ARTICLES
https://doi.org/10.1038/s41562-019-0672-9

nature human behaviour

Cognitive prostheses for goal achievement

Falk Lieder[1*], Owen X. Chen[2], Paul M. Krueger[3] and Thomas L. Griffiths[3,4]

Procrastination takes a considerable toll on people's lives, the economy and society at large. Procrastination is often a consequence of people's propensity to prioritize their immediate experiences over the long-term consequences of their actions. This suggests that aligning immediate rewards with long-term values could be a promising way to help people make more future-minded decisions and overcome procrastination. Here we develop an approach to decision support that leverages artificial intelligence and game elements to restructure challenging sequential decision problems in such a way that it becomes easier for people to take the right course of action. A series of four increasingly realistic experiments suggests that this approach can enable people to make better decisions faster, procrastinate less, complete their work on time and waste less time on unimportant tasks. These findings suggest that our method is a promising step towards developing cognitive prostheses that help people achieve their goals.

While artificial intelligence (AI) is progressing steadily and the computing power of our electronic devices continues to grow, the computing power of the human brain does not. Our bounded cognitive resources continue to constrain our decision-making and often lead to simple heuristics. Previous research has shown that these heuristics can fail miserably in certain scenarios[1–3] but perform very well in the environments they have evolved for[4–8]. These two observations suggest that in the future the human mind could be augmented with cognitive prostheses that use AI to automatically restructure situations in which people's heuristics perform poorly into situations in which those heuristics perform very well.

In line with this vision, previous work has found that human judgement and decision-making can be improved by restructuring how information is presented to people[9–13], and parallel work in operations research and computer science has developed decision-support systems[14,15] that use planning algorithms to solve complex, sequential decision problems for people[16–20]. These approaches have rarely been combined to help people overcome motivational obstacles and achieve their personal long-term goals.

One class of decision problems in which people systematically underperform involves choices whose proximal rewards are misaligned with their long-term value (for example, persevering with a frustrating challenge versus getting drunk and watching TV). In situations like these, people's heuristics tend to reach short-sighted decisions[21–23] that can manifest in procrastination[24] and impulsivity[25]. This apparently myopic nature of human decision-making suggests that decision environments can be repaired by aligning each action's immediate reward with the value of its long-term consequences.

While it is generally difficult to change how people experience the actions necessary to achieve their goals (for example, dieting, debugging or filing taxes) relative to actions that do not (such as eating chocolate or watching TV), it is possible to incentivize those actions with game elements such as points, levels and badges. This approach is known as gamification[26]. Previous research has found that gamification can have positive effects on motivation, engagement, behaviour and learning outcomes[27]. Yet determining which actions should be incentivized and by how much is still an art rather than a science, and misspecified incentives can have devastating consequences[28,29].

Here, we leverage ideas from AI to develop a mathematical framework to help people make more future-minded decisions. The basic idea is to align each action's immediate reward with its long-term value. The resulting system can be interpreted as a cognitive prosthesis that uses AI to solve people's complex sequential decision problems and uses gamification to restructure them in such a way that people can easily identify the course of action that is best for them in the long run. This approach offloads the computational challenges of long-term planning into the reward structure of the environment, and the underlying theory ensures that the added game elements will never incentivize counterproductive behaviour. We evaluate our approach in a series of four increasingly naturalistic experiments starting with controlled proof-of-concept experiments and culminating in longitudinal studies with a naturalistic to-do list app. We find that our optimal gamification method can mitigate the adverse effects of cognitive biases and is more effective at helping people get started on important tasks and waste less time on unimportant tasks than simpler approaches to incentivizing productivity.

An optimal gamification method for decision support

A sequential decision problem can be modelled as a Markov decision process (MDP)

$$M = (\mathcal{S}, \mathcal{A}, T, \gamma, r, P_0) \quad (1)$$

where S is the set of states, A is the set of actions and $T(s, a, s')$ is the probability that the agent will transition from state s to state s' if it takes action a. The discount factor γ can be interpreted as the probability that the decision maker can continue to act and gather more rewards when they arrive in state s'. Setting γ to a value less than 1 thereby captures the possibility that the episode described by the MDP can end early so that future rewards might become unavailable; for instance, $\gamma < 1$ could be used to model the probability that one cannot reap all of the long-term rewards of getting promoted because the company might go bankrupt. The reward generated by this transition is $r(s, a, s')$, and P_0 is the probability distribution of the initial state S_0 (ref. [30]). A policy $\pi: \mathcal{S} \mapsto \mathcal{A}$ specifies which

[1]Max Planck Institute for Intelligent Systems, Tübingen, Germany. [2]Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA, USA. [3]Department of Computer Science, Princeton University, Princeton, NJ, USA. [4]Department of Psychology, Princeton University, Princeton, NJ, USA. *e-mail: falk.lieder@tuebingen.mpg.de

action to take in each of the states. The expected sum of discounted rewards that a policy π will generate in the MDP M starting from a state s is known as its value

$$V_M^{\pi}(s) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t r\,(S_t, \pi(S_t), S_{t+1})\right]. \quad (2)$$

A rational decision maker should follow the optimal policy $\pi_M^{\star}$, which maximizes the expected sum of discounted rewards; that is,

$$\pi_M^{\star} = \arg\ \max_{\pi} \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t r\,(S_t, \pi(S_t), S_{t+1})\right]. \quad (3)$$

People's limited cognitive resources make it impractical for them to always plan out the optimal policy[31]. And even when people know which action would be best in the long term, they do not always exert enough self-control to override their more short-sighted impulses and habits[32]. The value function of the optimal policy satisfies the Bellman equation

$$V_M^{\star}(s_t) = \max_a \mathbb{E}\left[r\,(s_t, a, S_{t+1}) + \gamma V_M^{\star}(S_{t+1})\right]. \quad (4)$$

We can therefore rewrite the optimal policy as

$$\pi_M^{\star}(s) = \arg\ \max_a \mathbb{E}\left[r(s_t, a, S_{t+1}) + \gamma V_M^{\star}(S_{t+1})\right] \quad (5)$$

which reveals that it is myopic with respect to the sum of the immediate reward and the discounted value of the next state. Here, we leverage the MDP framework to model game elements such as points and badges as pseudo-rewards $f(s, a, s')$ that are added to the reward function $r(s, a, s')$ of a decision environment M to create a modified environment $M' = (\mathcal{S}, \mathcal{A}, T, \gamma, r', P_0)$ with a more benign reward function $r'(s, a, s') = r(s, a, s') + f(s, a, s')$ that aligns immediate reward with long-term value.

Designing an incentive system that aligns each action's immediate reward with its long-term value is non-trivial, and misspecified incentives can divert people even farther away from the optimal policy. From the perspective of our formal MDP framework, the problem with misspecified incentives is that they change the optimal policy $\pi_M^{\star}$ of the original decision problem M into a different policy $\pi_{M'}^{\star}$ that is optimal for the gamified environment M' but not for the original environment M. To avoid this problem, we have to ensure that each optimal policy of M' is also an optimal policy of M.

Research on machine learning has identified which conditions pseudo-rewards must satisfy to achieve this: according to the shaping theorem[33] adding pseudo-rewards retains the optimal policies of any original MDP if, and only if, the pseudo-reward function f is potential-based; that is, if there exists a potential function $\Phi: \mathcal{S} \mapsto R$ such that

$$f\,(s, a, s') = \gamma \Phi(s') - \Phi(s) \quad (6)$$

for all values of s, a and s'.

If gamification is to help people achieve their goals, then the pseudo-rewards added in the form of points or badges must not divert people from the best course of action but must make its path easier to follow. Gamification would otherwise lead people astray instead of guiding them to their goals. Hence, the practical significance of the shaping theorem is that it gives the architects of incentive structures a method to rule out incentivizing counterproductive behaviours:

1. Model the decision environment as an MDP.
2. Define a potential function Φ that specifies the value of each state of the MDP.
3. Assign points according to equation (6).

This method could thus be used to avoid some of the dark sides of gamification[28,29].

While the shaping theorem constrains pseudo-rewards to be potential-based, there are infinitely many potential functions one could choose. Given that people's cognitive limitations prevent them from fully incorporating distant rewards[22,23], the modified reward structure $r'(s, a, s')$ should be such that the best action yields the highest immediate reward, that is:

$$\pi_M^{\star}(s) = \arg\ \max_a r'(s, a, s'). \quad (7)$$

Here, we show that this can be achieved with our method by setting Φ to the optimal value function $V_M^{\star}$ of M:

$$\Phi^{\star}(s) = V_M^{\star}(s) = \max_{\pi} V_M^{\pi}(s). \quad (8)$$

First, note that the resulting pseudo-rewards are

$$f\,(s, a, s') = \gamma V_M^{\star}(s') - V_M^{\star}(s) \quad (9)$$

which leads to the modified reward function

$$r'(s, a, s') = r(s, a, s') + \gamma V_M^{\star}(s') - V_M^{\star}(s) \quad (10)$$

Hence, if the agent always maximized its immediate reward without regard for future rewards, then its policy would be

$$\begin{aligned} \pi(s) &= \arg\ \max_a \mathbb{E}\left[r(s, a, s') + \gamma V_M^{\star}(s') - V_M^{\star}(s)\right] \\ &= \arg\ \max_a \mathbb{E}\left[r(s, a, s') + \gamma V_M^{\star}(s')\right]. \end{aligned} \quad (11)$$

According to equation (5), this is $\pi_M^{\star}$ for the original decision environment M. Thus, people would act optimally even if they were completely myopic; this property by itself could be achieved by heuristic incentive structures that always put the largest reward on the best action. However, when this is done heuristically, there is a risk that decision makers could find ways to game the system by collecting the incentives in ways that undermine their intended purpose[34–36]. For instance, a heuristic incentive system (for example, pay proportional to the number of hours worked) that rewards a worker for completing certain tasks (such as reviewing paperwork) could create perverse incentives for them to create unnecessary work for themselves and others (for example, by inventing further bureaucracy to justify their job) even when the incentive system does not reward that directly. By contrast, the optimal incentive structures designed with our method cannot be gamed and are therefore unlikely to mislead decision makers. This safety guarantee is unique to incentive structures that obey the shaping theorem[33], and our optimal gamification method applies this theorem to support human decision-making. This suggests that potential-based pseudo-rewards derived from $V_M^{\star}$ should allow even the most myopic agent that considers only the immediate reward to perform optimally. In this sense, the pseudo-rewards defined in equation (9) can be considered optimal. Optimal pseudo-rewards have also been found to facilitate reinforcement learning in machines[33] and people[37] (F. Lieder et al., manuscript in preparation).

