How should we incentivize learning? An optimal feedback mechanism for educational games and online courses

Lin Xu
Rationality Enhancement Group, MPI for Intelligent Systems, Tübingen, Germany
Maria Wirzberger
Rationality Enhancement Group, MPI for Intelligent Systems, Tübingen, Germany
Falk Lieder
Rationality Enhancement Group, MPI for Intelligent Systems, Tübingen, Germany
Bernstein Center for Computational Neuroscience, Tübingen, Germany

Abstract

There are plenty of opportunities for life-long learning but people rarely seize them. Game elements are an increasingly popular tool to keep students engaged in learning. But gamification only works when it is done properly. Here, we introduce the first principled approach to gamifying learning environments. Our feedback mechanism rewards students' efforts and study choices according to how beneficial they are in the long run. The rewards are conveyed by game elements that we call "optimal brain points". In our experiment, these optimal brain points significantly increased the proportion of participants who attempted to learn a difficult skill, persisted through failure, and succeeded to master it. Our method provides a principled approach to designing incentive structures and feedback mechanisms for both educational games and online courses. We are optimistic that this can help people overcome the motivational obstacles to self-directed life-long learning.

Keywords: gamification; artificial intelligence in education; persistence; educational games; incentive structures

Introduction

As the technological development accelerates, self-directed life-long learning is becoming critically important. Massive Open Online Courses (MOOCs) and other digital resources provide unprecedented opportunities for life-long learning. However, only about 15% of the students who enroll in a MOOC actually finish it (Jordan, 2019). One of the reasons might be that learning something new often requires confronting one's own incompetence and persisting through several failed attempts to understand a new concept or do something new. Many people tend to irrationally avoid such hardships (Urdan & Midgley, 2001; Baker et al., 2008) even though they are often necessary to master new skills (Ericsson, Krampe, & Tesch-Römer, 1993). People who have become experts in using an outdated tool by doing the same work in the same way for many years may be especially resistant to learning how to use a new tool because in the short-run it is much more comfortable for them to exploit their outdated expertise than to become a novice again.

When students are given choices in online courses or educational software they sometimes procrastinate on learning something new by repeatedly practicing skills they already know (Baker, Corbett, & Koedinger, 2004; Mostow et al., 2002).

To help student's overcome such motivational obstacles, educational software increasingly relies on game elements, such as points, levels, and badges, to encourage continued engagement with the learning materials (Kapp, 2012; Dicheva, Dichev, Agre, & Angelova, 2015; Huang & Soman, 2013). The trend of gamification has outpaced the development of an adequate theoretical foundation, and it has been noted that gamification is often ineffective and sometimes even harmful (Toda, Valle, & Isotani, 2018). This raises the question how the incentive structures of digital learning environments such as educational games and online courses should be designed to optimally incentivize good study choices and effective learning strategies.

The points students receive in educational games usually convey performance feedback. But making performance feedback more gameful does not address the fundamental problem that – in the short run – performance feedback might discourage trying to learn something new. Rather, by making student's failures more salient to them, gamified performance feedback can have a negative effect on their study choices – thereby making things worse rather than better (Shute, 2008). O'Rourke, Haimovitz, Ballweber, Dweck, and Popović (2014) argue that to address this problem, gamification should give students "brain points" that reward effort and persistence rather than performance. In support of this view, they found that incentivizing students' effort and learning strategies in an educational game significantly increased their persistence and the total amount of time they spent in the game. However, the hand-crafted incentive system was imperfect and could be exploited by discovering easy ways to earn brain points without doing the hard work of learning a new skill (O'Rourke, Peach, Dweck, & Popovic, 2016). The high prevalence of students "gaming the system" across many intelligent tutoring systems (Baker et al., 2008) underlines that designing good incentives by hand is hard and fallible. This illustrates the deeper issue that we lack a principled theory for designing reward structures in learning systems that incentivize learning properly.

Recent work has begun to establish such principles in the domain of decision-support (Lieder & Griffiths, 2016; Lieder, Chen, Krueger, & Griffiths, 2019). There, the basic idea of this approach is to align the immediate reward of each decision with its long-term value. This addresses the problem that people's decisions are usually overly swayed by the anticipated immediate outcomes (e.g., the unpleasantness of struggling with a difficult math problem vs. the fun of watching a YouTube video) rather than their long-term consequences (e.g., the benefits of a good education). This, so called, *present bias* (O'Donoghue & Rabin, 1999) manifests in a wide range of sub-optimal, short-sighted decisions and problems such as impulsivity and procrastiation that have been explained in terms of hyperbolic discounting and temporal motivation theory (Steel & König, 2006; Steel, 2007).

