether they gradually escalated in difficulty (i.e., Scaffold) or inadvertently posed a difficult problem to the students (i.e., Challenge).

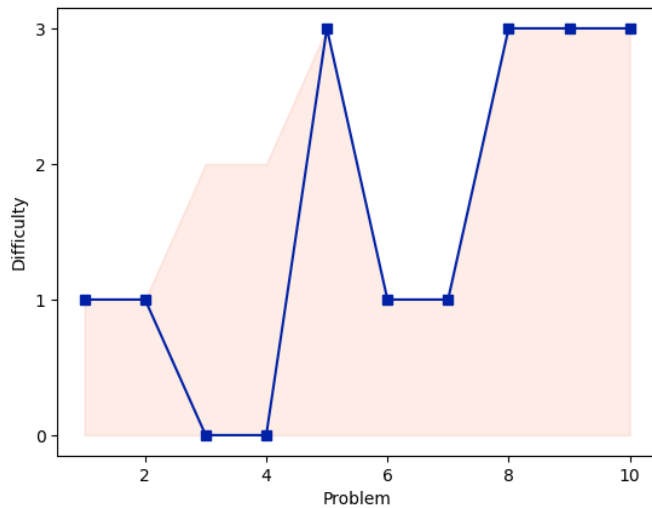
As the second part of the analysis, we conducted a logistic regression, as is appropriate for binary outcome variables like ours (whether they get the correct answer or not on the first problem of the reassessment). Specifically we focused on how a student's ability and the varying levels of difficulty within the first three problems interact and jointly affect the probability of reassessment correctness.

5. PRELIMINARY RESULTS

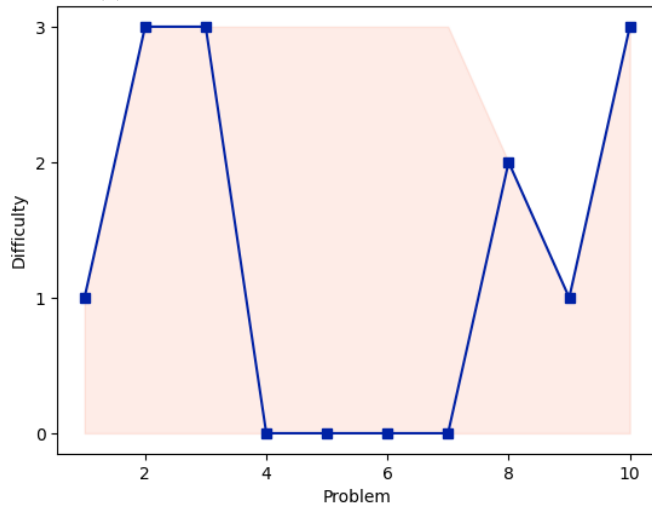
In this paper, we report on the preliminary findings on how problem sequence influences students' learning. We first present diverse graphical representations of different problem sequences that were exposed to the students, and provide the student's narrative around their challenges they accept to achieve mastery in math. Figure 1 shows a case where problems by random chance have been scaffolded by engaging students from easy, medium to hard problems in a step-by-step fashion. As this figure shows, students can struggle to solve the problem with an increased difficulty, but if the system adjusts the difficulty accordingly, then students can check their conceptual understanding on a concept, which are prerequisite for more advanced problems.

Figure 2 shows an interesting case where the student was challenged to solve a hard problem after solving an easy problem, and managed to achieve mastery after several incorrect responses. The question remains how to assess the skills tested on the first two hard questions that students got right and the subsequent series of hard problems students did not get right. Similarly, more data is needed to confirm that students have successfully mastered skill related to the medium-level problem the student solved considering that the student struggled in harder problems in Figure 3.