Computing the optimal pseudo-rewards requires perfect knowledge of the decision environment and the decision maker's preferences. This information may be unavailable in practice. Yet even when $V_M^{\star}$ cannot be computed, it is often possible to approximate it. If so, the approximate value function $\hat{V}_M$ can be used to approximate the optimal pseudo-rewards (equation (9)) by

$$\hat{f}\,(s, a, s') = \gamma \hat{V}_M(s') - \hat{V}_M(s). \quad (12)$$

For example, one can estimate the value of s from its approximate distance to a goal[33].

Here we develop and evaluate an approach that leverages AI to help people make better decisions. The basic idea is to automatically restructure decision environments in such a way that people's Pavlovian impulses to collect immediate rewards and avoid immediate losses lead to optimal decisions. To achieve this, our method leverages AI to compute optimal pseudo-rewards and delivers them through game elements. On the basis of previous simulations[33], we predict that adding approximate pseudo-rewards (equation (12)) improves people's decisions and that adding optimal pseudo-rewards is even more beneficial. We test these predictions in three behavioural experiments.

Results

Reward shaping is more effective than heuristic incentive structures at improving people's choices. As a proof of concept, we demonstrate that our mathematical framework for decision support makes it possible to automatically compute optimal incentive structures for a difficult sequential decision problem. In this problem, people often overlook the optimal course of action because they fail to plan beyond intermediate steps that incur a large loss[23]. Experiment 1 showed that the optimal incentives computed by our method allowed people to overcome this bias and perform better than when the incentives were designed heuristically or when no additional incentives were provided.

The sequential decision problem used in this study was the flight planning problem shown in Fig. 1a; it is based on the task that Huys et al.[23] used to demonstrate that people often fail to find optimal plans when this requires looking beyond short-term losses. To help people make better decisions in this task, the incentives computed with our method or heuristic approaches were added directly onto the rewards shown on the arrows in Fig. 1a. In the control condition, people were shown the true transition and reward structure of this task, and their incentives were identical to the task's reward function $r(s, a, s')$ (see Fig. 2a). By contrast, in the experimental conditions, the incentives shown to the participants ($r'(s, a, s')$) differed from the task's true reward function by the pseudo-rewards $f(s, a, s')$ (i.e., $r'(s, a, s') = r(s, a, s') + f(s, a, s')$). We evaluated three kinds of pseudo-rewards: optimal pseudo-rewards (see equation (9) and Fig. 2b), potential-based pseudo-rewards based on an approximate value function (equation (12) and Fig. 2c) and a heuristic incentive system (see Fig. 2d). The approximate value function and the heuristic incentive system both rewarded moving towards the most profitable location (that is, Smithsville in Fig. 2a), but the former satisfied the shaping theorem (equation (6)) whereas the latter violated it. Regardless of the incentives shown to the participants, we measured their performance according to the reward function $r(s, a, s')$ of the original task.

A Kruskal–Wallis analysis of variance (ANOVA) revealed that the type of pseudo-rewards added to the reward function significantly affected people's performance in the original task (Kruskal's $H(3) = 40.35$, $P < 0.001$, $\eta^2 = 0.218$, 95% confidence intervals: [−27.92, −10.42], [−28.75, −9.58], [−14.58, 0.83], [−2.08, 16.25] points per trial for the conditions without pseudo-rewards, with non-potential-based, with approximate pseudo-rewards and with optimal pseudo-rewards, respectively; see Fig. 1b). To elucidate this effect further, we performed a series of two-sided Mann–Whitney *U* tests. As expected, the unaided participants in the control condition performed very poorly, attaining a median loss of 18.75 points per trial (95% confidence interval: [−28.75, −9.58]). Aiding participants with optimal pseudo-rewards led to significantly better performance, enabling them to achieve a median gain of +5.00 points per trial (95% confidence interval: [−2.08, +16.25], standardized *U*-score: $Z = 4.76$, $P < 0.001$, effect size $r = 0.48$, 95% confidence interval on the difference between the medians: [+15.00, +36.25]). Potential-based pseudo-rewards derived from an approximate value function also improved people's performance ($Z = 2.86$, $P = 0.004$, $r = -0.30$, 95% confidence interval of the resulting increase in the median performance: [+3.12, +23.75]) but not as much as optimal pseudo-rewards ($Z = 2.68$, $P = 0.007$, $r = 0.29$, 95% confidence interval of the difference between the median performance with optimal and approximate pseudo-rewards: [1.67, 22.08]). By contrast, the non-potential-based pseudo-rewards failed to improve people's performance ($Z = 0.72$, $P = 0.469$, $r = 0.07$, 95% confidence interval: [−10.42, 12.50]). Optimal pseudo-rewards also accelerated the decision process (Supplementary Results and Supplementary Fig. 15), supporting the conclusion that optimal pseudo-rewards simplify decision problems. Inspecting the four groups' choice frequencies revealed that the optimal pseudo-rewards significantly changed the choice frequencies in each of the six states and successfully nudged participants to follow the optimal cycle: Smithsville → Jonesville → Williamsville → Bakersville → Smithsville (see Supplementary Results and Supplementary Fig. 16).

These results indicate that our mathematical framework makes it possible to automatically compute optimal incentives that can be more effective than manually designed incentive structures. To the extent that game elements such as points act as rewards, these findings suggest that optimal gamification should be safer and more effective than the prevailing heuristic gamification methods. We test this prediction in the following experiments.

Optimal pseudo-rewards can be effectively conveyed by game elements. Having found that optimal pseudo-rewards can substantially improve people's performance, we asked how they should be presented in practice. In Experiment 1, pseudo-rewards were embedded directly into the reward structure of the decision environment. In the real world, this kind of intervention could be implemented by changing tax rates or corporate compensation schemes, for instance. Most of the time such radical changes are beyond the scope of what a choice architect can do, but adding game elements, such as points and badges, is cheap and comparatively easy. Game elements could be used to convey the optimal pseudo-rewards computed with our computational method; here we refer to this idea as 'optimal gamification'. This could be done in many different ways, some of which might be more effective than others. As a first step towards designing optimal gamification apps for decision support in the real world, Experiment 2 evaluated candidate solutions in a more controlled online experiment. We augmented the task from Experiment 1 with game mechanics that conveyed optimal pseudo-rewards through stars and badges (see Fig. 3a and Supplementary Methods). In the first experimental condition, the number of stars awarded for each action was equal to its optimal pseudo-reward (see Fig. 4c). Thus, the player could, in principle, identify the optimal course of action by adding rewards and pseudo-rewards. Because this process can be tedious and prone to errors, the second experimental condition used an alternative presentation format in which stars represented the sum of pseudo-rewards and immediate rewards (integrated pseudo-rewards) so that people would no longer have to perform mental arithmetic to use the pseudo-rewards optimally (see Fig. 4d). We compared people's performance in these conditions with their performance in a control condition without pseudo-rewards (see Fig. 4a) and a condition where optimal pseudo-rewards were added directly to the pay-offs (embedded pseudo-rewards; see Fig. 4b).

A Kruskal–Wallis ANOVA revealed that whether and how pseudo-rewards were presented had a significant effect on people's performance in the task ($H(3) = 29.08$, $P < 0.001$, $\eta^2 = 0.092$, 95% confidence intervals on median performance: [−1.3, +0.4] without pseudo-rewards, [−0.1, +0.9] for embedded pseudo-rewards, [−1.5, +0.0] for separately presented pseudo-rewards and [−0.2, +0.8] for the integrated presentation format). To evaluate the relative effectiveness of these presentation formats, we performed a series of two-sided Mann–Whitney *U* tests. The results of this experiment replicated the finding that when optimal pseudo-rewards

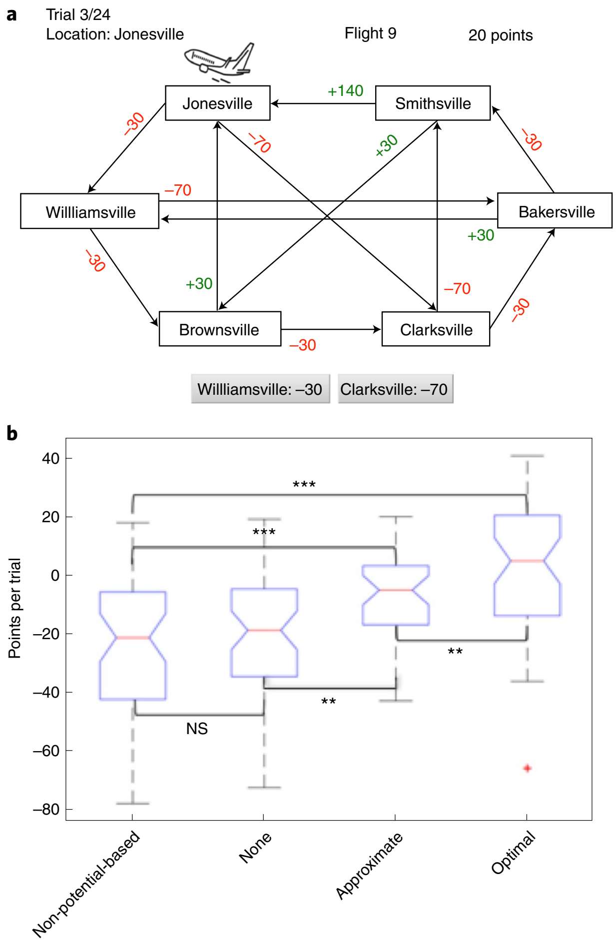

Fig. 1 | Experiment 1. a, Task in Experiment 1: screenshot of the control condition without pseudo-rewards. Participants can move the airplane along the arrows from one airport (e.g., Williamsville) to the next. The number on the arrow indicates the loss or profit of each flight. **b**, Box plots of performance by condition. The red line indicates the median. The blue lines are quartiles. The red crosses denote outliers, and the black whiskers extend to the most extreme data points not considered outliers. The number of stars indicates the significance level according to two-sided Mann–Whitney U tests: $^{*}P<0.05$; $^{**}P<0.01$; $^{***}P<0.001$; NS, not significant). As shown in this figure, optimal pseudo-rewards led to significantly higher performance than no pseudo-rewards (50 participants, $Z=4.76$, $P<0.001$, effect size $r=0.48$, 95% confidence interval on the difference between the medians: [+15.00, +36.25] points/trial), non-potential-based pseudo-rewards ($Z=5.34$, $P<0.001$, $r=0.545$, 95% confidence interval: [+15.42, +37.92]) or approximate pseudo-rewards ($Z=2.68$, $P=0.007$, $r=0.29$, 95% confidence interval: [1.67, 22.08]). Furthermore, approximate pseudo-rewards led to significantly higher performance than no pseudo-rewards ($Z=2.86$, $P=0.004$, $r=-0.30$, 95% confidence interval: [+3.12, +23.75]) or non-potential-based pseudo-rewards ($Z=3.61$, $P<0.001$, $r=0.383$, 95% confidence interval: [+4.17, +24.38]). Non-potential-based pseudo-rewards had no beneficial effect compared to no pseudo-rewards ($Z=0.72$, $P=0.469$, $r=0.07$, 95% confidence interval: [-10.42, 12.50]). These findings are based on data from 186 participants of which 50 received no pseudo-rewards, 47 received optimal pseudo-rewards, 49 received non-potential-based pseudo-rewards and 40 received approximate pseudo-rewards.

