



# A General Multi-method Approach to Design-Loop Adaptivity in Intelligent Tutoring Systems

Yun Huang<sup>(✉)</sup>, Vincent Aleven, Elizabeth McLaughlin,  
and Kenneth Koedinger

Carnegie Mellon University, Pittsburgh, PA 15213, USA  
{yunhuanghci, koedinger}@cmu.edu,  
{aleven, mimim}@cs.cmu.edu

**Abstract.** Design-loop adaptivity, which involves data-driven redesign of an instructional system based on student learning data, has shown promise in improving student learning. We present a general, systematic approach that combines new and existing data mining and instructional design methods to redesign intelligent tutors. Our approach is driven by the main goal of identifying knowledge components that are demonstrably difficult for students to learn and to optimize effective and efficient practice of them. We applied this approach to redesigning an algebraic symbolization tutor. Our classroom study with 76 high school freshmen shows that, compared to the original tutor, the redesigned tutor led to higher learning efficiency on more difficult skills, higher learning gain on unscaffolded whole tasks, and more robust transfer to less practiced tasks. Our work provides general guidance for performing design-loop adaptations for continuous improvement of intelligent tutors.

**Keywords:** Instructional design · Adaptivity · Data mining · Intelligent tutor

## 1 The Need for a General Data-Driven Redesign Approach

*Design-loop adaptivity* [1] uses student learning data to drive instructional decision making for design and iterative improvement of a course or system. It is part of a broader set of endeavors of data-driven instruction and learning designs for continuous improvement in classrooms and schools [2–5]. This paper focuses on the context of *intelligent tutoring systems* (ITSs), a widely adopted and proven technology, where empirical research on data-driven redesign and optimization is still lacking. Numerous data mining methods have been demonstrated to improve prediction accuracy using data from fielded ITSs [6–8], but most stop at better predictions without demonstrating whether and how these methods can improve student learning. One reason for a shortage of such “close the loop” experiments may be that there is no good general guidance for how to convert data-mining outcomes into better tutor design. Prior close-the-loop studies [9–11] were often driven by a limited set of methods or narrow redesign features. This paper demonstrates a general, systematic approach that combines new and existing data mining and instructional design methods to redesign ITSs.

We applied this approach to redesign an algebraic symbolization tutor, and provide empirical evidence of its effectiveness through a classroom study comparing the redesigned tutor to the original tutor.

2 Method: How to Use Data to Improve Intelligent Tutors

Our approach focuses on the continuous improvement of ITSs by mining tutor log data collected from previous iterations. It starts with knowledge component (KC) refinement, followed by content and task selection revisions aiming at more effective and efficient practice of KCs. Our process (Table 1) is generalizable to other domains and other ITSs grounded in a KC approach [12] to instructional design.

Table 1. A general multi-method approach to data-driven redesign of ITSs.

Goals	Methods
1 Refine the knowledge component (KC) model	
Identify difficulty factors to split KCs	Difficulty Factor Effect Analysis
Compare hypothesized KC models	AFM prediction and inspection [13]
2 Redesign content	
Estimate opportunities to mastery, under- and over-practice for each KC in the refined model	Probability-Propagation Practice Estimation
Create focused tasks for difficult KCs with better scaffolding and reduce over-practicing easier KCs	Focused Practice Task Design (with dynamic, composition scaffolding)
Add feedback messages to frequent errors	Error analysis [14, 15]
3 Optimize individualized learning	
Optimize student model parameters	Data-tuning BKT parameters [16]
Optimize task selection based on a student model	Task selection simulation [17]

We applied this process to redesigning the *Algebraic Expressions* unit in Mathtutor [18], a free online tutor based on prior instructional design research [19] (Fig. 1). We utilized prior log data from 356 students with 50,279 student steps. We describe our new methods below, and refer readers to prior work for existing methods.

**Difficulty Factor Effect Analysis.** A difficulty factor (DF) refers to a property that makes some tasks more difficult than other comparable tasks. We first identified a broad set of potential DFs by coding task features hypothesized to impact difficulty (e.g., requiring parentheses or not). Then we ran a regression for each targeted KC to examine the main and interaction effects of potential DFs on performance, controlling for student proficiencies and learning from prior opportunities. These regressions might be viewed as an efficient simplification of LFA [20]. A KC was split by a set of DFs when there was an interaction or by a DF when there was a main effect.

