A Multi-Pronged Redesign to Reduce Gaming the System

Yuanyuan Li[1], Xiaotian Zou[1], Zhenjun Ma[1], and Ryan S. Baker[2]

[1] Learnta, Inc., 1460 Broadway St., New York, NY 10036, USA
[2] University of Pennsylvania, 3700 Walnut St., Philadelphia, PA 19104, USA
leah.yuanyuan@gmail.com, vivian.xtzou@gmail.com,
will@learnta.com, rybaker@upenn.edu

Abstract. Despite almost two decades of interest in reducing gaming the system in interactive learning environments, gaming continues as a key factor reducing student learning outcomes and contributing to poorer learning outcomes. In this study, we redesigned the Kupei learning system by implementing a combined set of three interventions aimed at mitigating the impact of the two gaming behaviors we documented. Our results show evidence of a possible positive effect of the combined gaming prevention intervention at reducing the second type of gaming behavior within our system, however, it was not as successful at mitigating the first type of gaming behavior.

Keywords: Gaming the System, Learning engineering, Iterative redesign, Interactive learning environments

1 Introduction

Interactive learning environments are intended to create opportunities for students to learn but a substantial proportion of students choose instead to game the system, attempting to succeed by taking advantage of the regularities and properties of a system rather than by learning the material [1]. Considerable research has demonstrated negative correlations between gaming the system and student outcomes [2, 3].

Over the last 15 years, a range of interventions have been proposed and investigated. Several research groups attempted to mitigate the impact of gaming, in a more subtle fashion when gaming occurs and then adapt in real time to detection of gaming behavior [4, 5]. Other approaches attempted to prevent gaming behavior in the first place by adding delays or minimum amount of wait time between two actions [6, 7]. Some of these approaches [4, 5] improved learning outcomes but were not adopted at scale, while others may also hinder the usefulness of help-seeking for non-gaming students [7].

In this paper, we investigate a multi-pronged redesign of an AIED system, using three interventions in tandem to reduce students' propensity to game. We conduct a within-system quasi-experiment, investigating whether the redesigned version of the system leads to reduced gaming behavior and better within-system performance.

2 Method

2.1 Platform

Our intervention was developed in the context of the Kupei learning platform that supports the learning of math, English and science subjects. Rather than teacher-led instruction, the system uses algorithms that can automatically determine which content a student should work on next.

With the Kupei learning system, students usually take less than three practice sets to achieve basic mastery (probability of mastery falls between 80% and 95%) of each concept. Therefore, we define practices on the same concept after three practice sets as extra practices. We believe that a considerable proportion of extra practice will be the result of either gaming the system or struggling with the content.

Kupei uses Bayesian Knowledge Tracing (BKT) [8] to estimate student proficiency in real-time. When studying a concept using the Kupei system, Kupei assesses students' probability of mastering a concept after the first 3 items are completed. If the probability falls between 80% (a cut-off used by many commercial systems) and 95% (the original cut-off in) [8], then the concept is labeled as basic mastery, and the student will continue to work on two additional items. If the mastery probability is more than 95% (advanced mastery), then the system stops and advances the student to the next concept. If a student's probability of knowing a concept is less than 80% after the first three items are completed, the concept is labeled as unmastered, and the practice stops and displays the result. Students who did not master the concept (whether after 3 or 5 problems) are next required to complete an integrated review on the same concepts/skills (involving video and/or lecture notes). In all situations, the learning recommendation offered after each concept will change according to students' performance during the practice.

2.2 Gaming Behaviors

Prior to the integration of the gaming prevention intervention, gaming behaviors typically observed in the Kupei system can be divided into two types:

1. Students use an exhaustive method to obtain the correct answers of the practice sets by inputting random answers for each question of each practice set until earlier questions are re-shown.
2. Students open a practice set to obtain the set of questions, then quit the practice set midway to seek answers elsewhere.

2.3 Design and Method

Our design aimed to simultaneously accomplish two goals: first, by increasing the costs of gaming, it is hoped that students will game the system less often; second, with less gaming behaviors, we hope that students will engage in more productive behaviors and learn more effectively.

Aiming to achieve these objectives, we designed three gaming prevention interventions: first, we re-designed the system so that students may not complete more than two practice sets (of five problems each) on a concept more than three times a day, with a

pause of 36 hours before they can work on a concept again. Second, Kupei now provides meta-cognitive feedback which acts as a reminder to the students about the cost of gaming -- if they now game, they will have to wait 36 hours [5]. Third, the system now requires students who responded too quickly and failed to reach basic mastery to complete an integrated review on the same concepts/skills (involving video and/or lecture notes).

A within-subjects quasi-experiment was conducted comparing two 15-day math learning periods -- a control period before the new strategy for reducing gaming behavior was adopted, and an experimental period immediately following adoption of the new strategy within the system. We analyzed data (i.e. interaction logs) from a total of 343 students who studied at least 10 math concepts in both two periods.

