Boosting Engagement With Educational Software Using Near Wins

Mohammad M. Khajah[1,2], Michael C. Mozer[2], Sean Kelly[3], and Brent Milne[3]

[1] Kuwait Institute for Scientific Research,
mmkhajah@kisr.edu.kw
[2] University of Colorado Boulder, CO, USA,
mozer@colorado.edu
[3] Woot Math Inc., CO, USA,
sean.kelly@wootmath.com,
brent.milne@wootmath.com

Abstract. Boosting engagement with educational software has been promoted as a means of improving student performance. Various engagement factors have been explored, including choice, personalization, badges, bonuses, and competition. We examine two promising and relatively understudied manipulations from the realm of gambling: the *near-win effect* and *anticipation*. The near-win effect occurs when an individual comes close to achieving a goal, e.g., getting two cherries and a lemon in a slot machine. Anticipation refers to the build-up of suspense as an outcome is revealed, e.g., revealing cherry-cherry-lemon in that order drives expectations of winning more than revealing lemon-cherry-cherry. Gambling psychologists have long studied how near-wins affect engagement in pure-chance games but it is difficult to do the same in an educational context where outcomes are based on skill. In this paper, we manipulate the display of outcomes in a manner that allows us to introduce artificial near-wins largely independent of a student's performance. In a study involving thousands of students using an online math tutor, we examine how this manipulation affects a behavioral measure of engagement—whether or not a student repeats a lesson. We find a near-win effect on engagement when the 'win' indicates to the student that they have attained critical competence on a lesson—the competence that allows them to continue to the next lesson. Nonetheless, when we experimentally induce near wins in a randomized controlled trial, we do not obtain a reliable effect of the near win. We discuss this mismatch of results in terms of the role of anticipation on making near wins effective. We conclude by describing manipulations that might increase the effect of near wins on engagement.

Keywords: near-win, educational applications, anticipation

1 Introduction

The *near-win effect* is a manipulation that has been studied extensively in the gambling addiction literature. This effect occurs when a player almost wins a game, e.g., getting two cherries and a lemon in slot machine (Figure 1). Reid [1] cites several other examples of this effect in real-life: near-win number sequences in lottery tickets, close finishes in horse racing, and customers thinking they are half-way to success when completing half the winning sequence in a promotional sales campaign. Reid argues that near-win events are useful in skill-based games, such as darts, because they provide feedback that winning is close. Even though such feedback is useless in games of pure chance, Reid notes that gamblers still think they can influence in the outcome with behaviors such as whispering to the dice, choosing lottery numbers carefully, or consulting books of lucky numbers.

A complementary manipulation that often gets studied alongside the near-win effect is *anticipation*. Studies on the near-win effect usually find differences in engagement between losing early (also known as a *clear loss*) and nearly winning. Reid [1] cites one study where subjects preferred a near-win over a clear-loss (Reid's own study also showed a trend in that direction). Video game developers also use anticipation, usually in scenarios with close calls. For example, a scene from Half-Life 2 (Figure 1, right panel), a first-person shooter, shows a large chimney slowly falling down to block the path of the player. The scene is gripping because all the player can do is press full throttle—to escape pursuers—and hope that she can clear the falling chimney. Ultimately, the chimney falls before the player gets to it so she has to lift off the throttle and maneuver around the rubble. If the scene was different, say with the chimney already in ruins, it would not be as captivating.

Given the effectiveness of near-wins and anticipation in promoting engagement in gambling and video games, we explore whether the benefits of those manipulations transfer to an educational context. We designed a novel manipulation of the near-win effect that independently induces near-wins in a skill-based context, and we continuously manipulate anticipation using Bezier-curve-based animations. Finally, we analyze the impact of those two manipulations on a behavioral measure of engagement in a large-scale math-tutoring software used by thousands of student.

