

The impact of gaming (?) on learning at the fine-grained level

Yue Gong¹, Joseph E. Beck¹, Neil T. Heffernan¹, Elijah Forbes-Summers²

^{1,2} Computer Science Department, Worcester Polytechnic Institute
100 Institute Road, Worcester, MA, 01609, USA

¹{ygong, josephbeck, nth}@wpi.edu

²elijah.forbes.summers@gmail.com

Abstract. One of the common expectations of ITS designers is that students efficiently learn from every practice opportunity. However, when students are using an Intelligent Tutoring System, they can exhibit a variety of behaviors, such as “gaming,” which can strongly reduce learning. In this paper, we present a new approach to infer the impact of gaming on learning at the fine-grained level. We integrated a knowledge tracing model of the student’s knowledge with the student’s gaming state (as identified by our gaming detector). We found that when gaming, students learn almost nothing (on the order of one-twelfth to one-fiftieth as efficiently). A student’s gaming amount is associated with aggregate effects on his knowledge and learning, leading to less learning even in the practice opportunities where no gaming occurs. In addition, we found that students tend to game in those skills on which they have relatively low knowledge. Furthermore, we found that knowing the identity of the student is more important than knowing the skill for predicting whether gaming will occur.

Keywords: Gaming, Knowledge tracing, Influences on learning

1 Introduction

With more and more students using Intelligent Tutoring System (ITS) in their daily study activities, their strategies for how to use ITS are becoming an important issue. Although ITS have been shown to have positive effect on helping student learning, different strategies of using ITS can lead to different learning outcomes [1, 2, 3, 4]. There are a variety of strategies exhibited by students, including “gaming,” which has been received a great deal of attention. A student is gaming if he is attempting to systematically use the tutors’ feedback and help methods as a means to obtain a correct answer with little or no work [5].

There have been many prior works showing that gaming behavior is generally associated with a reduced learning rate. Baker, et al. [3] used a traditional analysis method, applying a pretest and a posttest, to show that student gaming results in a poorer learning gain. A similar trend has been found by Walonoski, et al. [5] in a different computer tutor environment (ASSISTment) and using a different analysis method: longitudinal data analysis. However, those previous works explored the impact of gaming on learning by focusing on the long-term effects, thus their conclusions are based on the aggregated data. In other words, during a period of time,

the researchers tracked a sequence of student’s performances, and also the conditions of whether and how much gaming occurred during that time. They then came to the conclusion based on examining the relation between aggregate gaming occurrence and student performance. Perhaps the students who game happen to be the ones who don’t learn, but gaming is not the direct cause of the poor learning. Recent work by Corea, et al. [6] showed gaming has both immediate and aggregate effects on learning. They assessed whether gaming behavior is associated with immediate poorer learning, by applying learning decomposition method [7], where performance on a given skill at a given time is predicted based on the number times where the student previously engaged in gaming behavior on this skill. They found that the number of previous gaming behaviors is associated with less learning. Therefore, lower performance is predictable in the next problem after the student engages in gaming actions. They also pointed out that the apparent immediate impact of gaming, at the step level, appears to be due to a lack of learning at that very step where the gaming occurred; in other words, by gaming, an opportunity to learn is wasted.

One of the objectives of this study is to give a closer look at gaming’s impact on learning at the fine-grained level, namely, rather than examining the cumulative effects of gaming on learning during an amount of time, or the immediate effect contributed by the number of gaming behaviors that occurred previously, we aimed to track gaming’s quantitative effects on learning at the problem-solving level. We used student modeling as our conceptual framework, as it matches our requirements in this study by taking observations of a student’s performance and gaming state, and then using those to estimate the student’s latent attributes. We chose the knowledge tracing model (KT) [8], which is one of the most broadly used student modeling approach. It takes student performances as observations and uses those to estimate the student’s level of knowledge. Our motivation for using the knowledge tracing model is that we assume that if a student games on a problem, it negatively impacts the amount of learning from that problem. In addition, gaming is a state that varies across time, similarly to student knowledge in KT model. Therefore the knowledge tracing model is a good technique for our goal of exploring gaming’s effect on learning at the fine-grained level.