Considering people's present bias, the real world is far from being an optimal learning environment because the immediate reward for practicing a new skill is usually failure and negative feedback – when it should be something much more positive, namely the value of learning. Conversely, neglecting skill development in favor of exploiting existing skills is usually rewarded because it leads to higher immediate productivity. This suggests that the present bias could be one of the major reasons why students often quit studying too soon or procrastinate on learning a difficult skill – especially when this requires persisting through a series of failed attempts. This suggests that the optimal gamification approach developed by Lieder et al. (2019) might also be applicable to support students' study choices in MOOCs and educational games.

Here, we leverage the framework of optimal gamification (Lieder & Griffiths, 2016; Lieder et al., 2019) to develop a formal mathematical theory of optimal incentives for self-directed learning and an automatic method for computing such incentives from basic assumptions about the skills to be learned and the process of skill acquisition. To achieve this, we develop a mathematical model of the value of practice and apply optimal gamification to it. Our method can be used to automatically compute *optimal* brain points that encourage learning behaviors that are consistent with the growth mindset that the intervention by O'Rourke et al. (2014) was meant to encourage. We postulate that optimal brain points can not only increase the amount of time students invest into learning, as has been demonstrated for hand-designed brain points (O'Rourke et al., 2014), but also their learning outcomes. We test this prediction in a behavioral experiment that simulates a scenario where people have to choose between exploiting their old skill (Skill 1) or learning a new skill (Skill 2) that would allow them to solve a recurring task more efficiently.

We found that participants incentivized with optimal brain points were less likely to give up on trying to learn a new skill, became more likely to master it, and consequently performed better at their tasks. This suggests that our method for computing optimal brain points can help us overcome the pitfalls of incentivizing students study choices manually.

These findings suggest that our principled approach to incentivizing skill acquisition can help people overcome the motivational challenges of self-directed learning and could be used to make educational games and online courses more effective and to avoid the pitfalls of previous attempts to gamify education. Optimal brain points are a principled way to incentivize good study choice and might be able to help students develop a growth mindset (Dweck, 2008).

The plan for this paper is as follows: We first derive the long-term value of practicing a new skill using a simple model of skill acquisition. Next, we translate the value of practice into an optimal gamification method for encouraging skill acquisition. We then evaluate the efficacy of this method in a behavioral experiment mimicking the motivational obstacles to life-long learning. We conclude with the implications of our findings for designing educational games and directions for future work.

Quantifying the value of practice

When should you complete a task using the skills you already have and when should you try to learn a better way to accomplish it? If you would like to invest into learning a new skill, which one should you pick? And if trying to learn this skill is proving difficult, then how long should you keep trying before you give up and do it in the old, familiar way? To help people make these difficult choices, we derive the value of practicing an unfamiliar skill.

The first step of our derivation postulates a simplistic but general and tractable model of skill acquisition through trial and error. If a task has k potential solutions – only one of which is correct – then the probability of discovering the skill in the first attempt is $\frac{1}{k}$. Conversely, the probability that the first attempt will fail is $\frac{k-1}{k}$. After a failure the probability of success increases to $\frac{1}{k-1}$.

Based on this probabilistic model, we can describe skill acquisition as a Markov Decision Process (Sutton & Barto, 1998)

$$M_{\text{skill}} = \{\mathcal{S} \times \mathcal{D}, \mathcal{A}, \gamma, T, r\} \tag{1}$$

where $\mathcal{A}$ includes one action for each skill, $\mathcal{S}$ is the set of all possible skill levels the learner could attain through practice and $\mathcal{D} \subset \mathbb{N}_0$ denotes how much more work is required to complete the current task. The learner's skill level $\mathbf{s_t} \in \mathcal{S}$ reflects how close they are to having mastered each of n different skills at time t and how likely they are to succeed at the task by using each of those skills in their next attempt. We formalize it by the tuple $(k_1, k_2, \cdots, k_n)$ where k_i is the number of potential ways in which the i^{th} skill might work given what the learner knows so far. The transition matrix T encodes that unsuccessfully attempting skill i decreases k_i by 1 and that discovering how it works sets k_i equal to 1. It also encodes how the successful application of each skill would reduce the amount of work required to complete the task and that unsuccessful attempts do not decrease it. The reward function $r((\mathbf{s_t}, d_t), a_t, (\mathbf{s_{t+1}}, d_{t+1}))$ encodes the immediate effort of using or attempting to learn a skill and the value of completing the current task. For simplicity, we assume that the cost of each action is -1 and add the value of completing the current task when d changes to 0. Finally, $1-\gamma \in [0,1]$ is the probability that the current type of task will become obsolete in the next time step.