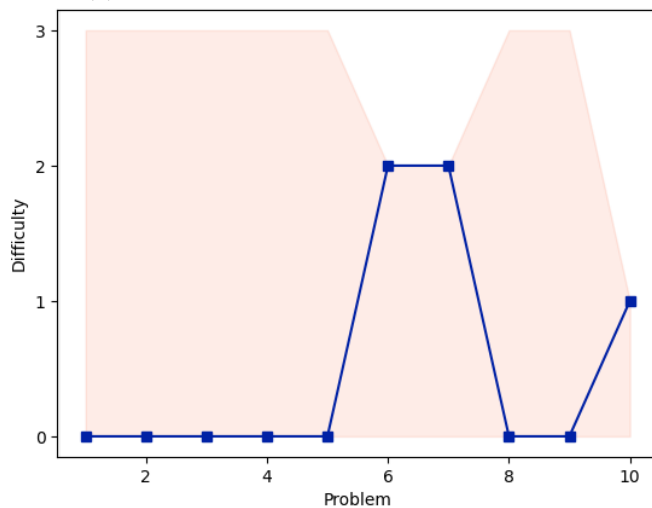
The regression was conducted to examine the effect of difficulty dynamics (i.e., difficulty of the first problem, second problem, and the third problem) as well as student ability



(a) Student Achieving Mastery through Scaffold



(b) Student Achieving Mastery despite Challenges



(c) Student Challenged by Challenging Problems

Figure 1: Examples of different types of difficulty dynamics

on learning outcome. Among other variables and their interaction, the significant predictors of reassessment correctness were identified as student ability with the coefficient of 0.26 ($SE = 0.085$, $p < 0.01$) and the problem difficulty of the third problem the coefficient of 0.14 ($SE = 0.059$, $p < 0.01$). This finding indicates that the higher level of ability the student initially has and the greater difficulty of third problem the student encounters are associated with increased log-odds of success, while controlling for other factors. Interestingly, the main effects of problem difficulty for the first and second problems as well as their interactions were not statistically significant. The intercept estimate of 0.61 ($SE = 0.12$, $p < 0.001$) also suggests that other variables may be at play. The Akaike Information Criterion (AIC) value of 887.49 further supports the reasonable fit of the model.

This study investigates the effect of difficulty dynamics on mastery learning using log data from the ASSISTments platform. The logistic regression model reveals that student ability and the difficulty of the third problem are significant predictors of mastery learning, with higher-ability students more likely to perform well on reassessments. These findings have important implications for the design and implementation of adaptive learning technologies. The significance of the student ability and third problem difficulty suggests that these systems should carefully consider the differences in starting point across students and ensure that the third challenge is well-matched to the student’s prior knowledge and skills. Doing so would establish foundation for success that capitalizes on students’ abilities and sets the stage for long-term mastery.

However, the lack of significant effects for subsequent problem difficulties raises questions about the optimal sequencing and the factors that contribute to successful learning outcomes. It is possible that factors such as student motivation, engagement, and the specific skills targeted by each problem play a more substantial role in determining long-term retention than the dynamics of problem difficulty alone. This highlights the need for a more comprehensive approach to designing adaptive learning systems that takes into account a wide range of variables and their interactions.

6. CHALLENGES

From our preliminary findings, we present a number of challenges toward our broader goal of promoting a better understanding of how problem sequences influence student learning as well as engagement. First, we can consider the ways in which the problem-level difficulty can be defined. There are different approaches to measuring the difficulty of a problem. In this work, specifically for the first part of the analysis, we used the percentage of correctness scored by the students who tackled the same problem. This, however, is limited in that it comes from a selective group of students who not only demonstrated mastery, and that it does not extract difficulty from the item itself. It is worth also noting that the difficulty of a problem can vary by students depending on what aspects of the problem primes students in ways that help or hinder their problem-solving capacity. While we use the Rasch model for the student ability and difficulty parameters, the simplicity of this approach can be reinforced with other supplementary measures of difficulty.