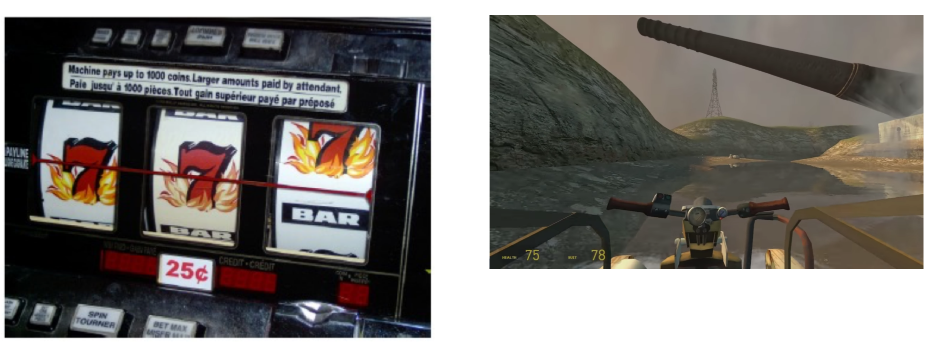

Fig. 1: (Left) Examples of the near-win effect in slot machines [2] and (right) anticipation in video games.

1.1 Related Research

In an educational setting, one ideally wants to maximize learning or performance on a test. However, conducting experiments that evaluate learning is difficult as it requires pre- and post-testing of students in a controlled setting. In this paper, we use *engagement* as a proxy for learning, with the working assumption being that extra engagement leads to greater learning. Engagement has no precise definition and researchers have operationalized it via surveys, such as the Game Engagement Questionnaire (GEQ) [3], physiological measures, such as galvanic skin response, and behavioral measures, such as voluntary duration of play, also known as *persistence* [4]. The literature on factors influencing engagement often draws on Csikszentmihalyi's theory of *flow* [5], a psychological state in which the person becomes fully absorbed in an activity, loses track of time and experiences high level of enjoyment.

Researchers have explored the impact of various task-irrelevant game-like features on engagement in educational contexts. For example, Cordova and Lepper [6] found that students who played a learning game which was enriched with motivational elements—contextualization, personalization, and choice—had greater learning gains and subjectively indicated higher levels of intrinsic motivation. Denny [7] studied virtual achievements, a common gamification feature, and found that while they increased participation in an online learning tool, they did not have an impact on performance. Contrary to the previous two studies, Katz et al. [8] investigated the *removal* of gamification features on performance and engagement in the classic n-back working memory task [9]. Katz found that removing the persistent display of scoring information led to significant improvements in training performance, and that there was otherwise no difference in subjective engagement and generalization performance among tasks with and without gamification features.

The studies mentioned in the previous paragraph are examples of *overt* manipulations that are salient to the user. More subtle manipulations have been explored in the context of video game engagement. For example, Denisova and Cairns [10] found that telling players that a game has adaptive artificial intelligence—when it actually doesn't—before playing results in greater immersion compared to a condition where players are told nothing about the game beforehand. Khajah et al. [4] found that covertly assisting the player in a simple two-dimensional game can lead to greater persistence.

The near-win effect is also a subtle manipulation but it has not received much attention outside of the gambling psychology literature. In gambling, near-wins can be induced by rigging a slot machine to deliver a certain frequency of almost-winning sequences. Kassinove and Schare [11] implemented this manipulation and found that subjects voluntarily played a slot machine for longer when the proportion of near-wins was medium (30%), not too small nor too large. They hypothesize that high near-win rates can lead to desensitization, i.e., crying wolf too many times, and low rates do not have enough impact In education, Lomas [12] studied the effect of close/far losses/wins on engagement in an fraction learning game used by thousands of students. In the game, students see a scorecard after completing a level which contains the score required to win the level and the student's actual score. Engagement was defined as the number of additional exercises attempted beyond the first two mandatory levels. Lomas found that the number of additional exercises attempted increased as the absolute difference between the target criterion and the actual score decreased, providing the famous inverted U relationship between difficulty and engagement (this inverted U relationship indicates that there is a level of difficulty, not too low or too high, which promotes maximum engagement). Lomas argued that these results indicated that the inverted U was a result of the drama of nearly winning or losing, rather than a moderate amount of challenge, as is often hypothesized by theories of flow. However, Lomas' work is observational, and without randomized controlled trials we cannot determine whether near wins have a causal effect on engagement.