Abstracting away the details of how specific skills are acquired makes this model very general and broadly applicable. It can therefore be used to incentivize student effort in any learning context. Our model can either be applied out of the box or tailored to specific learning contexts by measuring how specific learning activities increase the probability that the student will successfully learn a particular skill and plugging the measured probabilities into the model's transition matrix T .

Having modelled the process of skill acquisition as an MDP allows us to leverage standard dynamic programming methods (Sutton & Barto, 1998) to compute the value of practice. For instance, we can apply the value iteration algorithm to compute $V^{\star}((\mathbf{s},d))$ – which is the value of having the skill set s when the current task has difficulty d – and $Q^{\star}((\mathbf{s},d),a)$ which is the value of choosing action a (e.g., trying out a new tool versus reusing an old one). To work out under which conditions it is worthwhile to invest in extending one's skill set, we can then translate these value functions into the value of practice which we define as

$$\mathrm{VOP}\left((\mathbf{s},d),a\right) = Q^{\star}\left((\mathbf{s},d),a\right) - V^{\pi_{\text{stop learning}}}\left((\mathbf{s},d)\right), \quad (2)$$

where $V^{\pi_{\text{stop learning}}}$ is the expected return of the strategy that always exploits existing skills without making any investment into learning new skills.

The simplicity of our model allows us to derive the value of practice analytically for the dilemma of choosing between exploiting a mastered skill and attempting to learn a new skill that would make you more effective. When the value of completing the task is g, the mastered skill achieves it in d time steps, and the to be learned skill could achieve it in 1 time step, then the value of practicing the second skill is

$$\begin{aligned}\mathrm{VOP}\left((k_2,d),a_2\right) = &\frac{1}{k_2}\cdot\left[(g-1)+\gamma\cdot V^{\star}(((1,1),d))\right]\\ &+\left(1-\frac{1}{k_2}\right)\cdot\left[\gamma\cdot V^{\star}((1,k_2-1),d)-1\right]\\ &-V^{\pi_{\text{stop learning}}}\left(((1,k_2),d)\right),\end{aligned} \quad (3)$$

and the value of ceasing to learn and exploiting Skill 1 is

$$V^{\pi_{\text{stop learning}}}\left((\mathbf{s},d)\right) = g\cdot\frac{\gamma^{d-1}}{1-\gamma^{d}} - \frac{1}{1-\gamma}, \quad (4)$$

where $\frac{\gamma^{d-1}}{1-\gamma^{d}}$ is the expected number of times one can complete the task using only Skill 1, and $-\frac{1}{1-\gamma}$ is the expected cumulative cost of using the skill. This allows us to characterize under which conditions it is valuable to invest in learning a new skill and under which conditions it is better to exploit the skills one already has.

As shown in Figure 1, we found that the value of practice decreases with the relative effectiveness of the skill one has already mastered ($\frac{d_2}{d_1}=\frac{1}{d}$ if d_1 and d_2 are the number of time steps it takes to complete the task with Skill 1 vs. Skill 2 respectively in this example), but increases with the expected number of times one will have to perform the task in the future (i.e., $\frac{1}{1-\gamma}$). This means that learning a new skill

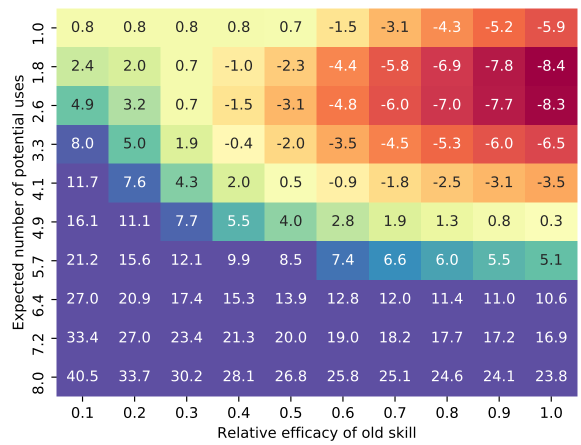

Figure 1: The value of practice. Each square of this heat map shows the difference between the value of attempting to learn a new skill versus exploiting an old skill (colors and numbers) depending on relative efficacy of the old skill (x-axis) and the expected number of occasions on which either skill could be used (y-axis; increasing from top to bottom).