2 Methodology

2.1 Detection of gaming

Since student gaming is a mental state or goal of the student, it is not directly observable and cannot be determined as precisely as other student attributes, such as the correctness of a student’s response. In order to build the model for exploring the effect of gaming on learning, we first must have some way of informing the model that gaming occurred. For simplicity, rather than treating gaming as a latent variable, we used a knowledge-engineering approach and tagged it using human-made heuristics.

We constructed a gaming detector that contains three criteria for gaming actions.

- Rapid Guessing: submit answers less than 2 seconds apart at least twice in a row.
- Rapid Response: perform any action after a hint or starting a problem before a reasonable amount of time has passed (where “reasonable” is a fast reading speed for the content of the hint or problem body. We chose a reading rate of 400wpm).
- Repeatedly Bottom-out Hinting: reach a bottom out hint on three consecutive problems.

There might be more than one action made by a student while he is solving a main question (a main question consists of an initial question and a number of scaffolding questions or hint messages), such as performing second attempt, requesting a hint message or solving scaffolding questions, etc. For each of such actions, we assigned a gaming score that ranges between 0 and 1 (0 is not-gaming and 1 is gaming). We assume the student starts with a gaming score of 0. If the student does some action (matching at least one of the three gaming criteria is required) that we think is gaming, the gaming score’s value goes straight to 1. The student later can “recover” from a gaming state by performing any non-gaming actions (any other actions where none of the criteria is satisfied). With a non-gaming action detected, the gaming score decreases by 0.5. Therefore, each main question might be associated with multiple gaming scores representing how well the student performed in every sub-step of that main question. We then tagged that question using the score calculated by averaging the gaming scores.

Since we chose the knowledge tracing model as our framework, and since discrete variables are more commonly used in BNs, we further converted the continuous gaming score into a discrete value by selecting 0.5 as the cut point. In other words, if a main question is tagged with the average gaming score as greater than or equal to 0.5, the gaming state of that question is labeled by 1; otherwise, 0 is assigned.

2.2 Student modeling framework

Knowledge tracing model

Knowledge tracing [2], shown in Fig. 1, is an approach for taking student observations and using those to estimate the student’s level of knowledge.

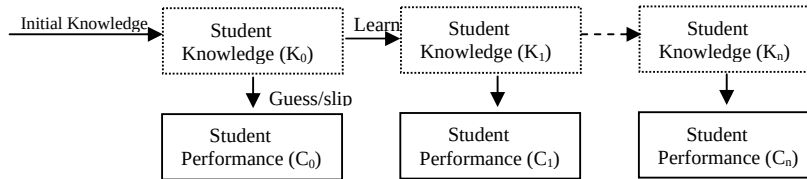


Fig. 1 Knowledge tracing model

$$\text{Initial Knowledge} = \Pr(K_0 = \text{True})$$

$$\text{Guess} = \Pr(C_n = \text{True} \mid K_n = \text{False})$$

$$\text{Slip} = \Pr(C_n = \text{False} \mid K_n = \text{True})$$

$$\text{Learning rate} = \Pr(K_n = \text{True} \mid K_{n-1} = \text{False})$$

There are two learning parameters. The first is initial knowledge (K_0), the likelihood the student knows the skill when he first uses the tutor. The second learning parameter is the learning rate, the probability a student will acquire a skill as a result of an opportunity to practice it. In addition to the two learning parameters, there are two performance parameters: guess and slip, which mediate student knowledge and student performance. In this paper we focus on the learning rather than the performance parameters.