In this paper, we manipulate near-wins independently of actual performance in a skill-based context. This enables assessment of the effect of near-wins on engagement. We also vary anticipation by manipulating the temporal dynamics of the animation that reveals a student's performance score. Our hope is that by

2 Experimental Manipulation

Woot Math® [13], an interactive web-based fraction learning software used by thousands of students, served as our platform to implement the near-win and anticipation manipulations. In Woot Math, students engage in a series of lessons where each lesson consists of a set of exercises that are chosen dynamically depending on the student's performance. Examples of these exercises include comparing fractions, placing decimals on the number line, adding and removing fractions, fraction multiplication, etc. After every lesson, a scorecard is shown with a *performance bar* indicating the score and three *goal posts* corresponding to thresholds for earning an additional *star* (Figure 2). Between zero and three stars can be awarded on any lesson, and the performance bar range is continuous in $[0, 3]$. An animation function is used to fill the performance bar (the yellow coloring in the figure) from left to right, and then award stars based on the goal posts and the student's score. Replay and continue buttons on the scorecard allow students to retry the current lesson or return to a main lesson-selection screen.

The awarding of one star is of critical importance to students, because they cannot advance to the next lesson unless they score at least one star. The additional stars may be intrinsically rewarding to a student, and for some classes using the software these additional stars may influence a teacher's grade, but over the population of students, the common goal is passing the threshold to obtain one star.

The near-win and anticipation manipulations are both implemented in the scorecard screen. The near-win manipulation is based on a single parameter $\nu \in [0.1, 0.9]$, assigned at random to each student in each session, which specifies the probability of *artificially* inducing a near-win event. Artificially induced near-

Fig. 2: The scorecard in the Woot Math software.

win events boost the score to within 0.1 of the next goal post: the manipulated score, S, is a draw from a uniform distribution over the range $[\lceil R \rceil - 0.1, \lceil R \rceil]$, where R is the true score. The score is not manipulated if the student's true score is already within 0.1 of the next goal post—a *natural* near win—or is just over the goal post. Consequently, under no circumstance does the score manipulation result in altering the number of stars a student receives. It is important to note that, contrary to gambling literature, the rate of actual wins is not controlled—it is determined by the student's skill. Nonetheless, the rate of *near* wins is influenced by ν.

The time course of animating the performance bar is determined by a parameter $\eta \in [0, 1]$, which we refer to as the *ease-out magnitude*; η controls the deceleration of the animation as it approaches the target score. Figure 3 graphs the proportion of the performance bar that is colored as a function of time for various ease-out magnitudes ν and for various target scores. All curves are generated using $y(t) = Sb(t/T)$ where S is the target score, $T = 0.3 + S$ is the animation duration, and $b = \text{Bezier}(0, 0, \eta, 1, 1, 1)$ is a unit cubic Bezier function. Animation duration depends on the score because otherwise the mean speed would change, enabling prediction of the target score early on. Small η values produce a linear animation which stops at the target score suddenly so the student is maximally uncertain about when and where the animation will terminate. In contrast, large η values produce an animation which quickly jumps to near the target score and then slowly approaches the target score. We hypothesized that $\eta = 1$ produces more anticipation than $\eta = 0$ because in the case of near-wins, the first creates more anticipation as the animation slows down while the latter distributes anticipation evenly over multiple goal posts. To see how we hypothesize anticipation plays out in the student's mind, consider a hypothetical model in which anticipation increases linearly if the score in the performance bar is within 0.25 units of the next goal post, otherwise the anticipation is set to zero (Figure 4). For a target displayed score of 2.90, the figure plots how anticipation

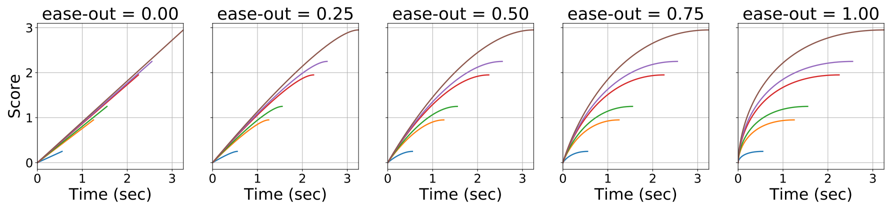

Fig. 3: The anticipation manipulation. Each plate plots the score in the performance bar as a function of time. Different curves correspond to different target scores.