becomes more valuable the more often it might be useful in the future and becomes less worthwhile the more effective the skill is that one has already mastered. By quantifying these effects, the value of practice reveals under which conditions it is worthwhile to learn something new (green-blue) and under which conditions it is better to stick with what one already knows (orange-red). Since the recommendations of our mathematical framework appear to be intuitively correct, we proceed to apply our model of the value of practice to automatically incentivize people's study choices. Furthermore, future work might leverage Equations 2–4 to assist people with decisions about their personal or professional development.

An optimal gamification method for incentivizing skill acquisition

Optimal brain points. Having quantified the value of practice with the skill acquisition MDP defined above, we can now use it to incentivize learning behaviors according to their expected contributions to the learner's competency. Formally, the expected increase in the value of the learner's skill set s achieved by action a is

$$\Delta V(\mathbf{s_t},a) = \gamma\cdot\mathbb{E}[V^{*}(\mathbf{S_{t+1}})|\mathbf{s_t},a] - V^{*}(\mathbf{s_t}),$$

where the random variable $\mathbf{S_{t+1}}$ denotes the learner's skill set after performing action a and $V^{\star}$ is the optimal value function of the skill acquisition MDP defined above. The discount factor γ accounts for the possibility that the practiced skill might become obsolete.

The value of learning by doing is twofold: it increases the value of the learner's skill set (ΔV) and it produces potentially valuable outcomes ($r(\mathbf{s_t},a)$). Our optimal brain points capture both sources of value, that is

$$\text{BrainPoints}(\mathbf{s},a) = \Delta V(\mathbf{s},a) + r(\mathbf{s},a). \tag{5}$$

The way in which optimal brain points are constructed is a direct application of the optimal gamification method developed by Lieder et al. (2019). It satisfies the necessary and sufficient conditions of the *shaping theorem* (Ng, Harada, & Russell, 1999) which thereby guarantees that the resulting incentives do not encourage sub-optimal learning strategies. Rather, by using the value of the learner's skill set ($V^{\star}$) as the basis for constructing the brain points, they are making optimal study choices immediately rewarding. We predict that they should therefore help learners overcome the present bias and invest more in acquiring difficult skills that will benefit them in the future. In the next section, we test this hypothesis with a simple behavioral experiment.

Optimal brain points improve learning and performance

To evaluate the potential of our approach to help people overcome the motivational obstacles to learning new skills, we conducted an online experiment where people repeatedly solve a task and can choose to either solve it using a skill that they already possess (Skill 1) or try to learn a new skill that, once mastered, would allow them to solve the task more efficiently (Skill 2). The experimental group received optimal brain points for their choices between exploiting Skill 1 versus attempting to learn Skill 2 whereas the control condition received no brain points. We predicted that a) most participants in the control condition would neglect investing the time and effort necessary to acquire the new skill – even if their investment in learning would pay off in the long run, and b) that optimal brain points can help them overcome this irrational bias.

Methods

We ran our experiment using psiTurk (Gureckis et al., 2016). We recruited a total of 450 participants from Amazon Mechanical Turk between 15:30 EST and 18:30 EST on January 19, 2019, and we restricted the worker region to the United States of America. Participants received \$0.75 for about 6 ± 2 minutes of work and could earn a bonus of up to \$1 (average bonus \$0.10, standard deviation \$0.10) for their performance in the task. Of our 450 participants, 226 were assigned to the control condition and 224 were assigned to the experimental condition with optimal brain points according to psiTurk's counterbalancing method.

