

Procrastination in Rational Agents

Peiyuan Zhang (pz580@nyu.edu)

Center for Neural Science, New York University, 4 Washington Place
New York, NY 10003, USA

Wei Ji Ma (weijima@nyu.edu)

Center for Neural Science, New York University, 4 Washington Place
New York, NY 10003, USA

Abstract

Procrastination is prevalent in people's daily lives. Empirical studies of procrastination have identified various contributing factors, including temporal discounting, perfectionism, and time pressure. Models of procrastination, however, have only considered temporal discounting and have ignored other factors. Also, they assume that people make a binary choice between working and delaying, which ignores the fact that people often also decide the amount of effort they put into their work. We bridge this gap by using reinforcement learning theory to re-examine procrastination. Our model predicts two scenarios in which a rational agent chooses to procrastinate: one is that people delay working until the last minute if their cost is sensitive to increases in effort when it is low, and another is that people do not work at all if they pursue perfection under time pressure. Our model raises questions about the common notion that procrastination is necessarily irrational, and suggests several experimental tests.

Keywords: procrastination; reinforcement learning; time pressure; perfectionism; cost function

Introduction

Procrastination is ubiquitous and is especially prevalent in academic tasks (Solomon & Rothblum, 1984). For example, graduate students delay writing their thesis and rush in the end. As often happens when a common-language term becomes the object of scientific study, there are multiple definitions of procrastination. However, they all agree on one essential element, namely delay in working.

Despite many empirical studies on procrastination (Steel, 2007), theories of procrastination fall short of explaining all aspects of procrastination behavior. First, they simplify actions as binary: either working or not working (O'Donoghue & Rabin, 1999). However, in daily life, when we work, we also decide how much effort we put into our work. For example, one might gradually work more each day as the deadline approaches. Second, the central idea of existing theories of procrastination is that future reward is temporally discounted (O'Donoghue & Rabin, 1999; Steel & König, 2006). This idea explains impulsiveness – spontaneity and a tendency to act on a whim – as one factor in procrastination. However, impulsiveness is not the whole story. Surveys have revealed that many other personality traits as well as task properties are

correlated with procrastination (Steel, 2007). In this paper, we highlight the roles of perfectionism (Frost, Marten, Lahart, & Rosenblate, 1990) and time pressure (Ferrari, 2001). People with perfectionism tend to set excessively high standards for their performance and evaluate their behavior overly critically (Flett & Hewitt, 2002). Time pressure is a type of psychological stress that occurs when a person has less time available than is necessary to complete a task (Maule, 1993). The current theories are unable to explain how these factors affect procrastination.

Third, procrastination is often considered to be irrational and a failure of self-regulation (Akerlof, 1991), but the criteria for rationality and successful self-regulation are not well-defined. Moreover, is it possible that procrastination is rational under certain circumstances?

To address these issues, we use reinforcement learning theory to study procrastination from a computational perspective. In our model, the agent not only decides to work or delay on a given day, but they also decide the amount of effort they want to make. The model incorporates perfectionism and time pressure as factors, and predicts a novel relevant factor, namely the form of the cost function. The model further predicts that procrastination is rational in two situations.

Model

We assumed discrete time; we arbitrarily refer to the unit of time as a day. We denote by T the total number of days provided to complete the task (deadline). The smaller T , the higher the time pressure. The agent chooses an action on every day from 1 to T . We define the task state s as the proportion of task completion (between 0 and 1).

Reward Associated with s is a reward. We assume a power-law relationship between reward and task completion:

$$R(s) = s^\beta.$$

Despite this simplicity, this relationship has the potential to capture aspects of perfectionism – a personality trait characterized by overly critical evaluation on the work. When β is large, the agent is a perfectionist: they only feel satisfied when the work is (nearly) perfect, with $\beta \rightarrow \infty$ being all-or-nothing (Fig 1A).

Reward Schedule We consider two reward schedules, which we believe are representative of real-world situations. The first reward schedule is **delayed reward**, where the agent



does not receive any external reward until $T + 1$. For example, homework is only graded until one day after the deadline. We formalize the reward schedule as follows:

$$r_t = \begin{cases} 0 & \text{when } t < T + 1 \\ R(s_{T+1}) & \text{when } t = T + 1 \end{cases} \quad (1)$$

where $t = 1, 2, \dots, T, T + 1$ (Fig 1B).

The second reward schedule represents another common scenario where the agent receives **instantaneous reward**. For example, the agent is internally driven by making progress on the task, so they receive instantaneous reward whenever s increases. A real-life example would be a student feels rewarded immediately when they write a chapter of their dissertation. We formalize the reward schedule as (Fig 1C)

$$r_t = R(s_{t+1}) - R(s_t). \quad (2)$$

Effort Cost Working on a task requires the agent to invest mental effort, and such effort is costly (Kool & Botvinick, 2018). We assumed that the cost function follows a power law,

$$C(a) = c_1 a^\lambda, \quad (3)$$

where a denotes the amount of effort. $c_1 > 0$. When $\lambda > 1$, the cost function is convex: increasing cost is associated with successive increments in effort (Fig 1D) (Navon & Gopher, 1979; Glimcher & Fehr, 2013). When $\lambda < 1$, the cost function is concave: the cost is relatively sensitive to increases in effort when it is low, but not when it is high (Fig 1D) (Kool & Botvinick, 2018).