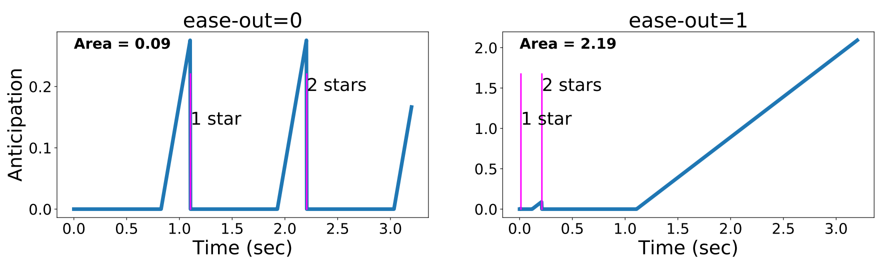

Fig. 4: Simulation of a simple model of anticipation using different animation functions. The magenta lines indicate when the progress in the performance bar clears a goal post. The area under the curve is shown for both cases.

changes over time for $\eta = 0$ (left) and $\eta = 1$ (right). Given the same score, both conditions will animate over the same duration, but there are clear differences in how anticipation builds-up. When $\eta = 1$, animation clears the first two goal posts quickly, causing very little anticipation, but then it slows down and spends of most the time within the 0.25 region of the next goal post, causing a large build-up in anticipation.

The engagement measure in this experiment is whether students replay the lesson. The experiment was conducted over a period of three months from March 2017 to June 2017 on the Woot Math platform. Students were randomly assigned to a design (ν, η) in the two-dimensional design space. Every time a student completed a lesson, a log entry was added to the dataset with information including the design, the actual score, whether a near-win was triggered, and whether the student replayed, continued, or quit afterwards. We limit the analysis to entries in the dataset where the score was less than 3.0. The analyses in the next section apply additional filters on the dataset, as we discuss. Although students typically practice over multiple sessions, we simply concatenate lessons from multiple sessions to form one sequence.

3 Results

Our dataset consists of 129,214 entries from 7,976 students. The median number of lessons attempted by a student is 9 lessons (std. 22.00, range 1-311), over a median of 3 sessions (std. 4.83, range 1-51), with a lesson duration of 2.8 minutes (std. 11.00, range 0.15-1,441.60). After excluding perfect and zero scores and instances where the student moved on before the progress bar animation had completed,[4] the dataset reduces to 29,470 entries from 5,953 students. The median number of lessons drops to 3 lessons (std. 6.54, range 1-124), over 2 sessions (std. 3.20, range 1-39), with a lesson duration of 3.8 minutes (std. 9.64, range 0.31-408.88).

Recall that *natural* near-wins occur when the student's actual score is close to the next goal post, whereas *artificial* or *induced* near-wins are the result of increasing the actual score to within 0.10 of the next goal post. This distinction means that we can look at (i) how *any* near-win (natural or induced) affects engagement, and (ii) how our specific manipulation, artificial near-wins, modulates engagement.

We answer the first question by examining how the score shown to the student—the displayed score—affects engagement. This score captures both natural and artificially induced near-wins. Because students can advance to the next lesson only if they achieve at least 1 star, we expect a near-win effect for students scoring below 1-star. For this situation, "winning" has meaningful impact within the application, whereas achieving 2- and 3-stars has only intrinsic reward value. In order to classify events as near-win or not, we must pick a performance threshold p, quantified in terms of the distance to the next goal post, that constitutes a near-win from the player's perspective. Because we have no way of determining this threshold, we evaluate for two reasonable possibilities: $p = 0.25$ and $p = 0.10$.

Figure 5 shows the replay probability as a function of displayed score for thresholds $p = 0.25$ and $p = 0.10$. The orange bars denote replay probability for a near-win situation (as defined by the threshold), and the blue bars for the no near-win situations. In the top row, we include all near-win lessons, regardless of whether they were induced or not. In the middle row, we exclude induced near-win lessons, focusing only on natural near-wins. In the bottom row, the orange bars only include induced near-wins and since those are always shown within 0.1 of the next goalpost, there is no difference in the height of the orange bars between the left and right plates in the bottom row. As the score increases, there is a clear downward trend in mean replay probability, for both near-win and non-near-win lessons. Students are most likely to replay following lessons in

[4] This information is not explicitly provided so we infer it by measuring how long the student stayed before continuing or quitting. If the duration is less than the animation duration, the entry is excluded. The dataset does not explicitly provide an animation duration field so it is computed, as described earlier in the Chapter, via $T = 0.3 + S$, where S is the displayed score. It turns out that this filter also eliminates all entries where the student quit the program afterwards. In other words, students who quit do so immediately, without waiting for the animation to finish.