Experimental paradigm. We created the *Spaceship Adventure* game shown in Figure 2 and used it to evaluate the efficacy of optimal brain points. The game world is a board with 6×6 cells. The task for the participants is to control the spaceship so as to move from its initial position (0, 0)

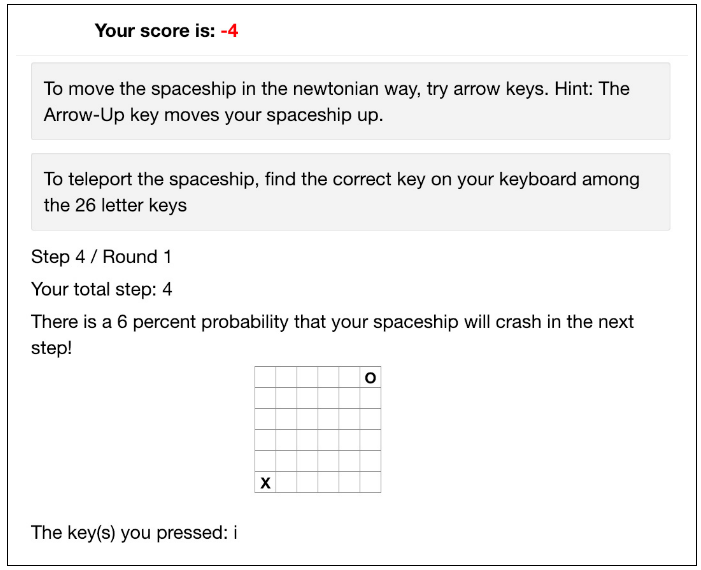

Figure 2: Screenshot of the spaceship game.

to its destination (5, 5). Participants play the game for several rounds. After each time they arrive at the destination, the game board is reset and the spaceship is returned to its initial position.

The instructions inform participants that they will be playing the game for multiple rounds. Participants are also informed that there are two modes of moving the spaceship: The spaceship can be moved one step at a time whereas an unknown letter key could be used to teleport the spaceship directly to its final destination. Each step (using the arrow keys or trying out a new letter key) incurs a cost of -1, whereas reaching the destination earns a reward of $+20$. Following each round there was a 6% chance that the game would end and a 94% chance that it would continue (i.e., $\gamma = 0.94$) and participants were informed about that.. The two skills involved in the game are using the arrow key to move the spaceship forward one square at a time (Skill 1) and teleporting the spaceship directly to the destination using one of the 26 letter keys (Skill 2). For each participant the letter key that would teleport their spaceship was independently selected at random before they started their first round and remained the same until the end of their last round.

In the control condition, the only points being shown were the cost of controlling the spaceship and the reward for reaching the goal. In the experimental condition, participants additionally received the optimal brain points described above. Brain points were given for each of the participant's choices between exploiting Skill 1 versus attempting to learn Skill 2. As illustrated in Figure 3, brain points were conveyed using a color-coded score that was accompanied by the image of a brain. The first time, the participant received brain points, those were explained as conveying the value of learning a new skill. To make the brain points more rewarding, a pleasant crystal sound, which is often used to convey a sense

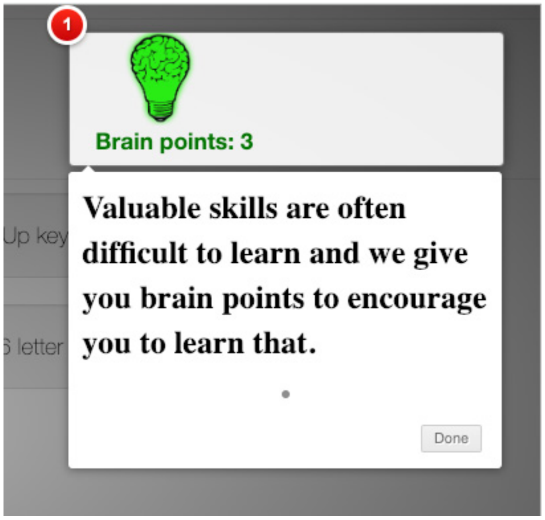

Figure 3: Screenshot illustrating the brain points a participant would receive for trying Skill 2 in the first step.