**Probability-Propagation Practice Estimation.** We estimated the number of opportunities needed for mastery by fitting parameters of a student model (e.g., BKT [21]) to the data and used the parameters to estimate knowledge for each step. We then compared the estimates to actual opportunities to get the over- or under-practice. Instead of

simulating many sequences by propagating simulated outcomes [22], we simulate one sequence by propagating the *probability* of succeeding, and use it as weights to update knowledge (i.e.,  $P(L)_{new} = P(C)P(L|C)_{new} + P(W)P(L|W)_{new}$ ). The extrapolation of a KC-student sequence stops when  $P(L)_{new} \geq .95$ , or the extrapolated opportunities reach a threshold (e.g., 20). Our offline experiments showed that our method reached similar estimations as the method in [22] with higher efficiency.

**Focused Practice Task Design.** We created new focused problems for hard KCs which eliminate interface steps of easier KCs (Fig. 2). These problems aim to reduce under-practicing hard KCs and over-practicing easier KCs, which were prevalent in the original tutor according to our practice estimation. We introduced *composition scaffolding* (inspired by prior work [14, 23, 24]) to break down problems (Fig. 2a), because our method estimated that many opportunities ( $\geq 58$ ) would be needed to master each difficult KC (two-operator KCs) in the original tutor.

Brady's Little League team is ranked first in the city, with a total of 25 points. For every game they win, the team will get another 5 points for their ranking.		The number of additional wins	The total number of points the team has	Show your work	
(1) If Brady's team wins another 3 games, how many total points will they have?	Unit	wins	points	--	
(2) If Brady's team wins another 8 games, how many total points will they have?		1	3	40	$5 \cdot 3 + 25$
In the row labeled Expression, define a variable for the number of additional wins and use that variable to write an expression for the total number of points the team has.		2	8	65	$5 \cdot 8 + 25$
	Expression	x	--	$5x + 25$	

Fig. 1. A table task in the original tutor (with cells filled in correctly and the toolbar excluded).

Brady's Little League team is ranked first in the city, with a total of 25 points. For every game they win, the team will get another 5 points for their ranking.

Write an expression for the total number of points the team has given that the number of additional wins is t.

Let's break it down!

(1) Let's solve a smaller problem. Brady's team had t wins and got 5 points per win. Write an expression for the number of points the team got by t wins.

(2) Let's solve another smaller problem. Brady's team had 25 ranking points originally and it earned x more points. Write an expression for the total number of points the team has.

(3) Let's solve the original problem by putting (1) and (2) together. Substitute 5\*t for x in x+25. Write the resulting expression.

a)

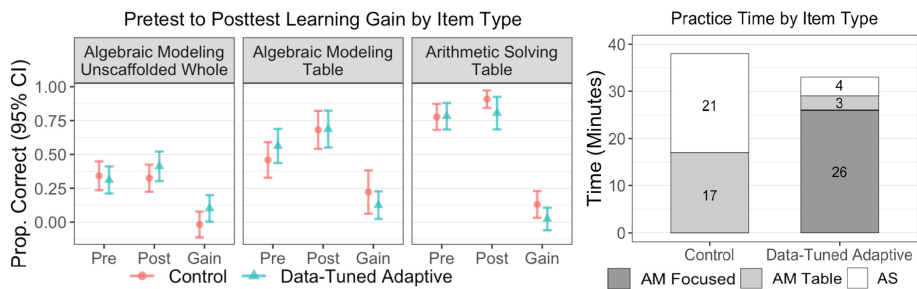
Substitute 10b for x in 600-x. Write the resulting expression.

b)

Fig. 2. Focused tasks target a KC that data reveals is particularly hard. a) A *focused whole task* practices this KC in more realistic problems that require mental steps of easier KCs along with the hard KC. If students fail on the whole task, dynamic composition scaffolding isolates the individual KCs (step 1–3) including the hard one (step 3). b) A *focused part task* practices the hard KC in isolation without the busy work (mental and interface steps) of the easier KCs.