Modified Knowledge tracing model

In order to explore the effect of gaming on learning at the problem-solving level, we first need to include student gaming state in the student model. We integrated it with knowledge tracing by putting in an additional node in the model structure, which indicates the gaming variable, shown in Fig. 2.

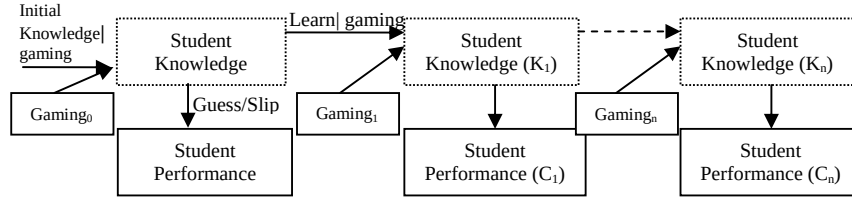


Fig. 2. Modified knowledge tracing model

The two performance parameters remain invariant after the modification, while the two learning parameters are changed. Initial knowledge is transformed from a prior probability to a conditional probability, thus there are two initial knowledge rates corresponding to the two given conditions: gaming and not-gaming. They indicate the likelihoods the student knows the skill when he first used the tutor, given whether he gamed on his first attempt or not. Similarly, the learning rate becomes two numbers after conditioning on the gaming state. The two learning rates are our major interests in this study. They indicate how much a student learns from a practice opportunity when he engages, or not engages in gaming behaviors in that practice. The difference between these two, if there is any, should be viewed as the immediate effect of gaming behaviors on learning.

2.3 Data set

For this study, we used data from ASSISTment, a web-based math tutoring system. The data are from 343 twelve- through fourteen- year old 8th grade students in urban school districts of the Northeast United States. They were from four classes. These data consisted of 193,259 main problems of ASSISTment usage during Nov. 2008 to Feb. 2009. Performance records of each student were logged across time slices for 106 skills (e.g. area of polygons, Venn diagram, division, etc).

For each student performance, we applied our gaming detector to identify the gaming state, and then fit the data to the modified knowledge tracing model. We used

the BNT-SM [9] and the expectation maximization (EM) algorithm to optimize data likelihood (i.e. the probability of observing our student performance data) in order to estimate the model’s parameters. We used the smoothing inference method [13] for using future data to estimate more plausible parameters. To address the problem of identifiability [10], we set Dirichlet priors [11] to initialize the EM algorithm.

3 Results

3.1 The impact of gaming

We trained a modified knowledge tracing model for each skill. I.e. observe all the data across all students for each skill and derive a set of 6 parameters (initial knowledge | gaming, initial knowledge | no-gaming, learning | gaming, learning | no-gaming, and guess and slip) for that particular skill. Thus, for 106 skills, we estimated 106 sets of parameters. Then, we calculated the mean values across all the skills (see Table 1). Other than the mean values, we also reported median to minimize the effect of outliers. However, in accordance with standard convention, our statistical analyses are based on the means rather than medians.

Table 1. Extended knowledge tracing model’s estimates of the learning parameters

Across Skills	Percent of gaming	Initial knowledge no gaming (K0_ng)	Initial knowledge gaming (K0_g)	Learning no gaming (Learning_ng)	Learning gaming (Learning_g)
median	0.139	0.527	0.149	0.158	0.003
mean	0.148	0.540	0.171	0.207	0.017
variance	0.004	0.019	0.010	0.022	0.002

From Table 1, we see that median and mean values suggest the similar trend that when students game in their first attempt with a skill, it is associated with lower initial knowledge. Meanwhile when a gaming behavior occurs, it immediately causes *much* less learning. The median value of Learning_g is 0.003, which is nearly 0, indicating essentially no learning. Although the corresponding mean is higher than the median, still, 0.017 is a very small number, especially compared to the counterpart 0.207, suggesting students learn 12 times faster when they are not gaming.