7. OPPORTUNITIES

Beyond the limited scope and space of this work, our work can be taken up in some interesting directions. One is to expand the student profiles, including the students who successfully completed the skill-builder assignments. The students in this dataset represent the group of students who have successfully mastered a skill, or found correct answers to three problems in a row. While these students offer relatively richer interaction data as well as the learning outcome measures, the story would not be complete if we were not to consider the cases where students stop out due to the problems that are challenging them beyond their comfort levels. More diverse groups of students can be explored beyond course completion, such as student profiles of prior knowledge as well as of affect. Examining a diverse range of students who opt out of the system, despite the challenge of working with limited data, is warranted to advance a nuanced understanding of the challenges students often encounter during their learning experience.

Another important and related strand of consideration is the complex nature of learning that goes beyond cognitive and metacognitive components. To have a holistic understanding of how problem sequences influences learning outcomes engagement, it requires a number of different factors at play. For example, students can either thrive or feel continually challenged as problems escalate in difficulty, but they can also be affected by their own performance as they monitor their own progress. From this perspective, it is reasonable to think that students may not start, if not maintain, their full capacity in the face of challenges. Therefore, it would be interesting to examine how student engagement, affect, and motivation interacts with problem difficulty and jointly influence learning processes and outcomes.

8. CONCLUSION

This short paper showcases initial efforts to better understand how difficulty dynamics of problems impact students' mastery learning experience. The preliminary findings call for more steps as follow-up, including but not limited to: (1) a better definition of a problem sequence, and (2) a richer consideration of the contextual factors at play. While more work is warranted, we believe that greater interests to explore the construct of difficulty from wider angles (i.e., the perceived difficulty that goes beyond students' capacity and comfort levels) would inform the design of adaptive learning systems that enhance learning and empower learners.

9. REFERENCES

- [1] J. E. Beck, K.-m. Chang, J. Mostow, and A. Corbett. Does help help? introducing the bayesian evaluation and assessment methodology. In *Intelligent Tutoring Systems: 9th International Conference, ITS 2008, Montreal, Canada, June 23-27, 2008 Proceedings 9*, pages 383–394. Springer, 2008.
- [2] A. F. Botelho, A. Varatharaj, E. G. V. Inwegen, and N. T. Heffernan. Refusing to try: Characterizing early stopout on student assignments. In *Proceedings of the 9th international conference on learning analytics & knowledge*, pages 391–400, 2019.
- [3] M. Feng, C. Huang, and K. Collins. Promising long term effects of assistments online math homework support. In *International Conference on Artificial Intelligence in Education*, pages 212–217. Springer, 2023.
- [4] N. T. Heffernan and C. L. Heffernan. The assistments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education*, 24:470–497, 2014.
- [5] A. Munshi, G. Biswas, R. Baker, J. Ocumpaugh, S. Hutt, and L. Paquette. Analysing adaptive scaffolds that help students develop self-regulated learning behaviours. *Journal of Computer Assisted Learning*, 39(2):351–368, 2023.
- [6] R. Murphy, J. Roschelle, M. Feng, and C. A. Mason. Investigating efficacy, moderators and mediators for an online mathematics homework intervention. *Journal of Research on Educational Effectiveness*, 13(2):235–270, 2020.
- [7] L. Paquette, A. de Carvahlo, R. Baker, and J. Ocumpaugh. Reengineering the feature distillation process: A case study in detection of gaming the system. In *Educational data mining 2014*, 2014.
- [8] L. Razzaq and N. T. Heffernan. Scaffolding vs. hints in the assistment system. In *International Conference on Intelligent Tutoring Systems*, pages 635–644. Springer, 2006.
- [9] J. Roschelle, M. Feng, R. F. Murphy, and C. A. Mason. Online mathematics homework increases student achievement. *AERA open*, 2(4):2332858416673968, 2016.
- [10] L. S. Vygotsky and M. Cole. *Mind in society: Development of higher psychological processes*. Harvard university press, 1978.
- [11] X. Xiong, S. Li, and J. E. Beck. Will you get it right next week: Predict delayed performance in enhanced its mastery cycle. In *The Twenty-Sixth International FLAIRS Conference*, 2013.