are embedded in the task's reward structure, participants achieve a significantly higher median score than participants who receive no pseudo-rewards ($Z=4.53$, $P<0.001$, $\eta^2=0.357$, 95% confidence interval on the difference between the medians: [+0.48, +1.73]). Furthermore, the results suggest that optimal pseudo-rewards can be effectively conveyed by game elements such as stars and badges. Integrated pseudo-rewards presented in the form of stars significantly increased people's performance from −US$0.73 per trial to +US$0.17 per trial ($Z=3.69$, $P<0.001$, $r=0.298$, 95% confidence interval: [+0.42, +1.65]), which was not significantly lower than the performance of the group presented embedded pseudo-rewards (US$0.42 per trial, $Z=0.52$, $P=0.610$, $r=-0.04$, 95% confidence interval: [−0.69, +0.50]). Presenting pseudo-rewards in this integrated format was critical to their effectiveness, since presenting them separately failed to significantly increase people's performance (median performance: −US$0.5 per trial; $Z=0.22$, $P=0.83$, $r=0.017$, 95% confidence interval: [−0.90, +0.88]). Inspecting participants' choice frequencies revealed that the three presentation formats had significantly different effects on people's decisions (see Supplementary Results and Supplementary Fig. 18). In summary, incentivizing good decisions with game elements can be as effective as redesigning the decision environment, and this approach is most effective when the game elements make it very easy for people to identify the best course of action.

Optimal gamification helps people overcome procrastination. Given that optimal gamification enabled the participants of Experiment 2 to act more far-sightedly, we hypothesized that this approach might be able to alleviate the myopic biases that give rise to procrastination in everyday life. To test this hypothesis, we designed a more naturalistic experiment in which participants used the to-do list app shown in Fig. 5. In Experiment 3, each participant's to-do list comprised five daunting writing assignments. Participants were free to complete as few or as many of those assignments as they wanted and could earn a US$20 bonus by completing all assignments by a distant deadline. The critical experimental manipulation was whether and how the to-do list items were incentivized with game elements. The experiment comprised a control condition where participants were presented with a regular to-do list without game elements and three conditions where participants could earn points and badges for completing the items on their to-do list (see Fig. 5). In one of these conditions, each task was assigned the same number of points; this was considered a second control condition. In the other two conditions, the number of points assigned to each task was computed by optimal gamification (see Fig. 5). The optimal point values displayed next to each to-do item conveyed how much closer completing that task would bring the participant to earning the US$20 bonus. Consequently, more difficult tasks would earn the participant more points than shorter and simpler ones. In one of these conditions, the incentives for completing each assignment were conveyed as points (see Fig. 5), and a participant's total number of points determined their level in the game. In the other condition, the optimal pseudo-rewards were displayed as virtual dollars.

To evaluate the effect of presentation format on the probability that participants would complete all the writing assignments on time, we conducted a series of χ^2 tests. As shown in Fig. 6, optimal pseudo-rewards significantly increased the completion rate from 56.1% in the control conditions to 85.2% in the experimental conditions with optimal pseudo-rewards ($\chi^2(1)=11.20$, $P<0.001$, goodness of fit $w=2.863$, 95% confidence intervals: [41.0%, 75.9%] versus [82.2%, 99.7%]). This benefit cannot be explained by the mere presence of incentives or game elements because adding constant point values failed to increase the completion rate (53.6% with constant points versus 58.6% without points, $\chi^2(1)=0.15$, $P=0.701$, $w=0.384$, 95% confidence intervals: [35.6%, 71.5%] versus [41.0%, 75.9%]). Framing optimal pseudo-rewards in terms of money led to a completion rate of 92.3% while presenting them as points led to a completion rate of 78.6%, but this difference was not statistically

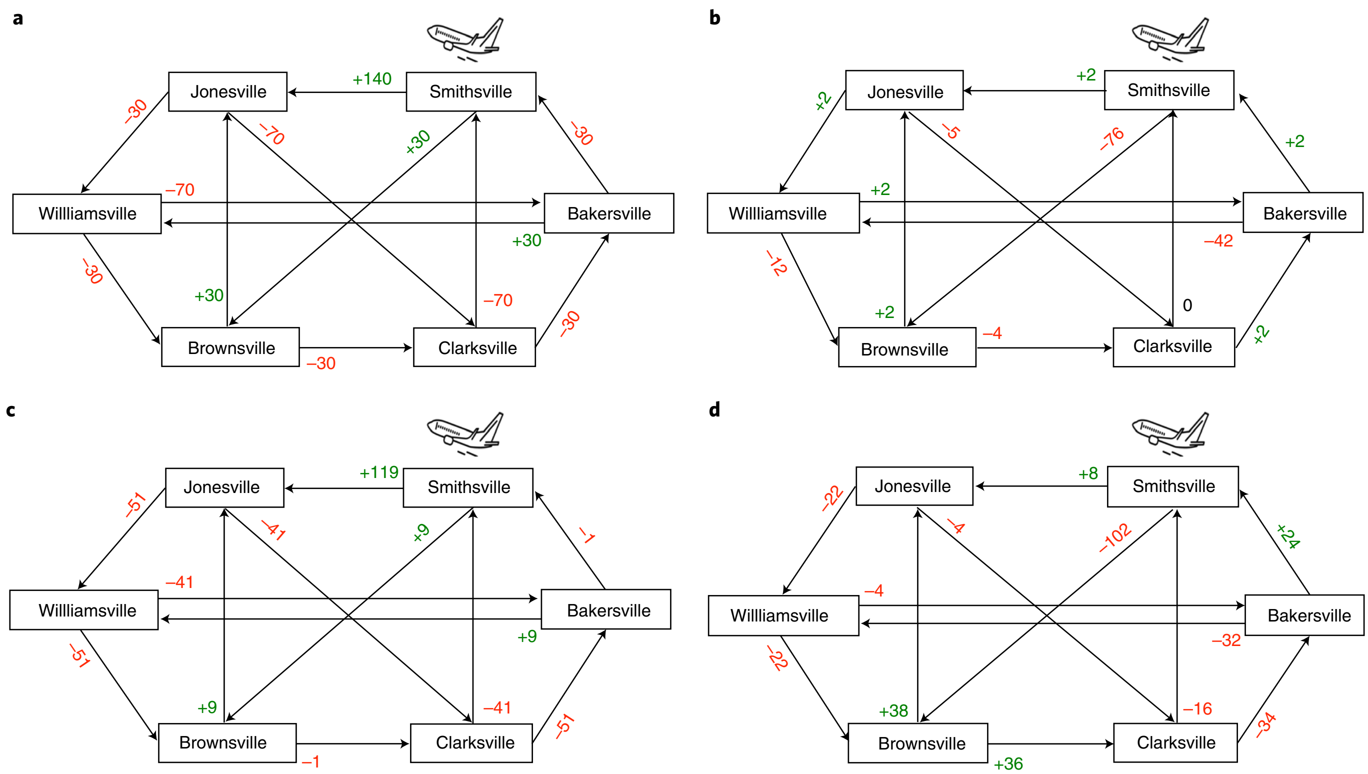

Fig. 2 | Conditions of Experiment 1. a, No pseudo-rewards. **b**, Optimal pseudo-rewards. **c**, Approximate pseudo-rewards. **d**, Non-potential-based pseudo-rewards.

significant ($\chi^2(1)=2.01$, $P=0.156$, $w=1.42$, 95% confidence intervals: [82.2%, 99.7%] versus [63.5%, 92.4%]). Along with the increase in the completion rate, the average number of completed assignments increased from 2.93 out of 5 without optimal gamification to 4.30 out of 5 with optimal gamification according to a two-sided Mann–Whitney U test ($Z=3.38$, $P=0.001$, $r=0.321$, 95% confidence intervals on the means: [2.30, 3.54] without optimal gamification versus [3.81, 4.72] with optimal gamification), and the average total number of words written by each participant significantly increased from 458.4 ± 59.0 without optimal gamification to 765.46 ± 71.45 with optimal gamification according to a Mann–Whitney U test ($Z=3.37$, $P=0.001$, $r=0.320$, 95% confidence intervals on the means: [345.6, 576.1] without optimal gamification versus [634.2, 911.2] with optimal gamification). Further analyses reported in the Supplementary Information suggested that the primary benefit of optimal gamification was to increase the probability that a participant would complete the first task from 59.65% to 87.04% (Kruskal–Wallis ANOVA, $\chi^2(1)=11.01$, $P<0.001$, $\eta^2=0.09$, 95% confidence intervals: [48.8%, 73.7%] versus [80.5%, 96.4%]), because regardless of gamification 95.1% of all participants who completed the first task went on to complete all of the tasks (Kruskal–Wallis ANOVA: $\chi^2(1)=0.69$, $P=0.408$, $\eta^2=0.006$, 95% confidence intervals: [82.2%, 98.9%] without gamification versus [90.2%, 99.9%] with gamification), and their motivation and behaviour seemed to be unaffected.