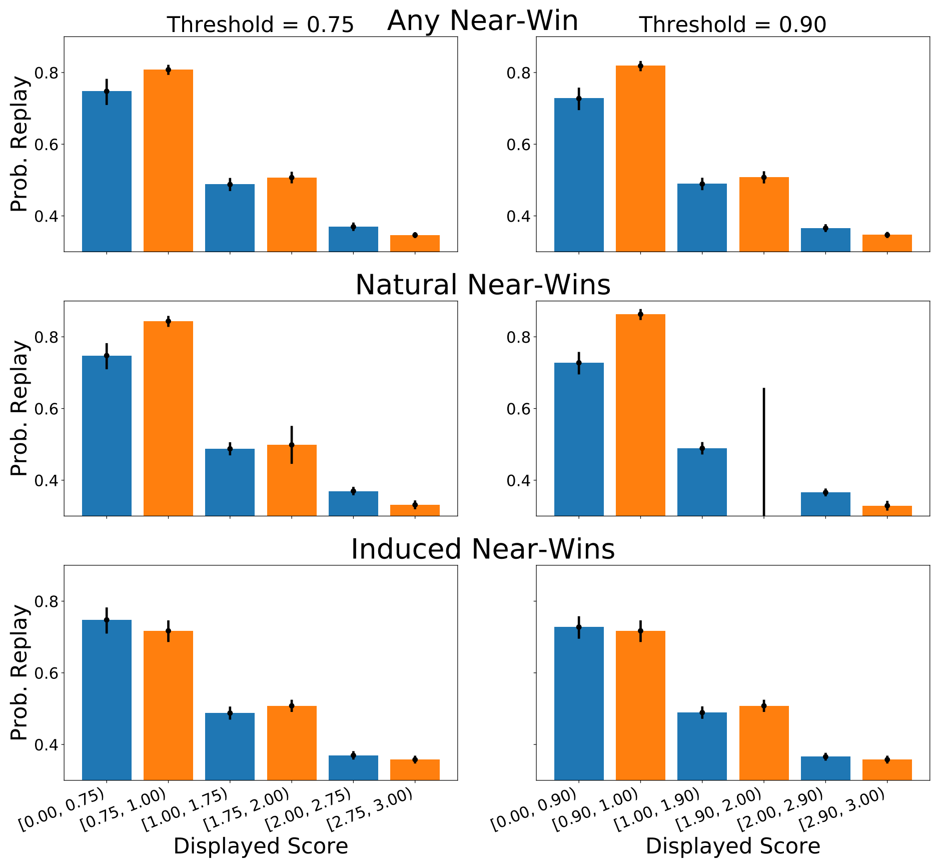

Fig. 5: Analysis of replay rates as a function of the displayed score. Each bar corresponds to a score range $[a, b)$, with a included and b excluded. Bar heights indicate the mean replay probability of observations in the corresponding score range. Blue and orange bars correspond to no-near-win and near-win events, respectively. (Top) orange bars contain all events that fall within the corresponding bin, (Middle) orange bars contain events that are *not* induced near-wins, and (Bottom) orange bars correspond to induced near-wins. Error bars correspond to the 95% confidence interval for a binomial proportion, using Wilson's score interval.

which their displayed score is lower than 1.0. These lessons also show a reliable near-win effect: as the score approaches but does not reach 1.0, students are most eager to try the lesson again. As we hypothesized a priori, near-wins have the largest effect for the students attempting to reach the one-star criterion, which is a prerequistie for advancing in the software. Interestingly, the middle row in Figure 5 suggests that the boost in engagement for low-performance lessons is mostly due to natural near-win events.