Another interesting observation is across 106 skills, our detector found that students gamed approximately 14% of the time. This number is much higher than what Baker reported in [3], suggesting that gaming is a much more common behavior. The difference can be possibly explained as our study was on a different experimental population, with the different classroom environment, and using a different computer tutor. However, given the results shown in Table 1 (0.158 vs. 0.003 and 0.207 vs. 0.017) it is plausible that our gaming detector just successfully captured certain kinds

of non-learning behaviors. Another possibility is that our gaming detector captures more than pure gaming behaviors (according to the definition of gaming given in [3]), our view is it would be a better goal for researchers to focus on the question of “what types of behaviors result in little or no learning,” rather than specific, named, behaviors.

3.2 The impact of gaming amount

After examining the immediate effect of gaming at the level of individual problems, we now inspect its aggregate effect at the student level. In other words, are there any differences between the students who appeared to game more in their performances and the students who behaved more seriously?

In order to make claims about students, we trained one model for each student by observing his responses in all questions across all skills. For each student, the model estimated a set of learning parameters corresponding to his individual initial knowledge and his learning rate, given whether he was gaming.

Based on how much a student gamed overall, we divided students into three equal sized groups having relatively high, medium and low gaming level. To avoid the potential impact of outliers, for each group, we employed the more robust measure of central tendency, the median (see Table 2). Also, the range, mean value and variance of the amount of gaming of each group are listed.

Table 2. Knowledge tracing parameters disaggregated by amount of overall gaming

Overall amount of gaming	Initial knowledge no gaming (K0_ng)	Initial knowledge gaming (K0_g)	Learning no gaming (Learning_ng)	Learning gaming (Learning_g)	Amount of gaming		
					range	mean	var.
High	0.337	0.235	0.114	0.021	19.7% - 85.3%	33.2%	0.012
Medium	0.538	0.345	0.151	0.049	8.4% -19.6%	13.3%	0.001
Low	0.705	0.358	0.219	0.089	0 - 8.3%	4.1%	0.001

As shown in Table 2, students who don’t game much start with more incoming knowledge. When comparing initial knowledge given no gaming (leftmost column of numbers), we see a clear trend, which is higher frequently gaming students are associated with lower initial knowledge. However, it is not the case in initial knowledge given gaming. The three numbers in the third column are close to each other, which suggests students tend to game on their weaker skills. In other words, when encountering skills they know little about (say, having only 35.8% chance of knowing the skill), even those infrequent gaming students also conduct some gaming behaviors. We are interested to see whether there exists a threshold for the occurrence of gaming: is it the case that whenever a student’s incoming knowledge is absolutely lower than a certain level, he tends to game when using the tutor? As can be seen in Table 2, low gaming students game with an average initial knowledge of 0.358, when high gaming students are *not* gaming their average initial knowledge is 0.337 (bolded

values). Therefore, it appears such threshold does not exist, and the low knowledge that determines whether a student games is relative. In other words, students are inclined to game on their relatively weaker skills.

For learning rate, there is a consistent trend that students who game less learn more quickly both when gaming and when not. We noticed for those frequently gaming students, even for those skills they don't appear to game, their initial knowledge are still fairly low and the learning rates remain the lowest among the three groups. We think those students who are found to game here probably also game in other contexts, including before our study. Therefore, they are estimated with lower initial knowledge due to the possibility that a lot of practice opportunities were wasted. Another possibility for their lower initial knowledge is their learning rate is not as high as the serious students'.

3.3 Which is more useful for predicting gaming: Student or Skill?

What causes student gaming is another research question in this study. Prior work inspected lesson vs. student [12], finding that for determining gaming, knowing the students has much less predictive power than knowing the tutor lessons. Our objective is to compare between student and skill. Is this just the fact that some students game more than others? Or are some skills just too hard to solve, so the students game on those problems? In order to resolve this problem, we did two ANOVAs. For each student, for each skill he attempted, we calculated the percent of gaming that occurred across all the questions he solved for that skill. E.g. one row of the well-prepared data is "Tom, Venn Diagram, 15%".