Optimal gamification helps people prioritize. While Experiment 3 showed that optimal gamification is more effective than assigning the same number of points to every item on one's to-do list, there might be other simple heuristics that are just as effective as our computationally intense optimal gamification method. One intuitive heuristic is that long and difficult tasks should be assigned more points than tasks that are short and easy. On the surface, this is exactly what optimal gamification did in Experiment 3. However, optimal gamification also takes into account the value of the goals that each task contributes to and how much it contributes to that value. Experiment 4 shows that this makes optimal gamification more effective than a simple heuristic that incentivizes each task only according its length (i.e., the number of points for completing task A is \$20 times the length of task A divided by the total length of all tasks combined). Participants were given a to-do list of ten tasks. Five of those tasks belonged to a valuable project (Project 1) whose completion would earn the participant a US\$20 bonus. The other five tasks were equally difficult but belonged to an unworthwhile project (Project 2) whose completion would earn the participant only US\$1. Participants were free to complete either one of the projects, both or neither.

According to a χ^2 test of independence, gamification significantly reduced the proportion of participants who wasted their time by working on the unworthwhile second project($\chi^2(2)=10.85$, $P=0.004$,$w=0.726$, 95% confidence intervals: [19.1%, 46.5%] without gamification, [27.5%, 56.4%] with heuristic gamification and [3.0%, 20.2%] with optimal gamification). Consistent with our prediction, we found that the proportion of participants who worked on Project 2 was lowest in the optimal gamification condition (11.1%), highest in the condition with heuristic gamification (41.9%) and intermediate in the control condition (32.6%). Mann–Whitney U tests confirmed our hypothesis that participants who are supported by optimal gamification perform fewer wasteful tasks of Project 2 (avg. 0.38) than participants supported by heuristic gamification (avg. 1.47; $Z=3.22$, $P=0.001$, $r=-0.343$, 95% confidence intervals on the means: [0.88, 2.09] with heuristic gamification versus [0.04, 0.80] with optimal gamification) and participants in the control condition (avg. 1.00, $Z=2.38$, $P=0.017$, $r=-0.254$, 95% confidence intervals: [0.51, 1.56] without gamification versus [0.04, 0.80] with optimal gamification). Likewise, the proportion

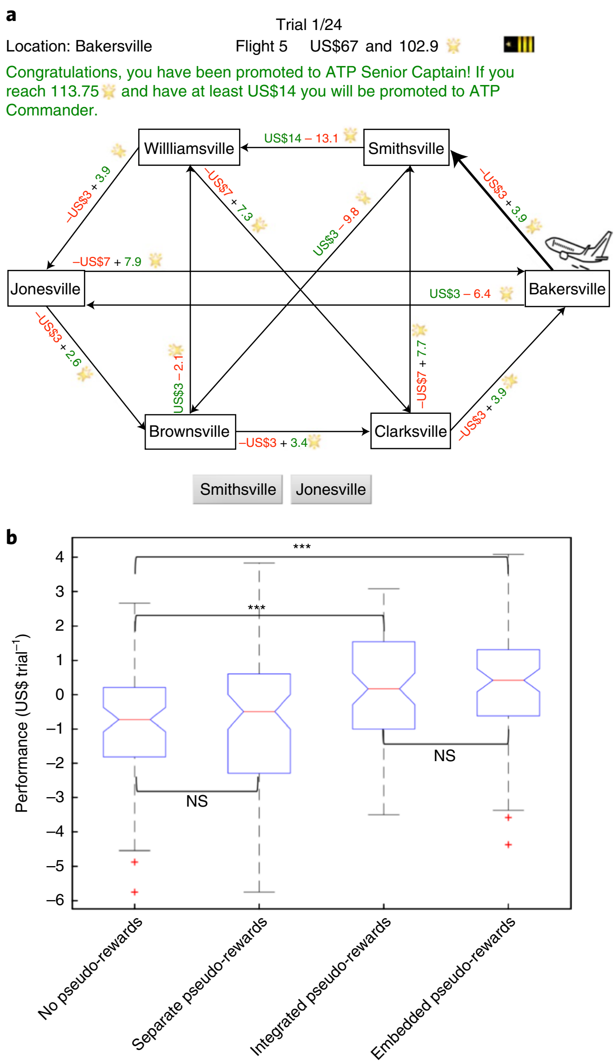

Fig. 3 | Experiment 2. **a**, Screenshot from the pilot game with separately presented pseudo-rewards. **b**, Box plots of performance in Experiment 2 by condition. When game elements conveyed the optimal pseudo-rewards in the integrated format, participants performed significantly better than when they were presented no pseudo-rewards ($Z = 3.69$, $P < 0.001$, $\eta^2 = 0.298$, 95% confidence interval: [+0.48, +1.73] points/trial) or when the optimal pseudo-rewards were presented separately ($Z = 3.21$, $P = 0.001$, $r = 0.258$, 95% confidence interval: [+0.17, +1.83]), and not significantly better or worse than when optimal pseudo-rewards were embedded directly into the environment's reward structure ($Z = 0.52$, $P = 0.610$, $r = -0.04$, 95% confidence interval: [-0.69, +0.50]). By contrast, separately presented pseudo-rewards did not significantly improve people's performance compared to the control condition ($Z = 0.22$, $P = 0.829$, $r = 0.017$, 95% confidence interval: [-0.90, +0.88]) and led to significantly lower performance than embedded pseudo-rewards ($Z = -3.63$, $P < 0.001$, $r = -0.285$, 95% confidence interval: [-1.96, +0.25]). As in Experiment 1, embedding optimal pseudo-rewards into the task's reward structure significantly improved participants' performance relative to the control condition ($Z = 4.53$, $P < 0.001$, $\eta^2 = 0.357$, 95% confidence interval: [+0.48, +1.73]). All statistical comparisons reported in this figure are two-sided Mann–Whitney U tests. These findings are based on data from 316 participants (about 80 in each condition).

of participants who completed the unworthwhile project was highest in the condition with heuristic gamification (20.9%), lowest in the optimal gamification condition (10.0%) and intermediate in the control condition (14.0%), although these differences were not statistically significant according to a χ^2 test of independence ($\chi^2(2) = 3.78$, $P = 0.151$, $w = 1.943$, 95% confidence intervals: [4.6%, 24.3%] without gamification, [9.6%, 33.1%] with heuristic gamification and [0.8%, 13.9%] with optimal gamification). These findings suggest that optimal gamification can not only help people get started on important tasks, as shown in Experiment 3, but can also help them avoid wasting time on unimportant ones.

Consistent with the findings in Experiment 3, an increasing trend in the completion rate of Project 1 from the control condition (34.9%) to the optimal gamification condition (42.2%) via the heuristic gamification condition (37.2%) pointed in the predicted direction, but these differences were not statistically significant according to a Kruskal–Wallis ANOVA ($\chi^2(2) = 0.53$, $P = .769$, $w = 0.726$, 95% confidence intervals: [21.2%, 49.0%] without gamification, [23.2%, 51.5%] with heuristic gamification and [28.2%, 56.5%] with optimal gamification); we think this is primarily because the addition of five more writing assignments drastically reduced the overall completion rate in all conditions, thereby reducing the statistical power of our experimental design. See the Supplementary Results for more detail on how optimal gamification affected people's engagement with Project 1.

According to a one-way ANOVA, gamification also had a significant effect on participants' self-reported motivation to complete the tasks ($F(2, 46) = 3.35$, $P = 0.044$, $\eta^2 = 0.127$, 95% confidence intervals: [3.94, 5.22] without gamification, [4.83, 5.69] with heuristic gamification and [5.91, 6.51] with optimal gamification) with motivation higher in the optimal gamification condition than in the control condition ($t(29) = 2.58$, $P = 0.015$, $d = 0.951$ according to two-sided t-test) and potentially higher than in the heuristic gamification condition ($t(35) = 1.82$, $P = 0.077$, $d = 0.598$).

Contrary to the concern that people might experience gamification as a manipulative intrusion on their freedom of choice, a series of one-way ANOVAs revealed that gamification had no effect on the level of autonomy ($F(2, 46) = 1.57$, $P = 0.219$, $\eta^2 = 0.127$, 95% confidence intervals: [4.74, 5.65] without gamification, [4.56, 5.38] with heuristic gamification and [5.53, 6.36] with optimal gamification) or perceived intrusion ($F(2, 46) = 0.27$, $P = 0.761$, $\eta^2 = 0.012$, 95% confidence intervals: [1.55, 2.37] without gamification, [1.15, 1.85] with heuristic gamification and [1.19, 2.10] with optimal gamification) that our participants experienced. To the contrary, in our sample participants supported by optimal gamification gave the highest median rating of autonomy (6.2 versus 5.1 in the control condition and 4.9 in the heuristic condition) and the lowest median rating of perceived intrusion (0.50 versus 1.75 in the control condition and 1.50 in the heuristic gamification condition).

Discussion

The results of Experiments 1–4 suggest that optimal gamification can help people make better decisions and act more far-sightedly (Experiments 1 and 2), get started on daunting tasks (Experiment 3) and waste less time on unimportant tasks (Experiment 4). These findings jointly show that aligning each action's immediate reward with its long-term value succeeds in improving people's choices when heuristic gamification methods fail. We found that optimal gamification is more effective than uniformly incentivizing progress towards a goal (Experiment 1), assigning the same number of points to each behaviour the designer wants to encourage (Experiment 3) or making the number of points proportional to how long a task takes (Experiment 4). Our optimal gamification method achieves this alignment between immediate reward and long-term value by leveraging AI to solve sequential decision problems that are challenging for people and translating the solutions into incentives that align each action's immediate reward with its long-term value. The resulting incentive structures are implemented using game elements such as points and levels that motivate people to do what is best for them in the long run. While each of the heuristics we evaluated