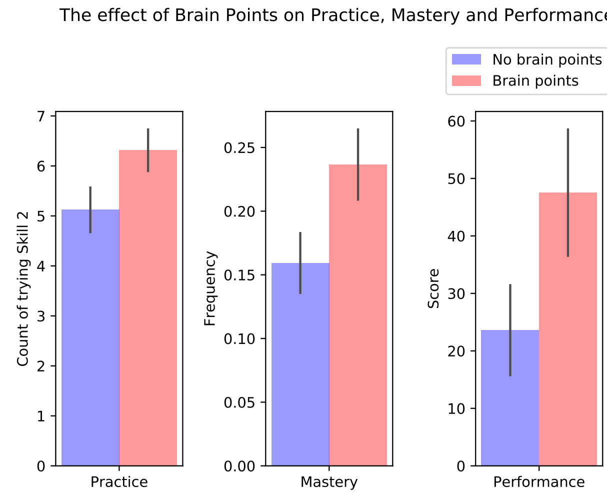

Figure 4: Effect of brain points on practice, learning, and performance. The left bar chart shows the average number of times that people who tried to teleport (skill 2) at least once attempted to figure out how it works until they discovered it or gave up. The middle bar chart shows the proportion of participants in either condition who succeeded to learn skill 2 by discovering which letter key would teleport their spaceship to its target location ("Mastery"). The right bar chart shows the average total score by condition. The error bars represent the standard error of the mean for the bar charts on Practice and Performance and the standard error of the proportion for the Mastery bar chart.

of enlightenment when the players find something valuable in a video game, was played when the number of brain points increased, whereas an unpleasant sound, which could be intuitively perceived as something shrinking, was played when it decreased. Additionally, in both conditions a cheerful sound is played when the spaceship reaches its destination. The brain points score was cumulative as is customary in computer games.

Our code, the experiment, and the data are available on the Open Science Framework at `https://osf.io/k6wjp/`.

Results

As predicted, we found that, when left to their own devices, 42% of the participants never even tried to learn Skill 2 and relied exclusively on Skill 1, although learning Skill 2 could have allowed them to reap higher rewards; that is always attempting Skill 2 would have yielded 154 points on average whereas always exploiting Skill 1 yielded only 8 points on average. This highlights that while there are some situations where people adequately invest into exploring new things (Wilson, Geana, White, Ludvig, & Cohen, 2014), the choice between solving a recurring task with a skill one has already mastered versus using trial-and-error to learn a new skill to be able to handle future occurrences of the task more efficiently might not be one of them for many people.

Encouragingly, we found that optimal brain points significantly increased the proportion of people who attempted to learn the difficult skill (i.e., teleportation, henceforth "Skill 2") from 32% to 46% of participants who had not already tried it in the first step ($\chi^2(1) = 5.74, p = .0165$)[1].

As illustrated in Figure 4, optimal brain points also increased the amount of effort people invested into acquiring Skill 2 from 2.8 to 3.9 attempts on average ($t(448) = 2.52$, $p = .006$; the median number of attempts were 1 and 2 respectively, $Z = 2.59$, $p = .0048$). Furthermore, our optimal brain points also made the people who tried learning Skill 2 at least once more persistent, doubling their median number of additional attempts at learning Skill 2 from 2 to 4 ($Z = 2.49$, $p = .0064$; 4.1 vs. 5.3 on average, $t(448) = 1.86$, $p = .0323$). As a consequence, the proportion of participants who mastered Skill 2 increased from 15% to 24% ($\chi^2(1) = 3.77, p = .0523$), and their average total score doubled from 24 points to 48 points ($t(448) = 1.74, p = .0414$).

These findings suggest that optimal brain points successfully motivated our participants to learn the more difficult skill and thereby improved their learning outcomes and performance.

Conclusion

We derived the expected value of attempting to learn a new skill and translated it into an optimal feedback mechanism for encouraging students to persist in learning valuable skills. Our results suggest that optimal brain points could be useful for helping people overcome the motivational obstacles towards life-long learning. Its basic idea is to reward people's efforts to learn a new skill according to the long-term value of having mastered it and the expected progress towards mas-

[1] We excluded the first action from this analysis because the conditions are identical up until the first feedback is displayed after the participant's first action.

tery.

Our principled computational method for incentivizing learning might become part of the theoretical foundation for the gamification of digital learning environments such as MOOCs or educational games. We hope that the approach illustrated in this article will eventually help people overcome the motivational obstacles that stand in the way of life-long self-directed learning.

Our admittedly simplistic experiment was merely the first step towards evaluating the potential of optimal brain points for increasing student effort. Follow-up experiments should use more naturalistic skill acquisition paradigms and evaluate the proposed feedback mechanism against simpler, heuristic approaches to the gamification of learning environments (Huang & Soman, 2013; Dicheva et al., 2015; Kapp, 2012; O'Rourke et al., 2014). Before we can make any practical recommendations randomized field experiments will have to evaluate our intervention with real students learning real skills.