We assigned percent of gaming as the dependent variable in two models, and student as the independent variable in one ANOVA, and skill as the independent in the other one. We compared the two models to see which one accounts for more variance of the dependent. We found that the R^2 from the student ANOVA is 0.61, which is more than 5 times greater than the R^2 from the skill ANOVA, 0.11. Thus, student is more closely related to, and more predictive of gaming than, skill.

4. Contributions

In this work, we performed a quantitative comparison between how quickly students learn while they are engaged in gaming the system and presented several novel empirical results. We found when gaming occurs in a practice opportunity, the learning rate is near 0, indicating gaming has large, immediate, negative impact on learning. Meanwhile, a student's gaming amount causes aggregate effects on his knowledge and learning. Students who game more frequently have lower estimated initial knowledge, especially in those skills they game (but even in skills where they don't). They also learn more slowly compared to students who game less. Furthermore, we found gaming varies more at the student than at the skill level. It is more than five times as useful to know students. This finding is contrary to the results in [12] where they looked at lesson vs. students, and found there is little predictive power from knowing students.

In addition, we presented a new approach for detecting the effect of gaming at the fine-grained level. There have been many prior works using different data analyses and experimental methods for examining gaming's impact and showing student gaming is associated with substantially less learning. Most of them focused on the cumulative effects of gaming, we, however inspected the immediate impact of gaming when it occurs. We combined the goal of this study with the concept of student modeling, as rather than measuring the effects of gaming at a coarse-grained level, we aimed to make inference about it at a problem-by-problem level. Therefore, we augmented the knowledge tracing model with the student gaming state, by modifying the regular Bayesian network structure and enabling it to handle gaming as an additional variable. This approach is easily generalized and applied to other attributes which may have real-time impact on student knowledge in student's practices, such as motivation, attitude, etc.

In this study, we employed the smoothing [13] inference method for using future data to estimate more plausible model parameters. Traditionally, the monitoring method is used for student modeling, because the common objective of the model training is about prediction and future data are unavailable for this goal. For ITS, however, making scientific claims about students is also a desired aspect. For estimating the impact of gaming, looking at future data is helpful as it contains a great amount of information. Related work [14] explored hand-derived formulas to use the future two actions for estimating parameters. On the contrary, we are able to take advantage of the smoothing inference method to peek at the future data for free. Therefore, this approach does not require any algebra calculations and its scalability to complex models is not limited.

5. Future work

Student gaming is a hidden behavior that is difficult to determine externally. Therefore, most prior works detected gaming either by using human observers or utilizing machine learning processes. Ultimately, gaming should be inferred similarly to student knowledge and estimated by Bayesian inference. Although, we still constructed a gaming detector based on knowledge engineering to identify gaming, this study is the first step towards that final goal as we have explored the impacts of gaming via inferences.

Our intuition in reusing knowledge tracing to model gaming is that we believe that when students game it largely prevents them from acquiring knowledge from that practice opportunity. Therefore we built the model with an arrow oriented from gaming state to student knowledge and both of them lay in the same time slice, indicating they happen simultaneously. We think it is a productive step to look at other model structures which are inspired by other hypotheses, such as student gaming in a practice opportunity perhaps influences the cumulative knowledge in his next practice; or how much a student knows may have predictive power to determine whether gaming will occur in his next practice opportunity. According to those hypotheses, by tweaking model structures, we might be able to make more interesting findings in terms of student gaming.

6. Conclusions

This paper has presented an approach for estimating the effect of gaming at the fine-grained level. We integrated student gaming state with the knowledge tracing model. We used the smoothing inference method to take advantage of future data, without any human-made heuristics, in order to produce plausible parameter estimates.