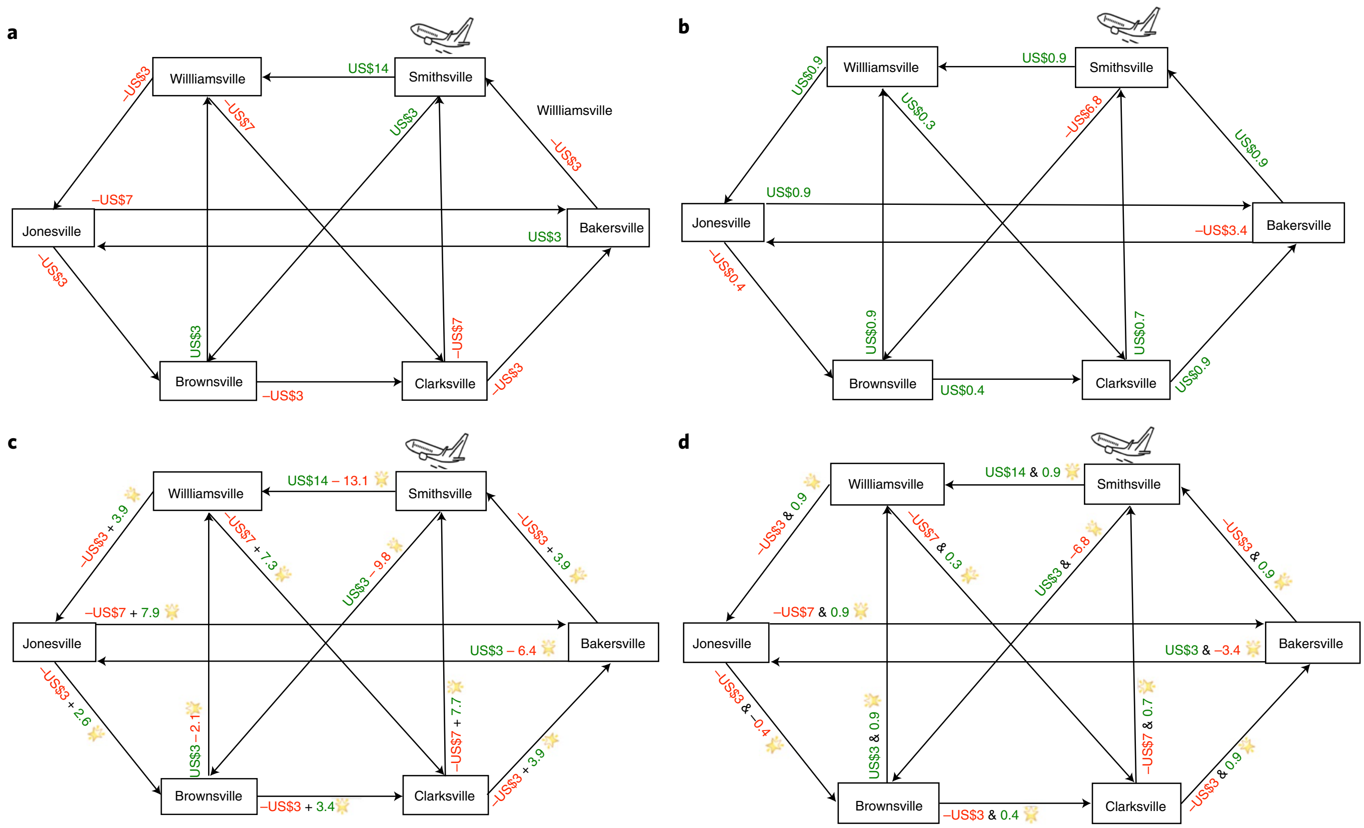

Fig. 4 | Conditions of Experiment 2. a, Control condition. **b**, Embedded pseudo-rewards. **c**, Separate pseudo-rewards. **d**, Integrated pseudo-rewards.

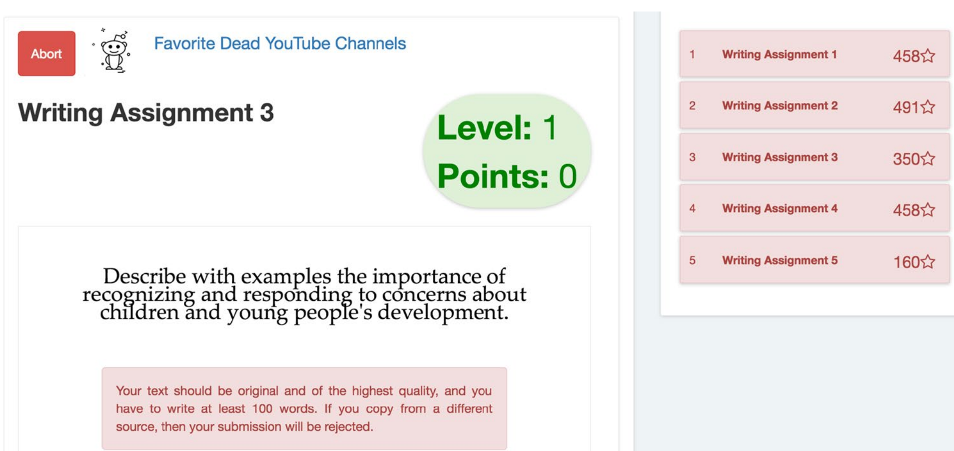

Fig. 5 | Screenshot from Experiment 3. Screenshot.

optimal gamification against may work in some cases, our results suggest that optimal gamification is more reliable across a wider range of potential applications.

More generally, our results illustrate that AI can be used to automatically restructure decision problems in such a way that people's heuristics work well. This approach is in line with an extensive literature on bounded rationality that emphasizes that decision quality depends on the fit between people's heuristics and the structure of their environment[4–8,38]. While optimal gamification accommodates the myopic nature of many heuristics, AI can also be leveraged to adapt the way in which decision problems are presented to other characteristics of heuristic decision-making. For instance, one could accommodate people's tendency to select the first option they find good enough[38] by leveraging AI to sort the alternatives in descending order of their predicted value to the decision maker. This approach could be combined with optimal gamification to help people choose the option that is best in the long term even when there are too many alternatives to be considered and many of them are decent. Furthermore, an intelligent personal assistant could try to ensure that the default decision is always an action with high expected utility rather than the status quo[12].

There are already many decision support systems that solve MDPs to compute optimal decisions and advise people to execute them[16–20]. However, research in psychology suggests that this

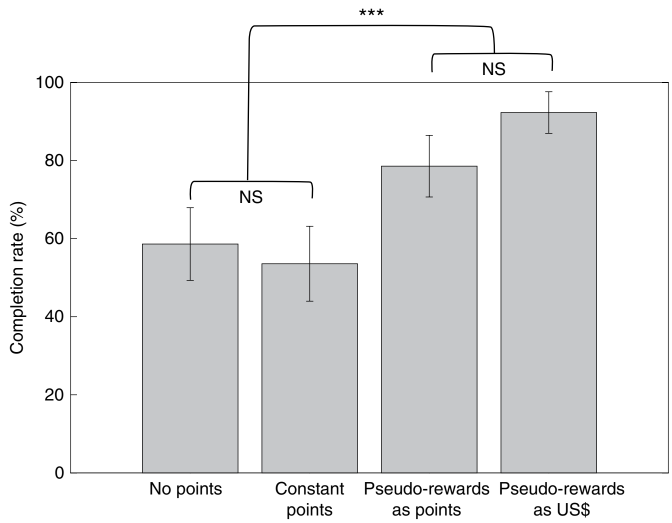

Fig. 6 | Proportions of participants in Experiment 3 who completed all assignments by the deadline. Error bars show ±1 s.e.m. The completion rate was significantly higher when participants were incentivized by optimal gamification than when all tasks were incentivized equally or no points were provided ($\chi^2(1) = 11.20$, $P < 0.001$, $w = 3.35$, 95% confidence intervals: [67.6%, 91.5%] without optimal gamification, [75.7%, 93.9%] with optimal gamification). The difference between the conditions with no points and those with constant point values was not statistically significant ($\chi^2(1) = 0.15$, $P = 0.701$, $w = 0.384$, 95% confidence intervals: [35.6%, 71.5%] versus [41.0%, 75.9%]). The difference between framing optimal pseudo-rewards in terms of points and framing them in terms of money was not statistically significant ($\chi^2(1) = 2.01$, $P = 0.156$, $w = 1.42$, 95% confidence intervals: [82.2%, 99.7%] versus [63.5%, 92.4%]). All statistical tests reported in this figure are χ^2 tests. These results are based on 111 participants (about 28 per condition).

approach to decision support would probably undermine people's intrinsic motivation because it runs counter to the fundamental human need for self-determination and autonomy[39]. Optimal gamification, by contrast, gives people complete freedom over what to do and can be applied to help people motivate themselves to take action towards their own goals. Using game elements to boost motivation is not a new idea, but optimal gamification is based on a rigorous mathematical theory for determining which actions should be incentivized and by how much. This theory guarantees that optimal gamification will never incentivize counterproductive behaviour[33]. Here we have shown that it can be leveraged to avoid the perils of less principled approaches to motivating people with incentives and game elements[28,29].

The results of Experiment 3 suggest that optimal gamification might be able to help people align their actions (for example, whether or not to work on a writing assignment) with their long-term goal (for example, to earn a US$20 bonus). The primary problem that optimal gamification solved in this setting was to help people overcome the motivational barriers of immediate effort that would be rewarded much later. This suggests that optimal gamification might be useful for helping people overcome the myopic biases affecting their motivation[40], avoid self-control failure and support the pursuit of long-term goals. While Experiment 3 presented pseudo-rewards numerically, future work will investigate whether optimal gamification can also be effectively implemented using other presentation formats that are more commonly used in practice.

Beyond motivational issues, many decision problems that arise in the pursuit of long-term goals are simply too large and too complex for people to solve them optimally. Our approach could be used to overcome such challenges by augmenting people's bounded cognitive resources with the power of computing and leveraging planning algorithms developed in AI[41] to build the solution of complex decision problems into the reward structure of the environment. Future work will investigate these hypotheses and explore optimal gamification as an interface between artificial and human intelligence. By integrating the power of computing with psychological insight into human motivation and decision-making, this line of research could lead to cognitive prostheses that might substantially enhance human productivity and self-mastery. Our approach illustrates how advances in AI can be leveraged to help people make better decisions. In this way, the continuing progress in AI could enable a parallel growth in human effectiveness. Given that people learn to adapt their planning horizon to the structure of the environment[42], optimal gamification could have adverse effects on how people make decisions in the absence of game elements. A preliminary follow-up experiment reported in the Supplementary Information suggests that this might not be the case when the environment remains the same, but transfer effects from a gamified environment to decision-making in other environments remain to be investigated. Furthermore, we agree with earlier studies[43,44] that nudges and decision-support systems should be complemented by improving people's decision-making competencies, and ongoing work suggests that AI and optimal gamification can be combined in a more sophisticated way to teach people better decision-making strategies (Lieder, Callaway, Das, Jain, Gul, Krueger, & Griffiths, manuscript in preparation)[37].

Methods

The experiments reported in this Article were approved by the Committee for the Protection of Human Subjects of the University of California, Berkeley under protocol number 2016-02-8359 (To-Do-List Gamification), and Experiment 4 was also approved by the Independent Ethics Council of the medical faculty of the University of Tübingen as IEC Project Number 668/2018BO2 (To-Do-List Gamification).