3 Results

3.1 Frequency of Gaming Behavior by Condition

In this study, 93 students were control-gamers (they gamed the system during the control period) and 250 students were control-non-gamers (they did not game the system during the control period). After the gaming prevention interventions were integrated, the average gaming frequency per student decreased from 0.124 in control period to 0.064 in experimental period, a statistically significant difference, $V=5411$, $p<0.01$.

In terms of specific behaviors, there was a statistically significant reduction in gaming by quitting to seek answers, from 0.085 during control period to 0.031 during experimental period, $V=2511$, $p<0.01$. However, there was not a significant reduction in gaming by memorizing answers, from 0.040 during the control period to 0.032 during the experimental period, $V=2086$, $p=0.51$.

3.2 Other Behavior Changes

The Proportion of Extra Practice. The average proportion of extra practice decreased from 12.20% in the control period to 7.66% in the experimental period, which is statistically significant, $V=25588$, $p<0.01$. The proportion of extra practice in the control-gamers decreased from 23.1% in the control period to 13.4% in the experimental period, while for control-non-gamers, the proportion of extra practice decreased from 8.1% control to 5.5% experimental. According to a Wilcoxon rank sum test, the control-gamers' decrease in extra practice is significantly steeper than the control-non-gamers, $W=7928$, $p<0.01$.

Average Time Spent Per Item. Starting from the second practice set, there was a statistically significant increase in the average time spent on each item in the experimental period compared to the time spent in the control period, especially in the second practice set. The average time spent per item in the first practice set decreased from 98.87s in the control period to 91.06s in the experimental period, which is significantly different, $t(342)=3.32$, $p = 0.001$ for a paired t-test. In the second practice set, the average time a student spent answering each math item increased from 79.66s in the control period to 116.90s in the experimental period, which is statistically significant, $t(339)=12.75$, $p<0.01$. In the third practice set, the average time a student spent on each

math item increased from 85.05s in the control period to 97.41s in the experimental period, which is statistically significant, $t(216)=2.98$, $p<0.01$.

4 Discussion and Conclusions

In this paper, we attempted to address the two gaming behaviors that we documented within Kupei learning platform. We found that the multi-pronged gaming prevention intervention appear to have been successful at dissuading students from gaming the system. We detected a lower frequency of gaming behaviors in learning math after the integration of the gaming intervention. In addition, we found that fewer students used extra practice on a concept after the implementation of the gaming intervention. Instead, students spent more time on later items during the experimental period, possibly indicating students are practicing each item more seriously than students in the control period.

However, there appear to be some limitations to this approach that should be considered in future work. The intervention was not successful at reducing at addressing the first type of gaming behavior. Another possible limitation is that even if some students reduce their frequency of gaming the system, they may not replace gaming with the most desirable behaviors.

Ultimately, we hope that our research will inform the design of systems that will reduce students' motivation to game the system, and, in turn, increase the frequency of effective self-regulated learning strategies that lead to better student learning.

References

1. Baker, R. S., Corbett, A. T., Koedinger, K. R., & Wagner, A. Z. (2004). Off-task behavior in the cognitive tutor classroom. Proceedings of the 2004 Conference on Human Factors in Computing Systems - CHI '04 (pp. 383-390).
2. Aleven, V., McLaren, B.M., Roll, I., & Koedinger, K. (2006). Toward Meta-cognitive Tutoring: A Model of Help Seeking with a Cognitive Tutor. International Journal of Artificial Intelligence in Education. 16. 101-128.
3. Pardos, Z.A., Baker, R.S., San Pedro, M.O.C.Z., Gowda, S.M., & Gowda, S.M. (2014). Affective states and state tests: Investigating how affect and engagement during the school year predict end of year learning outcomes. Journal of Learning Analytics, 1(1), 107-128.
4. Baker, R.S., Corbett, A.T., Koedinger, K.R., Evenson, S., Roll, I., Wagner, A.Z., Naim, M., Raspat, J., Baker, D.J., Beck, J.E. (2006). Adapting to when students game an intelligent tutoring system. In Proc. of the Int'l Conf on Intelligent Tutoring Systems (pp. 392-401).
5. Arroyo, I., Ferguson, K., Johns, J., Dragon, T., Mehranian, H., Fisher, D., Barto, A., Mahadevan, S., Woolf, B. (2007). Repairing Disengagement With Non-Invasive Interventions. In Proc of the Int'l Conf on Artificial Intelligence in Education, 195-202.
6. Murray, R. C. & Vanlehn, Kurt. (2005). Effects of Dissuading Unnecessary Help Requests While Providing Proactive Help. In Proc of the Int'l Conf on Artificial Intelligence in Education, 887-889.
7. Aleven, V., McLaren, B., Roll, I., & Koedinger, K. (2004). Toward tutoring help seeking. Proc. of the Int'l Conf. on Intelligent Tutoring Systems, 227–239.
8. Corbett, A. T., & Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. User modeling and user-adapted interaction, 4(4), 253-278.