This brings us to our second question, how *induced* near-wins affect engagement. The bottom row of Figure 5 shows no difference in engagement when a near-win is artificially induced through a randomized controlled trial, supporting our earlier observation that *natural* near-wins are responsible for the increase of engagement on low-performing lessons. Students may be able to self-assess their true performance, and thus be biased by their actual scores. Our data set does not support investigating this self-awareness hypothesis, but it seems implausible to us that students are internally modeling the Woot Math score keeping algorithm. Alternatively, near-win events may interact with anticipation, and the effectiveness of a near win may depend on appropriate build-up of anticipation. We examined this possibility by limiting the analysis to trials with low and high ease-out values (left and right plates of Figure 6, respectively). Here, we only show the interaction with anticipation for $p = 0.1$, but the findings are similar for the $p = 0.25$ case. We see a trend that suggests that near-win events increase engagement at all goal posts when anticipation is high and fails to do so when anticipation is low. However, this trend is not statistically reliable. We note that the induced near-win analysis at the one-star level (bottom row of Figure 5) is based on about 40% of the data as the full analysis (top row of Figure 5), which might explain failure to find reliable near win effects.

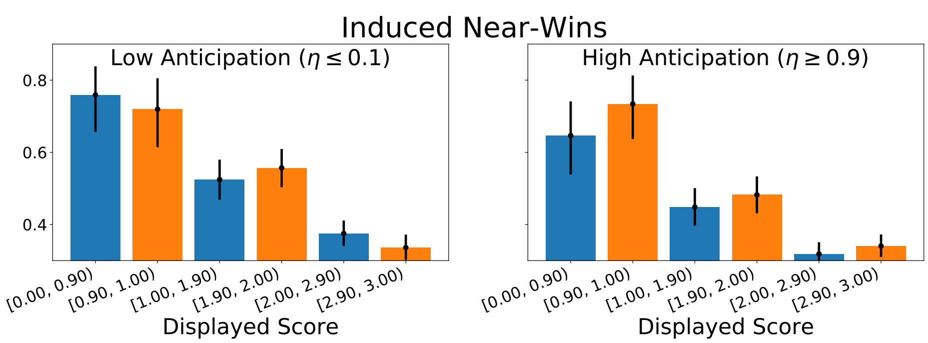

Fig. 6: Analysis of replay rates as a function of the displayed score and the occurrence of an induced near-win, restricted to trials with low and high anticipation (left and right plates, respectively). See the caption of Figure 5 for a description of Figure elements.

So far, we have coarsely characterized engagement as a function of the scoring bin which may obscure finer-grained relationships within individual bins. To look at these relationships in greater detail, we built and evaluated logistic regression models that predict whether the student replays a lesson or not as a function of the induction of a near-win N_w (binary variable), the actual un-manipulated distance to the next goal post D and its square D^2 (continuous variables), the star level S (one-hot encoded variable), the ease-out value η (continuous variable), and the log of the time it took to complete the lesson T (continuous variable). Since there are 6 inputs, the number of possible models is $2^6 = 64$ models. We constrained each model to contain first-, second-, and third-order interaction terms, in addition to the intercept term. We evaluated the 64 models via 10 replications of 5-fold cross-validation on our dataset. Entries where the student did not experience a natural near-win were the only ones to be included in the analysis. The top row of Figure 7 plots the results of the top 5 models in terms of mean test loglikelihood and area-under-the-curve (AUC). There are no differences in test performance among models that include manipulation terms—the near-win N_w and easing η—and those that do not. This is consistent with the previous analysis which found no difference in engagement as a result of induced near-wins.

One possible explanation for the null effect of near-wins on engagement is the low median number of lessons (3) which means that students may disengage, regardless of the manipulation. To investigate this further, we limited the analysis to students who completed at least 10 lessons in which they did not achieve a perfect score. This reduces the dataset to 8,330 entries from 503 students. The median number of lessons attempted is 14 lessons (std. 9.18, range 10-83), over 7 sessions (std. 3.66, range 1-34), with a lesson completion time of 3.34 minutes (std. 11.00, range 0.54-409). As before, we evaluated various logistic regression models using 10 replications of 5-fold cross-validation. Because students are doing a minimum of 10 trials, we added the near-win probability ν as an additional factor to explore, since students will do enough trials to get a sense of the frequency of near-win events. The bottom row of Figure 7 plots the results of the top 5 logistic regression models in terms of negative mean test log-likelihood and AUC. Similarly to the previous analysis, models which include the manipulation factors—N_w, E, η—do not have a predictive advantage over models that lack those factors.