Experiment 1. On the basis of our experience with similar online experiments, we recruited 200 adult participants on Amazon Mechanical Turk to obtain a sample size of about 60 participants per condition. No statistical methods were used to predetermine sample sizes, but our sample sizes are larger than those reported in previous publications[23,45]. Participants received US$0.50 and a performance-dependent bonus of up to US$2 for playing 24 rounds of the flight planning game shown in Fig. 2. In this game, the player receives points for routing an airplane along profitable routes between six cities. In each round, the initial location of the airplane is chosen at random. Participants then choose which of two possible destinations to fly to, receive the profit or loss of that flight, and choose the next flight until the round ends. After each flight there was a 1 in 6 chance that the round would end. Participants were instructed to score as highly as possible, and their financial bonus was proportional to the rank of their score among all participants in their condition. This game is based on the planning task that Huys et al.[23] used to demonstrate how cognitive limitations shape human planning and induce cognitive biases. We modelled our version of this task as an infinite horizon MDP with a discount factor of $\gamma = 1-1/6$ that models the probability that the current round will end after each move (see Supplementary Methods). This MDP model of the task is defined such that maximizing the expected sum of discounted rewards over an infinite horizon was equivalent to maximizing the expected sum of points the player earns until the game ends.

Participants were randomly assigned to one of four conditions (see Fig. 2 and Supplementary Table 1). In the control condition, there were no pseudo-rewards (Fig. 2a), and finding the optimal path required planning four steps ahead. In the three experimental conditions, the reward structure was modified by adding pseudo-rewards that summed to zero. In the first experimental condition, the pseudo-rewards were derived from the optimal value function according to the shaping theorem (see equation (9); Fig. 2b). In this condition, looking only one step ahead was sufficient to find the optimal path. This makes it possible for people to make optimal decisions by simply following the Pavlovian impulses that draw them towards immediate gains and push them away from immediate losses without having to engage in any planning or self-control. The second experimental condition used potential-based pseudo-rewards based on an approximate value function (equation (12)). This approximate value function (equation (3) in the Supplementary Methods) was designed by identifying the most profitable location (that is, Smithsville in Fig. 2a) as a goal and then scoring each location by its distance to that goal. The resulting pseudo-rewards were positive for actions that brought the airplane closer to the goal and negative for actions that moved it away from the goal. The resulting pseudo-rewards simplified planning but not as much as the optimal pseudo-rewards. Finding the optimal path required planning two or three steps ahead, and the immediate losses were smaller. In the third experimental condition (Fig. 2d), the pseudo-rewards were also designed to encourage participants to move towards the most profitable location. The heuristic pseudo-reward was +50 for each transition that reduced the distance to the most valuable state (Smithsville). But unlike the approximate potential-based pseudo-rewards, they imposed no penalty for moving away from the goal. The resulting pseudo-rewards violated the shaping theorem and were slightly too high relative to the true rewards. As a consequence, they incentivized taking a bad short cut to Smithsville (Jonesville → Clarksville → Smithsville) and rendered the optimal path (Jonesville → Williamsville → Bakersville → Smithsville) suboptimal in the gamified environment. This made it impossible for participants to recognize the best path as optimal even with extensive learning or perfect long-term planning. To ensure that all conditions were comparable in terms of the total reward, the pseudo-rewards of each condition were shifted such that their sum was zero. Since the experimental manipulation only affected the flights' pay-offs, participants were unaware of the pseudo-rewards in Experiment 1.

Inclusion criteria. The average completion time of the experiment was 13.6 min, and the median response time was 1.3 s per choice. The median of our participants' relative scores (that is, $(R-r_{\min})(r_{\max}-r_{\min})$ where R is the sum total of the player's points) was 79%. We excluded 3 participants who invested less than one-third of the median response time of their condition and 11 participants who scored lower than 95% of all participants in their condition (5.5%), leading to a total exclusion rate of 7% (14/200). Of the 186 participants included in the analysis, 50 were in the condition without pseudo-rewards, 47 were in the condition with optimal pseudo-rewards, 40 were in the condition with approximate pseudo-rewards and 49 were in the condition with non-potential-based pseudo-rewards.

This exact experiment was run only once, but its main finding was replicated in Experiment 2. Data collection and analysis were not performed blind to the conditions of the experiments.

Experiment 2. Since we expected the effect of varying the format in which optimal pseudo-rewards were presented to be smaller than the effect of presenting optimal pseudo-rewards at all, we recruited 100 participants per condition (that is, 400 participants in total) on Amazon Mechanical Turk. No statistical methods were used to predetermine sample sizes, but our sample sizes are larger than those reported in previous publications[23,45]. We paid our participants US$2.50 for about 20–25 min of work plus a performance-dependent bonus of up to US$2. The average value of the bonus was US$1. The median completion time of the experiment was 21.2 min.

The task was equivalent to the one used in Experiment 1 except that all rewards were scaled down by a factor of 10 to keep the arithmetic operations required to solve the task simple. Participants were randomly assigned to one of four conditions. In the control condition, no pseudo-rewards were presented (Fig. 4a). Three experimental conditions presented the optimal pseudo-rewards in three different formats. In the first experimental condition, the pseudo-rewards were embedded into the decision environment by adding them directly onto the flights' profits and losses (Fig. 4b). In the second experimental condition, the pseudo-rewards were presented separately from the monetary rewards in the form of stars (Fig. 4c). In the third experimental condition, the number of stars communicated the sum of the shifted optimal pseudo-reward and the immediate reward (Fig. 4d). In the conditions with stars, participants were informed that the stars were designed to help the pilots make better, less short-sighted decisions. The instructions explained the meaning of the stars. In the second experimental condition, participants were told that the difference in the number of stars awarded for flying to destination A versus B predicted the difference in the amount of money that could be earned from there onward in the long run. In addition, these participants were given the tip that the flight with the highest sum of stars plus dollars was most profitable in the long run. In the third experimental condition, participants were told that the difference between the number of stars awarded for flying to destination A versus B predicted the difference in how much profit they were going to make in the long run if they chose destination A over destination B. Participants in this condition were given the tip that they could earn the most by always flying the route with the larger number of stars. In all conditions, each flight's pay-off and number of stars were rounded to one decimal digit. Stars had no monetary value, but they determined the player's level in the game (see Supplementary Methods). Screenshots of the instructions for the experimental conditions are shown in the Supplementary Methods.

The optimal pseudo-rewards presented in the three experimental conditions were computed according to equation (9) and then shifted by a constant such that, on average across all states, the sum of the immediate reward and pseudo-reward for the optimal action was equal to the expected discounted long-run reward of the optimal strategy averaged across all possible starting states. This is appealing because it makes the pseudo-rewards assigned to each action predict how much money players will earn in the long run if they choose that action and then continue optimally.

Attention checks and inclusion criteria. To start the experiment, participants had to pass a quiz comprising three questions on how their financial bonus would be determined and three questions testing their understanding of the mechanics of the task. If participants got one or more questions wrong, they were asked to reread the instructions and retake the quiz until they answered all questions correctly.

Out of the 400 participants, 65 had participated in previous flight planning experiments and were therefore excluded from this study. Out of the 335 remaining participants, we excluded participants whose median response time was less than one-third of the median response time across all included participants. In addition, we excluded the 5% of participants with the lowest scores of each group. These two criteria led to the exclusion of 19 of the 335 included participants (5.7%), leaving us with 316 included participants: 80 in the control condition, 81 in the condition with embedded pseudo-rewards, 81 in the condition with separate pseudo-rewards and 74 in the condition with integrated pseudo-rewards.

This experiment was run only once, but its main finding was replicated in the follow-up study to Experiment 2 presented in the Supplementary Information. Data collection and analysis were not performed blind to the conditions of the experiments.

Experiment 3. We recruited 120 participants by posting a sign-up form on Amazon Mechanical Turk. We aimed for a sample size of 30 participants per condition so that a χ^2 test would be able to detect an increase from an expected completion rate of about 60% in the control condition to about 90% in the condition with monetary pseudo-rewards at $\alpha = 0.05$ with a power of 0.85. According to the exit survey (see below), the ages of our participants ranged from 20 to 68 yr (average: 36.7 yr, standard deviation: 10.6 yr), and our sample was roughly gender balanced (54% of our participants were women and 46% were men).

The sign-up form told potential participants that the study would comprise the five writing assignments shown in Supplementary Table 2. Potential participants were informed that they would earn a US$20 bonus if, and only if, they completed all five assignments by a deadline 10 d later. They were told that participants who failed to complete all assignments by the deadline would receive only US$3 for each hour's worth of completed tasks. The sign-up form let potential participants choose either to sign up for the experiment and receive an immediate compensation of US$0.05 or to receive US$0.15 and forego the opportunity to participate in the study. People who chose to participate were directed to create an account on the to-do list website where the study would be conducted.

The study website presented participants with a to-do list comprising five writing assignments (see Fig. 5; Supplementary Methods). Participants typed each essay into a text box below the corresponding assignment. When a participant submitted an assignment, its length would be checked against the required number of words. If the essay was long enough, the corresponding task was crossed off from the person's to-do list. Participants could freely choose whether and when to start their first assignment, how to spread out their assignments over the 10 day study period and in which order to complete their assignments. They were also free to quit the study at any time, and the website allowed them to log out and log back in at their convenience.

The experiment comprised four conditions that differed in whether and how the tasks on the participant's to-do list were incentivized. On creating their account on the study website, the nth participant was assigned to condition n modulo 4. In the two experimental conditions, the tasks were incentivized by applying the optimal gamification method described above to an MDP model of the sequential decision problem entailed by the to-do list (see below). In the first experimental condition, the resulting optimal pseudo-rewards for completing each assignment were conveyed as points (see Fig. 5), and the participant's total number of points determined their level in the game. In the second experimental condition, the optimal pseudo-rewards were displayed as dollars rather than points. The first control group was shown a plain to-do list without any incentives or game elements, and in the second control condition, the number of points was constant across all tasks. The incentives were displayed next to each entry of the participant's to-do list, and the level and current number of points were displayed above the current task (see Fig. 5).

Of the 111 people who created an account on the to-do list website, 28 were assigned to the control condition with constant points, 29 were assigned to the control condition without points, 26 were assigned to the condition where pseudo-rewards were framed as dollars and 28 were assigned to the condition were pseudo-rewards were framed as points.