Future work will evaluate the practical utility of our optimal feedback mechanisms for increasing the student retention rates of MOOCs, encouraging students to use educational games and intelligent tutoring systems more effectively, and building apps that facilitate deliberate practice. These applications may use our method as it is or refine its model of skill acquisition with domain-specific learner models.

While there is a lot of value in being able to motivate students to practice a specific skill inside a digital learning environment, it would be even more valuable if we could help them internalize the value of learning new skills. Future work will therefore investigate whether giving people optimal brain points for their efforts to learn a new skill in one environment can also improve their motivation to learn other skills in different environments and help them develop a growth mindset (Dweck, 2008).

References

Baker, R. S., Corbett, A. T., & Koedinger, K. R. (2004). Detecting student misuse of intelligent tutoring systems. In *International conference on intelligent tutoring systems* (pp. 531–540).

Baker, R. S., Walonoski, J., Heffernan, N., Roll, I., Corbett, A., & Koedinger, K. (2008). Why students engage in gaming the system behavior in interactive learning environments. *Journal of Interactive Learning Research*, *19*(2), 185–224.

Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: A systematic mapping study. *Journal of Educational Technology & Society*, *18*(3).

Dweck, C. S. (2008). *Mindset: The new psychology of success*. Random House Digital, Inc.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363.

Gureckis, T. M., Martin, J., McDonnell, J., Rich, A. S., Markant, D., Coenen, A., ... Chan, P. (2016). psiTurk: An open-source framework for conducting replicable behavioral experiments online. *Behavior Research Methods*, *48*(3), 829–842.

Huang, W. H.-Y., & Soman, D. (2013). Gamification of education. *Research Report Series: Behavioural Economics in Action, Rotman School of Management, University of Toronto*.

Jordan, K. (2019). *MOOC Completion Rates: The Data.* Retrieved from `www.katyjordan.com/MOOCproject.html`

Kapp, K. M. (2012). *The gamification of learning and instruction: game-based methods and strategies for training and education*. John Wiley & Sons.

Lieder, F., Chen, O. X., Krueger, P. M., & Griffiths, T. L. (2019). Cognitive prostheses for goal achievement. *Nature Human Behavior*.

Lieder, F., & Griffiths, T. L. (2016). Helping people make better decisions using optimal gamification. In A. Papafragou, D. Grodner, D. Mirman, & J. Trueswell (Eds.), *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* (pp. 2075–2080). Austin, TX: Cognitive Science Society.

Mostow, J., Beck, J., Chalasani, R., Cuneo, A., Jia, P., Kadaru, K., et al. (2002). A la recherche du temps perdu, or as time goes by: Where does the time go in a reading tutor that listens? In *International conference on intelligent tutoring systems* (pp. 320–329).

Ng, A. Y., Harada, D., & Russell, S. (1999). Policy invariance under reward transformations: Theory and application to reward shaping. In I. Bratko & S. Dzeroski (Eds.), *Proceedings of the 16th Annual International Conference on Machine Learning* (pp. 278–287). San Francisco, CA: Morgan Kaufmann.

O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review*, *89*(1), 103–124.

O'Rourke, E., Haimovitz, K., Ballweber, C., Dweck, C., & Popović, Z. (2014). Brain points: a growth mindset incentive structure boosts persistence in an educational game. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 3339–3348). ACM.

O'Rourke, E., Peach, E., Dweck, C. S., & Popovic, Z. (2016). Brain points: A deeper look at a growth mindset incentive structure for an educational game. In *Proceedings of the Third (2016) ACM Conference on Learning@ Scale* (pp. 41–50).

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153–189.

Steel, P. (2007). The nature of procrastination: a meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin*, *133*(1), 65–94.

Steel, P., & König, C. J. (2006). Integrating theories of motivation. *Academy of Management Review*, *31*(4), 889–913.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Toda, A., Valle, P. H., & Isotani, S. (2018). The dark side of gamification: An overview of negative effects of gamification in education. In A. I. Cristea, I. I. Bittencourt, & F. Lima (Eds.), *Higher Education for All. From Challenges to Novel Technology-Enhanced Solutions.* Cham, Switzerland: Springer Nature.

Urdan, T., & Midgley, C. (2001). Academic self-handicapping: What we know, what more there is to learn. *Educational Psychology Review*, *13*(2), 115–138.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General*, *143*(6), 2074.