We found that gaming has strong negative effects on learning. When gaming occurs, students basically don't learn. The amount of gaming has aggregate effects on student knowledge and learning as well. The more the student games, the slower the student learns in general. Even in the situations where the frequently gaming students don't exhibit any gaming behaviors, their estimated initial knowledge is lower and they learn less efficiently. In addition, we found there is no absolute level of initial knowledge that once lower than that, students tend to game. Rather, students tend to game on those skills on which they have relatively lower incoming knowledge. Thus even for those well-behaved students, they may engage in gaming if they are required to solve the questions beyond their knowledge. Finally, with respecting to gaming behavior, we compared the predictive power from knowing the student and knowing the skill. We have shown that being aware of student is more helpful for predicting whether gaming will occur.

References

1. Aleven, V., McLaren, B.M., Roll, I., Koedinger, K.R.: Toward tutoring help seeking: Applying cognitive modeling to meta-cognitive skills. *Proceedings of the 7th International Conference on Intelligent Tutoring Systems (ITS 2004)*, pp.227-239.
2. Arroyo, I., Woolf, B.: Inferring learning and attitudes from a Bayesian Network of log file data. *Proceedings of the 12th International Conference on Artificial Intelligence in Education*. pp.33-40.
3. Baker, R.S., Corbett, A.T., Koedinger, K.R., Wagner, A.Z.: Off-Task Behavior in the Cognitive Tutor Classroom: When Students "Game The System". *Proceedings of ACM CHI 2004: Computer-Human Interaction*. pp.383-390.
4. Stevens, R., Soller, A., Cooper, M., Sprang, M.: Modeling the Development of Problem-Solving Skills in Chemistry with a Web-Based Tutor. *Proceedings of the 7th International Conference on Intelligent Tutoring Systems (ITS 2004)*, pp.580-591.
5. Walonoski, J., Heffernan, N.: Detection and analysis of off-task gaming behavior in intelligent tutoring systems. *Proceedings of the 8th International Conference on Intelligent Tutoring Systems (2006)*. pp.382-391.
6. Cocea, M., HersHKovitz, A., Baker, R.S.J.d.: The Impact of Off-task and Gaming Behaviors on Learning: Immediate or Aggregate? *Proceedings of the 14th International Conference on Artificial Intelligence in Education*. pp.507-514.
7. Beck, J. E., & Mostow, J. (2008). How who should practice: Using learning decomposition to evaluate the efficacy of different types of practice for different types of students. *9th International Conference on Intelligent Tutoring Systems, Montreal*, 353-362.
8. Corbett, A., Anderson, J.: Knowledge tracing: Modeling the acquisition of procedural knowledge. *User modeling and user-adapted interaction*, 1995. 4: pp. 253-278.
9. Chang, K.-m., Beck, J., Mostow, J., Corbett, A.: A Bayes Net Toolkit for Student Modeling in Intelligent Tutoring Systems: *Intelligent Tutoring Systems: Intelligent Tutoring Systems Volume 4053/2006*

10. Beck, J. E., Chang, K.-m.: Identifiability: A Fundamental Problem of Student Modeling. Proceedings of the 11th International Conference on User Modeling, Greece
11. Rai, D, Gong, Y, Beck, J. E. : Using Dirichlet priors to improve model parameter plausibility. Proceedings of the 2nd International Conference on Educational Data Mining, Cordoba, Spain, pp141-148.
12. Baker, R.S.J.d.: Differences between Intelligent Tutor Lessons, and the Choice to Go Off-Task. Proceedings of the 2nd International Conference on Educational Data Mining, pp.11-20.
13. Russell, S., Norvig, P.: Artificial Intelligence A Modern Approach (Second edition)
14. Baker, R.S.J.d., Corbett, A.T., Aleven, V.: More Accurate Student Modeling Through Contextual Estimation of Slip and Guess Probabilities in Bayesian Knowledge Tracing. Proceedings of the 9th International Conference on Intelligent Tutoring Systems. pp. 406-415