To compute optimal pseudo-rewards, we modelled the sequential decision problem entailed by the to-do list as a finite horizon MDP whose state encodes which tasks have already been completed. This model includes one action for each task and an additional action for taking a break. The essence of this model is that completing a task incurs an immediate cost because it requires time and effort but transports the worker into a state that is closer to the desired final state in which they can claim the US$20 bonus for having completed all of their assignments on time. The subjective costs of completing the assignments were measured in the prestudy described in the Supplementary Information. The resulting optimal pseudo-rewards conveyed how much completing each task would contribute towards achieving the goal to earn the US$20 bonus. Because of this principle, tasks that workers perceived to be harder were incentivized with a larger number of points. A detailed description of this model is included in the Supplementary Methods.

When participants completed all tasks, they were shown a bonus code. After the deadline, we posted a reimbursement Human Intelligence Task on Amazon Mechanical Turk that included an exit survey. The exit survey asked participants to rate their motivation to complete the tasks and how rewarding it felt to complete the second task on nine-point Likert scales. This experiment was run only once because it was relatively expensive and time-consuming. No participant's data was excluded from the analysis. Data collection and analysis were not performed blind to the conditions of the experiments.

Experiment 4. We recruited 200 participants on Amazon Mechanical Turk following the procedure of Experiment 3. The sample size was selected to achieve a statistical power of at least 0.8 in a two-tailed Mann–Whitney *U* test on the number of tasks completed in Project 2 with $\alpha=0.05$ and an anticipated effect size of about 0.5 (as observed in a pilot experiment), while accounting for the fact that not all participants would opt into the second part of the experiment. According to the exit survey (see below), the ages of our participants ranged from 24 to 77 yr (average: 38.3 yr, s.d.: 11.6 yr), and about 61% of our participants were women.

Experiment 4 was similar to Experiment 3, but the to-do list now comprised ten daunting writing assignments rather than just five (see Supplementary Table 4). The ten assignments were grouped into two unequal projects. Completing Project 1 was worth US\$20, whereas completing Project 2 was worth only US\$1. Project 1 comprised five of the ten writing assignments, and Project 2 comprised the remaining five. Both projects were comparable in the total amount of work they required, according to the estimates Mechanical Turk workers gave in our prestudy (85 min for Project 1 and 88 min for Project 2; Supplementary Table 4) and their estimates of each task's fair price (US\$11.50 in total for all tasks of Project 1 and US\$10.50 in total for all tasks of Project 2; Supplementary Table 4). This made Project 1 highly lucrative (US\$20 pay for work worth only US\$11.50) and Project 2 highly unprofitable (only US\$1 pay for work worth US\$10.50). Participants were informed about the projects and the payments they would receive for completing them (see Supplementary Fig. 10a). But they were free to complete whichever tasks they wanted or none at all. The assignment of participants to the experiment's three conditions followed the same counterbalancing procedure we used in Experiment 3; that is, the *n*th participant was assigned to condition *n* modulo 3.

When we applied our optimal gamification method to the to-do list of Experiment 4, it automatically assigned positive points to the tasks of Project 1 and negative points to the tasks of Project 2 (see Supplementary Fig. 10d) based on the value of each project, the amount of time required to complete each task and the average wage our participants could earn by working on another task on Amazon Mechanical Turk. The heuristic to incentivize each task according to its duration, by contrast, was insensitive to the projects' values and incentivized the tasks of Project 2 just as much as the tasks of Project 1 (see Supplementary Fig. 10c).

Experiment 4 was conducted over a period of 7 d starting on Monday 4 February 2019 and ending on Monday 11 February 2019. Participants could fill out the exit survey at any point before the deadline regardless of how many tasks or projects they had completed. The exit survey was available until the day after the deadline. On the following day, each participant was paid a bonus based on which project(s) they had completed plus an extra US\$0.50. Participants who had completed some assignments of a project but not all of them received an extra US\$0.50 for each such assignment. In addition to the questions asked at the end of Experiment 3, the exit survey of Experiment 4 included Likert scales for measuring participants' motivation, how rewarding it felt to complete tasks, to what extent they felt they could choose their tasks freely (P. Michaelsen et al., manuscript in preparation)[46], and to what extent they felt that the to-do list website manipulated them (P. Michaelsen et al., manuscript in preparation)[46,47].

No participant's data were excluded from the analysis. Data collection and analysis were not performed blind to the conditions of the experiments. We tested the normality assumptions and heteroscedasticity assumptions of all ANOVAs and *t*-tests using the Kolmogorov–Smirnov test and Bartlett's test, respectively. We found no statistically significant violations of these assumptions for the data on people's motivation ($K_1=0.12$, $P_1=0.990$, $K_2=0.15$, $P_2=0.761$, $K_3=0.17$, $P_3=0.610$ and $\chi^2(2)=2.12$, $P=0.347$), perceived autonomy (Kolmogorov–Smirnov test, $K_1=0.13$, $P_1=0.975$, $K_2=0.10$, $P_2=0.987$, $K_3=0.14$, $P_3=0.775$ and $\chi^2(2)=0.24$, $P=0.887$ for Bartlett's test) or perceived intrusion ($K_1=0.19$, $P_1=0.731$, $K_2=0.20$, $P_2=0.432$, $K_3=0.25$, $P_3=0.143$ and $\chi^2(2)=2.12$, $P=0.347$).

More details on Experiment 4 are included in the Supplementary Methods.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data that support the findings of this study are available at https://osf.io/h7vqy/.

Code availability

The code used to conduct Experiment 1, Experiment 2 and its follow-up experiments is available at https://osf.io/h7vqy/. The code used to conduct Experiment 3 and Experiment 4 is available on GitHub at https://github.com/BrowenChen/Cognitive-Tools-for-Self-Mastery. The code used to analyse the data is available at https://osf.io/h7vqy/ and in the Supplementary Software file.

Received: 14 December 2017; Accepted: 27 June 2019;
Published online: 19 August 2019

References

1. Ariely, D. *Predictably iIrrational*. (Harper Collins, 2008).
2. Tversky, A. & Kahneman, D. Judgment under uncertainty: heuristics and biases. *Science* **185**, 1124–1131 (1974).
3. Gilovich, T., Griffin, D. & Kahneman, D. *Heuristics and Biases: The Psychology of Intuitive Judgment* (Cambridge Univ. Press, 2002).
4. Todd, P. M. & Gigerenzer, G. Environments that make us smart: ecological rationality. *Curr. Dir. Psychol. Sci.* **16**, 167–171 (2007).
5. Oaksford, M. & Chater, N. A rational analysis of the selection task as optimal data selection. *Psychol. Rev.* **101**, 608–631 (1994).
6. Chater, N. & Oaksford, M. Ten years of the rational analysis of cognition. *Trends Cogn. Sci.* **3**, 57–65 (1999).
7. Todd, P. M. & Gigerenzer, G. *Ecological Rationality: Intelligence in the World*. (Oxford Univ. Press, New York, 2012).
8. Griffiths, T. L., Lieder, F. & Goodman, N. D. Rational use of cognitive resources: levels of analysis between the computational and the algorithmic. *Top. Cogn. Sci.* **7**, 217–229 (2015).
9. Gigerenzer, G. & Hoffrage, U. How to improve Bayesian reasoning without instruction: frequency formats. *Psychol. Rev.* **102**, 684–704 (1995).
10. Hoffrage, U., Lindsey, S., Hertwig, R. & Gigerenzer, G. Communicating statistical information. *Science* **290**, 2261–2262 (2000).
11. Gigerenzer, G. & Edwards, A. Simple tools for understanding risks: from innumeracy to insight. *BMJ* **327**, 741 (2003).
12. Thaler, R. H. & Sunstein, C. R. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. (Yale Univ. Press, New Haven, 2008).
13. Johnson, E. J. et al. Beyond nudges: tools of a choice architecture. *Mark. Lett.* **23**, 487–504 (2012).
14. Aronson, J. E., Liang, T.-P. & Turban, E. *Decision Support Systems and Intelligent Systems* (Pearson Prentice-Hall, 2005).
15. Power, D. J., Sharda, R. & Burstein, F. in *Wiley Encyclopedia of Management* Vol. 7 (eds Straub, D. & Welke, R.) 1–11 (Wiley, 2015).
16. Aviv, Y. & Pazgal, A. A partially observed Markov decision procBhatnagar, S., Fernández-Gaucherand, E., Fu, M. C., He, Y. & Marcus, S. I. A Markov decision process model for capacity expansion and allocation. InIEEE Conference on Decision and Control Vol. 2 (ed Altman, E.) 1380–1385 (IEEE, 1999).ess for dynamic pricing. *Manag. Sci.* **51**, 1400–1416 (2005).
17. Song, H., Liu, C.-C., Lawarrée, J. & Dahlgren, R. W. Optimal electricity supply bidding by Markov decision process. *IEEE Trans. Power Syst.* **15**, 618–624 (2000).
18. Gadomski, A. M., Bologna, S., Costanzo, G. D., Perini, A. & Schaerf, M. Towards intelligent decision support systems for emergency managers: the IDA approach. *Int. J. Risk Assess. Manag.* **2**, 224–242 (2001).
19. Bhatnagar, S., Fernández-Gaucherand, E., Fu, M. C., He, Y. & Marcus, S. I. A Markov decision process model for capacity expansion and allocation. In *IEEE Conference on Decision and Control* Vol. 2 (ed Altman, E.) 1380–1385 (IEEE, 1999).
20. Nunes, L. G. N., de Carvalho, S. V. & de Cássia Meneses Rodrigues, R. Markov decision process applied to the control of hospital elective admissions. *Artif. Intell. Med.* **47**, 159–171 (2009).
21. Ainslie, G. & Haslam, N. in *Choice over Time* (eds Loewenstein, G. & Elster, J.) 57–92 (Russell Sage Foundation, 1992).
22. Myerson, J. & Green, L. Discounting of delayed rewards: models of individual choice. *J. Exp. Anal. Behav.* **64**, 263–276 (1995).
23. Huys, Q. J. M. et al. Bonsai trees in your head: how the Pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Comput. Biol.* **8**, e1002410 (2012).
24. Steel, P. The nature of procrastination: a meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychol. Bull.* **133**, 65–94 (2007).
25. Mischel, W., Shoda, Y. & Rodriguez, M. L. Delay of gratification in children. *Science* **244**, 933–938 (1989).
26. Deterding, S., Dixon, D., Khaled, R. & Nacke, L. From game design elements to gamefulness: defining gamification. In *Proc. of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (eds Lugmayr, A., Franssila, H., Safran, C. & Hammouda, I.) 9–15 (ACM, 2011).
27. Hamari, J., Koivisto, J. & Sarsa, H. Does gamification work? A literature review of empirical studies on gamification. In *Proc. of the 47th Hawaii International Conference on System Sciences* (ed Sprague, R. H., Jr.) 3025–3034 (IEEE, 2014).
28. Callan, R. C., Bauer, K. N. & Landers, R. N. in *Gamification in Education and Business* (eds Reiners, T. & Wood, L. C.) 553–568 (Springer, 2015).
29. Devers, C. J. & Gurung, R. A. in *Gamification in Education and Business* (eds Reiners, T. & Wood, L. C.) 417–430 (Springer, 2015).

30. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction*. (MIT Press, 1998).
31. Callaway, F. et al. A resource-rational analysis of human planning. In *Proc. of the 40th Annual Conference of the Cognitive Science Society* (eds Kalish, C. et al.) 178–183 (Cognitive Science Society, 2018).
32. Baumeister, R. F., Heatherton, T. F. & Tice, D. M. *Losing Control: How and Why People Fail at Self-Regulation* (Academic, 1994).
33. Ng, A. Y., Harada, D. & Russell, S. Policy invariance under reward transformations: theory and application to reward shaping. In *Proc. of the 16th Annual International Conference on Machine Learning* (eds Bratko, I. & Dzeroski, S.) 278–287 (Morgan Kaufmann, 1999).
34. Rieley, J. B. Are your employees gaming the system? *Natl Product. Rev.* **19**, 1–6 (2000).
35. Bevan, G. & Hood, C. What's measured is what matters: targets and gaming in the english public health care system. *Public Adm.* **84**, 517–538 (2006).
36. Thom, J., Millen, D. & DiMicco, J. Removing gamification from an enterprise SNS. In *Proc. of the ACM 2012 Conference on Computer Supported Cooperative Work* (eds Poltrock, S., Simone, C., Grudin, J., Mark, G. & Riedl, J.) 1067–1070 (ACM, 2012).
37. Lieder, F., Callaway, F., Krueger, P., Griffiths, T. & Gul, S. Discovering and teaching optimal planning strategies. In *KOGWIS2018: Computational Approaches to Cognitive Science* (eds Rothkopf, C. et al.) 38 (GK, 2018).
38. Simon, H. A. Rational choice and the structure of the environment. *Psychol. Rev.* **63**, 129 (1956).
39. Gagné, M. & Deci, E. L. Self-determination theory and work motivation. *J. Organ. Behav.* **26**, 331–362 (2005).
40. Steel, P. & König, C. J. Integrating theories of motivation. *Acad. Manag. Rev.* **31**, 889–913 (2006).
41. Puterman, M. L. *Markov Decision Processes: Discrete Stochastic Dynamic Programming* (John Wiley & Sons, 2014).
42. Krueger, P. M., Lieder, F. & Griffiths, T. L. Enhancing metacognitive reinforcement learning using reward structures and feedback. In *Proc. of the 39th Annual Conference of the Cognitive Science Society* (eds Gunzelmann, G. et al.) 2469–2474 (Cognitive Science Society, 2017).
43. Hertwig, R. & Grüne-Yanoff, T. Nudging and boosting: steering or empowering good decisions. *Perspect. Psychol. Sci.* **12**, 973–986 (2017).
44. Hertwig, R. When to consider boosting: some rules for policy-makers. *Behav. Public Policy* **1**, 143–161 (2017).
45. Keramati, M., Smittenaar, P., Dolan, R. J. & Dayan, P. Adaptive integration of habits into depth-limited planning defines a habitual-goal–directed spectrum. *Proc. Natl Acad. Sci. USA* **113**, 12868–12873 (2016).
46. Michaelsen, P., Johansson, L.-O. & Hedesström, M. Experiencing nudges: autonomy, intrusion and choice satisfaction as judged by people themselves. In *Society for Judgment and Decision Making: The 2017 38th Annual Conference* (eds Shu, S. et al.) 30 (SJDM, 2017).
47. Dillard, J. P. & Shen, L. On the nature of reactance and its role in persuasive health communication. *Commun. Monogr.* **72**, 144–168 (2005).

Acknowledgements

Preliminary versions of Experiments 1 and 2 were presented at the 38th Annual Meeting of the Cognitive Science Society. This material has been substantially revised and expanded for the present Article. We would like to thank J. R. Nill and E. Q. Zhang for their contributions to developing the to-do list gamification app, P. Michaelsen for sharing his self-report measures of autonomy and intrusion, T. Ntounis for technical support, and E. Kon, R. Antonova, P. Krueger, M. Pacer, D. Reichman, S. Russell and J. Suchow for feedback and discussion. This work was supported by grant number 1757269 from the National Science Foundation, grant number ONR MURI N00014-13-1-0341, and a grant from the Templeton World Charity Foundation to T.L.G. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

F.L. and T.L.G. designed the research. F.L., P.M.K. and O.C. performed the research. F.L. and P.M.K. analysed the data. F.L. and T.L.G. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41562-019-0672-9.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to F.L.

Peer review information: Primary Handling Editor: Stavroula Kousta.

nature research

Corresponding author(s): Falk Lieder

Last updated by author(s): Apr 24, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The code used to conduct Experiment 1, Experiment 2, and its follow-up experiments is available via the Open Science Framework: https://osf.io/h7vqy/ The code used to conduct Experiment 3 and Experiment 4 is available on GitHub at https://github.com/BrowenChen/Cognitive-Tools-for-Self-Mastery.
Data analysis	The code used to analyze the data is available through the Open Science Framework: https://osf.io/h7vqy/

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data presented in this article is available at https://osf.io/h7vqy/

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

Experiments 1 and 2 and their follow-up experiments were quantitative online experiments that measured people's decisions and reaction times in a sequential decision-making task. Experiments 3 and 4 were naturalistic online experiments that measured people's decision about which task to work on.

Research sample

Adults working on Amazon Mechanical Turk who are located in the USA who have a HIT approval rate of at least 95%.

Sampling strategy

We posted HITs on Amazon Mechanical Turk and let any member of our sample participate in the posted experiment until we had reached the intended number of participants.

The sample size was chosen to achieve about 60 participants per condition. No statistical methods were used to predetermine sample sizes but our sample sizes are larger than those reported in previous publications (Huys, et al., 2012; Keramati, et al., 2016}. Having 60 participants per condition would achieve a power of about 0.8 for a two-sided, two-sample t-test with alpha=0.05 assuming an effect size of 0.50 (exact power: 0.78).

For Experiment 2 we increased the sample size to about 100 participants per condition, because the expected some of the effect sizes to be lower than in Experiment 1..No statistical methods were used to predetermine sample sizes but our sample sizes are larger than those reported in previous publications (Huys, et al., 2012; Keramati, et al., 2016}. The selected sample size achieved a power of about 0.70 for a two-sided, two-sample t-test assuming an effect size of 0.35.

For Experiment 3, we aimed for a sample size of 30 participants per condition so that a chi^2 test would be able to detect an increase from an expected completion rate of about 60% in the control condition to about 90% in the condition with monetary pseudo-rewards at an alpha-level of 0.05 with a power of 0.85.

For Experiment 4, we aimed for a sample size of 42 participants per condition to achieve a power of at least 0.8 in a two-sided Mann-Whitney U-tests on the number of tasks completed in Project 2 with alpha=.05 and an anticipated effect size of about 0.5 (as observed in a pilot experiment). Taking into account the proportion of participants who chose not to participate in Part 2 in our pilot experiment, we increased the sample size per condition to 50.

Data collection

Data collection was performed via websites that people were directed to via Amazon Mechanical Turk. All data was collected via keyboard and/or mouse input from the participants.

Timing

Experiment 1: January 8 2016
Experiment 2: March 17 2016; Follow-Up Experiment 1: April 6 2016; Follow-Up Experiment 2: December 22 2016
Experiment 3: April 24 2017 -- May 4 2017
Experiment 4: February 4 2019 -- February 12 2019

Data exclusions

The exclusion criteria of each experiment are described in the Supplementary Information.
In Experiment 1, We excluded 3 participants who invested less than one third of the median response time of their condition and 11 participants who scored lower than 95% of all participants in their condition (5.5%), leading to a total exclusion rate of 7%.

To start Experiment 2, the experiment participants had to pass a quiz comprising three questions on how their financial bonus would be determined and three questions testing their understanding of the mechanics of the task. Out of the 400 participants, 335 had not participated in any of our previous flight planning experiments and were included for in this study. Out of those 335 participants, we excluded subjects whose median response time was less than one third of the median response time across all included subjects. In addition, we excluded the 5% of participants with the lowest scores of each group. This led to the exclusion of 19 out of the 335 included participants (5.7%).

No participants were excluded from Experiment 3.
No participants were excluded from Experiment 4.

Non-participation

In Experiments 1 and 2 data was only recorded to our data base when a participant finished the experiment. We therefore cannot tell how many people started the experiment but dropped out before finishing it; however these experiments were very short.

In Experiment 3, nine out of 120 participants opted not to participate in the second part of the experiment and to receive $0.10 instead.

In Experiment 4, 45 out of 200 participants opted not to participate in the second part of the experiment and to receive $0.10 instead.

Randomization	The assignments of participants to groups was performed at random. In Experiments 1 and 2 this was was implemented by drawing a random number from a uniform distribution over the conditions independently for each participant. In Experiment 3 the first participant who registered was assigned to condition 1, the second participant was assigned to condition 2, the third participant was assigned to condition 3, the fourth participant was assigned to condition 4, the fifth participant was assigned to condition 1, and so on. Experiment 4 used a counterbalancing procedure that was equivalent to the one used in Experiment 3.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	Participants were adults located in the United States of America who work on Amazon Mechanical Turk.
Recruitment	All participants were recruited online by posting HITs to Amazon Mechanical Turk. Anyone who was located in the USA and had an acceptance rate of at least 95% on Amazon Mechanical Turk was allowed to participate.
Ethics oversight	The experiments reported in this article were approved by the Committee for the Protection of Human Subjects of the University of California, Berkeley under protocol number 2016-02-8359 (``To-Do-List Gamification'') and Experiment 4 was additionally approved by the Independent Ethics Council of the medical faculty of the University of Tübingen as IEC Project Number 668/2018BO2 (``To-Do-List Gamification'').

Note that full information on the approval of the study protocol must also be provided in the manuscript.