3 Experiment

We conducted a classroom study to investigate whether the data-driven redesigned tutor (treatment condition) yields better learning than the original tutor (control condition). We ran the study in two high school freshman Algebra I classes during three 40-min periods for one week in 2019. Students were randomly assigned to conditions within each class, with 38 students per condition. We used linear mixed models to examine learning gains and two-sample t-tests to examine practice time (Fig. 3).



**Fig. 3.** Redesigned tutor showed advantages in targeted algebraic modeling (AM) skills in terms of learning gains on unscaffolded whole tasks and learning efficiency on table tasks.

Overall, both conditions produced significant learning gains ( $ps < .01$ ). Students in the treatment condition spent 13% less practice time than students in the control condition ( $p < .1$ ), with no difference in gains ( $p = .94$ ). In particular, on more difficult skills, two-operator skills, they spent 19% less time ( $p < .05$ ) with no difference in gains ( $p = .78$ ). Treatment students had significant gains on algebraic modeling (AM) in both unscaffolded whole tasks (Fig. 2a without scaffolding steps) and table tasks (Fig. 1) ( $ps < .05$ ) even with little practice on table tasks. Control students had a significant gain on AM table tasks ( $p < .001$ ) with no difference from treatment students ( $p = .14$ ), and no improvement on unscaffolded whole tasks ( $p = .58$ ) with the gain different from that of treatment students ( $p < .05$ ). These results suggest that treatment students acquired more robust, transferable learning. Control students had marginally higher gains on arithmetic solving (AS) ( $p = .096$ ); the treatment condition was designed to shift practice away from these (easier) skills to the harder AM skills.

4 Discussion and Conclusion

We demonstrate a general multi-method approach to design-loop adaptivity and provide empirical evidence of its effectiveness. The results are encouraging, although they are not as pronounced as anticipated. Students spent much less time on the tutor than planned, but our theoretical predictions were based on longer time. Thus, a more stringent test of our approach requires a future longer span study. After all, design-loop adaptivity is intended as an iterative process. Our work provides general guidance for

how to convert data-mining outcomes into better tutor design, an important need in AIED/EDM research and practice. Our work may also help define and enhance data-driven learning engineering processes.<sup>1</sup>

## References

1. Aleven, V., McLaughlin, E.A., Glenn, R.A., Koedinger, K.R.: Instruction based on adaptive learning technologies. In: *Handbook of Research on Learning and Instruction*, 2nd edn., pp. 522–560. Routledge, New York (2017)
2. Halverson, R., Grigg, J., Prichett, R., Thomas, C.: The new instructional leadership: creating data-driven instructional systems in school. *J. School Leadership* **17**(2), 159–194 (2007)
3. Lockyer, L., Heathcote, E., Dawson, S.: Informing pedagogical action: aligning learning analytics with learning design. *Am. Behav. Sci.* **57**(10), 1439–1459 (2013)
4. Bodily, R., Nyland, R., Wiley, D.: The RISE framework: using learning analytics to automatically identify open educational resources for continuous improvement. *Int. Rev. Res. Open Distrib. Learn.* **18**(2), 103–122 (2017)
5. Kaufman, T.E., Graham, C.R., Picciano, A.G., Popham, J.A., Wiley, D.: Data-driven decision making in the K-12 classroom. In: *Handbook of Research on Educational Communications and Technology*, pp. 337–346. Springer, New York (2014)
6. Piech, C., et al.: Deep knowledge tracing. In: *Advances in Neural Information Processing Systems*, pp. 505–513 (2015)
7. Lindsey, R.V., Khajah, M., Mozer, M.C.: Automatic discovery of cognitive skills to improve the prediction of student learning. In: *Advances in Neural Information Processing Systems*, pp. 1386–1394 (2014)
8. González-Brenes, J., Huang, Y., Brusilovsky, P.: General features in knowledge tracing to model multiple subskills, temporal item response theory, and expert knowledge. In: *The 7th International Conference on Educational Data Mining*, pp. 84–91 (2014)
9. Lovett, M., Meyer, O., Thille, C.: The Open Learning Initiative: Measuring the effectiveness of the OLI statistics course in accelerating student learning. *J. Interactive Media Educ.* **2008** (1) (2008). <https://doi.org/10.5334/2008-14>. Article no. 13
10. Koedinger, K.R., Stamper, J.C., McLaughlin, E.A., Nixon, T.: Using data-driven discovery of better student models to improve student learning. In: Lane, H.C., Yacef, K., Mostow, J., Pavlik, P. (eds.) *AIED 2013. LNCS (LNAI)*, vol. 7926, pp. 421–430. Springer, Heidelberg (2013). [https://doi.org/10.1007/978-3-642-39112-5\\_43](https://doi.org/10.1007/978-3-642-39112-5_43)
11. Mostafavi, B., Barnes, T.: Evolution of an intelligent deductive logic tutor using data-driven elements. *Int. J. Artif. Intell. Educ.* **27**(1), 5–36 (2017)
12. Aleven, V., Koedinger, K.R.: Knowledge component approaches to learner modeling. In: Sottolare, R., Graesser, A., Hu, X., Holden, H. (eds.) *Design Recommendations for Adaptive Intelligent Tutoring Systems*, vol. 1, *Learner Modeling*, pp. 165–182. US Army Research Laboratory, Orlando (2013)
13. Stamper, J.C., Koedinger, K.R.: Human-machine student model discovery and improvement using datashop. In: Biswas, G., Bull, S., Kay, J., Mitrovic, A. (eds.) *AIED 2011. LNCS (LNAI)*, vol. 6738, pp. 353–360. Springer, Heidelberg (2011). [https://doi.org/10.1007/978-3-642-21869-9\\_46](https://doi.org/10.1007/978-3-642-21869-9_46)