4 Discussion

We introduced a novel manipulation of near-win events in an educational skill-based context. This manipulation was coupled with an anticipation manipulation that controls how much build-up of suspense occurs before students see their scores. A basic analysis of the displayed scores revealed that the lowest-performing students are more likely to replay when experiencing a near-win event. We believe this is due to the requirement that a student achieves a minimum of one star to unlock new levels. In other words, when a goal post has

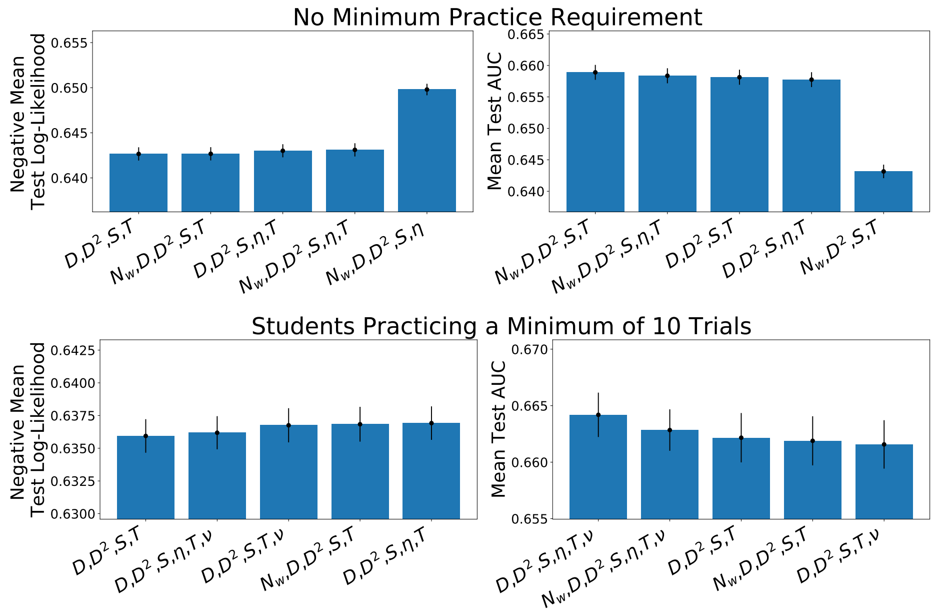

Fig. 7: Cross-validation results of logistic regression models trained to predict the probability of replay. The top row includes all non-natural-near-win trials, while the bottom only includes trials from students who completed a minimum of 10 trials. The top 5 models in terms of negative test loglikelihood (lower is better) and AUC (higher is better) are shown in the left and right columns, respectively. Each bar corresponds to a model and error bars correspond to standard errors. Models are represented by the factors included in the model.

meaningful impact within the application, students are more motivated to replay if they experience a near-win. Achieving 2- or 3-stars does not unlock extra game levels so students may not have real incentive to score higher.

In contrast to previous observational studies [12], our procedure allows us to conduct randomized controlled trials to determine if the perception of near wins *causes* an increase in engagement. To our surprise, we found no evidence for this causal relationship. One explanation is that students may be aware of their actual score, which would attenuate the effectiveness of the manipulation. However, we find this explanation unlikely because the software does not display a running tally of correct answers nor does it show the number of questions the student will answer in advance. Another possibility is that near-wins are effective when only there is anticipation or build-up. Subsequent analysis showed a trend in this direction. The lack of predictive power of near-win events may be due to engagement being primarily a function of student skill, which is encoded as the score on a lesson. This explains why artificially induced near-win events do not predict a different level of engagement, compared to non-near-win events.

A limitation of our anticipation manipulation was that our maximum animation duration was 3.6 seconds, and stretching that time seems to amplify anticipation. We were unable to increase the duration due to the fact that it would waste students' time on the scorecard screen. Our informal testing indicates that a minimum of about 7-8 seconds is required to buildup anticipation, which could be too long given that the median duration students stay on the scorecard screen is 8.3 seconds. One way to solve this problem is to accentuate the ease-out manipulation by adding sounds—such as the ticking sound of a roulette wheel—to increase the impact of the effect.