<sup>1</sup> This work was supported by Bill and Melinda Gates Foundation Prime Award #OPP1196889.

14. Koedinger, K., McLaughlin, E.: Seeing language learning inside the math: cognitive analysis yields transfer. In: *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 32, no. 32 (2010)
15. Koedinger, K.R., Baker, R.S., Cunningham, K., Skogsholm, A., Leber, B., Stamper, J.: A data repository for the EDM community: the PSLC DataShop. *Handbook Educ. Data Mining* **43**, 43–56 (2010)
16. Cen, H., Koedinger, K.R., Junker, B.: Is over practice necessary? improving learning efficiency with the cognitive tutor through educational data mining. *Front. Artif. Intell. Appl.* **158**, 511 (2007)
17. Doroudi, S., Aleven, V., Brunskill, E.: Robust evaluation matrix: towards a more principled offline exploration of instructional policies. In: *Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale*, pp. 3–12 (2017)
18. Aleven, V., Sewall, J.: The frequency of tutor behaviors: a case study. In: Micarelli, A., Stamper, J., Panourgia, K. (eds.) *ITS 2016. LNCS*, vol. 9684, pp. 396–401. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-39583-8\\_47](https://doi.org/10.1007/978-3-319-39583-8_47)
19. Koedinger, K.R., Anderson, J.R.: Illustrating principled design: the early evolution of a cognitive tutor for algebra symbolization. *Interact. Learn. Environ.* **5**, 161–180 (1998)
20. Cen, H., Koedinger, K., Junker, B.: Learning Factors Analysis – A General Method for Cognitive Model Evaluation and Improvement. In: Ikeda, M., Ashley, K.D., Chan, T.-W. (eds.) *ITS 2006. LNCS*, vol. 4053, pp. 164–175. Springer, Heidelberg (2006). [https://doi.org/10.1007/11774303\\_17](https://doi.org/10.1007/11774303_17)
21. Corbett, A.T., Anderson, J.R.: Knowledge tracing: modeling the acquisition of procedural knowledge. *User Model. User-adapted Interact.* **4**(4), 253–278 (1994)
22. Lee, J.I., Brunskill, E.: The impact on individualizing student models on necessary practice opportunities. In: *International Educational Data Mining Society* (2012)
23. Heffernan, N.T., Koedinger, K.R.: The composition effect in symbolizing: the role of symbol production vs. text comprehension. In: *Proceedings 19th Annual Conference Cognitive Science Society*, pp. 307–312 (1997)
24. Heffernan, N.T., Heffernan, C.L.: The ASSISTments ecosystem: building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *Int. J. Artif. Intell. Educ.* **24**(4), 470–497 (2014)