Students often score three stars—potentially due to the adaptive nature of the program—and attempt a small number of lessons, making it hard to analyze the effect of near-win probability on engagement. Given the high exclusion rate of the near-win manipulation, a redesign is necessary so that the manipulation is independent from student performance. One possibility is to define "winning" as getting a bonus award (e.g., a virtual pen color in Woot Math) or getting to play a round of a simple video game. The new manipulation would then control how often students win, nearly-win, or lose. Another possibility is to show the student a fake list ranking the student's performance relative to students in other schools (so that students in the same class don't notice the manipulation). The list would contain targets or goal posts and the location of the target or goal post changes depending on whether a near-win event is triggered or not.

The marginal effect of the design space on engagement suggests two modifications for future studies. First, the near-win manipulation should be decoupled completely from student performance so that the student can get a better sense of the rate of near-wins. Second, the ease-out manipulation should be complimented with additional effects, such as sound or other animations, to accentuate the buildup of anticipation.

While our procedure did not prove anything beyond what was already shown in Lomas' study—that near-wins increase engagement—we believe that our primary contribution here is the novel methodology. Being able to assess a casual relationship between nearly-winning and engagement not only helps inform educational application design, but it can also be used to control engagement. For instance, one possible extension of this work is to include an adaptive near-win controller that manipulates the score based on a student model of engagement. If the model predicts that the student will not replay if they experience a near-win, then maybe it is a good idea to *reduce* the score so that the student is more likely to play.

Acknowledgments

The authors would like to thank Krista Marks, Bill Troxel, and Adam Holt from Woot Math for their help and cooperation in conducting this study.

Bibliography

[1] R. Reid, "The psychology of the near miss," *Journal of Gambling Studies*, vol. 2, no. 1, pp. 32–39, 1986.

[2] K. A. Harrigan, K. Collins, M. J. Dixon, and J. Fugelsang, "Addictive gameplay: What casual game designers can learn from slot machine research," in *Proceedings of the International Academic Conference on the Future of Game Design and Technology*. ACM, 2010, pp. 127–133.

[3] J. H. Brockmyer, C. M. Fox, K. A. Curtiss, E. McBroom, K. M. Burkhart, and J. N. Pidruzny, "The development of the game engagement questionnaire: A measure of engagement in video game-playing," *Journal of Experimental Social Psychology*, vol. 45, no. 4, pp. 624–634, 2009.

[4] M. M. Khajah, B. D. Roads, R. V. Lindsey, Y.-E. Liu, and M. C. Mozer, "Designing engaging games using bayesian optimization," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2016, pp. 5571–5582.

[5] J. Chen, "Flow in games (and everything else)," *Communications of the ACM*, vol. 50, no. 4, pp. 31–34, 2007.

[6] D. I. Cordova and M. R. Lepper, "Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice." *Journal of educational psychology*, vol. 88, no. 4, p. 715, 1996.

[7] P. Denny, "The effect of virtual achievements on student engagement," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '13. New York, NY, USA: ACM, 2013, pp. 763–772. [Online]. Available: http://doi.acm.org/10.1145/2470654.2470763

[8] B. Katz, S. Jaeggi, M. Buschkuehl, A. Stegman, and P. Shah, "Differential effect of motivational features on training improvements in school-based cognitive training," *Frontiers in Human Neuroscience*, vol. 8, pp. 1–10, 2014.

[9] W. K. Kirchner, "Age differences in short-term retention of rapidly changing information," *Journal of experimental psychology*, vol. 55, no. 4, p. 352, Apr 01 1958, last updated - 2013-02-21.

[10] A. Denisova and P. Cairns, "The placebo effect in digital games: Phantom perception of adaptive artificial intelligence," in *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*. ACM, 2015, pp. 23–33.

[11] J. I. Kassinove and M. L. Schare, "Effects of the" near miss" and the" big win" on persistence at slot machine gambling." *Psychology of Addictive Behaviors*, vol. 15, no. 2, p. 155, 2001.

[12] J. D. Lomas, "Optimizing motivation and learning with large-scale game design experiments," Unpublished Doctoral Dissertation, HCI Institute, Carnegie Mellon University, November 2014.

[13] Wootmath, "Woot math - engaging, research-based tools for the math classroom," aug 2017. [Online]. Available: https://www.wootmath.com/