Optimal Policy In a state s , at time t , the goal of the agent is to maximize the sum of the task reward gained by making progress, and discounted value of the next state. Meanwhile, the agent needs to minimize the effort cost. To reach this goal, the optimal policy is derived from the Bellman equation (Bellman, 1957).

If the agent chooses not to work ($a = 0$), then they stay in the same state, and the value of the state-action pair is

$$Q_t(s, a = 0) = r_t + \gamma V_{t+1}(s), \quad (4)$$

where the value of s is specified as

$$V_t(s) = \max_a Q_t(s, a). \quad (5)$$

γ is the discount rate, which determines how much the state value is discounted on the next day ($\gamma \in [0, 1)$). The larger γ , the more the agent takes future value into account. In instantaneous reward condition, as the state does not change, $r_t = 0$ for all the days (r_t is given by equation (2)) (Fig 1C).

By contrast, if the agent decides to work ($a > 0$), then the agent moves the task forward to a new state s' , and pays effort cost immediately. The value of the state-action pair is

$$Q_t(s, a > 0) = r_t - C(a) + \gamma V_{t+1}(s'). \quad (6)$$

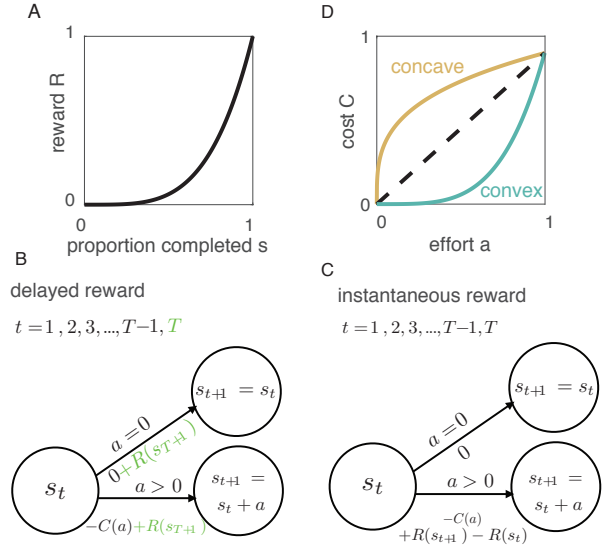


Figure 1: Functions and model. (A) reward as a function of task state. (B) Form of cost function (either concave or convex). (C) Model in delayed reward condition. The reward received at deadline T is marked in green. (D) Model in instantaneous reward condition.

Here we assumed a simple deterministic linear relationship between s and s' , i.e., $s' = s + a$: every additional unit of effort adds a fixed amount of progress.

$Q_t(s, a)$ is derived from solving the Bellman equation using dynamic programming. Policy consists of the amount of effort made on each day. The optimal policy maximizes $Q_t(s, a)$ on each day:

$$\pi_t(s) = \operatorname{argmax}_a Q_t(s, a). \quad (7)$$

Net Utility We evaluate the optimal policy with net utility, which is the reward the agent has gained at the end of the task minus the total effort cost.

$$U_{\text{net}} = R(s_{T+1}) - \sum_{t=1}^T C(\pi_t(s)). \quad (8)$$

Results

We define procrastination as delay in working before task completion, i.e., $a = 0$ when $s < 1$. We explored the parameter space in our model to look for procrastination behavior in the optimal policies. We further explored the effect of task properties (two reward schedules and total time T), and of parameters λ, β, γ , which might represent cognitive and personality traits of an agent (table 1). We found two patterns of procrastination: 1) binge working if the agent's mental cost is sensitive to increases in effort when it is low (i.e., concave cost function), and 2) not working at all if the agent strives for perfection under high time pressure.

Table 1: List of parameters and their psychological interpretations.

parameters	psychological interpretations
λ	convex versus concave cost function
β	level of perfectionism
γ	reward discount rate
r_t	delayed or instantaneous reward
T	total time (time pressure)

Binge Working If an agent has a concave cost function ($\lambda < 1$), they only work for one day (binge working), regardless of their discount rate, their perfectionism level, and the total time. On which day they choose to work depends on the reward schedule. The agent only works on the first day if they receive reward instantaneously, whereas if the reward is delayed, they only work on the last day (last-minute worker) (Fig 2A first row). They do not work until the last day and finish all work on that day.

Binge working is the optimal policy for an agent with a concave cost function. Intuitively, for an agent whose effort cost is very sensitive under low effort, it is costly for them to divide their effort across days instead of expending their total effort in one day.

We illustrate this optimality by comparing the net utility between the optimal policy and three alternative suboptimal policies (Fig 2A second to fourth row). The optimal policy indeed has the greatest net utility. Separating the total effort into pieces, for example two halves, is more costly and results in having lower net utility (Fig 2A second row).

Not Working at All For an agent with a convex cost function ($\lambda > 1$), we first explore whether or not they procrastinate even if they do not discount future reward ($\gamma \rightarrow 1$). When $\gamma \rightarrow 1$, the optimal policy is the same in both delayed and instantaneous reward conditions.

If the total time is limited, the optimal policy for an agent with high perfectionism ($\beta > \lambda > 1$) is to not work at all (Fig 2B bottom row in right panel, no work when $T < 4$). Intuitively, if the agent receives little reward before task completion, and meanwhile has less time available than necessary to complete the task, they should not even start. We again illustrate this optimality by comparing the net utility between the optimal policy (Fig 2C first row) and three alternative suboptimal policies (Fig 2C second to fourth row). The net utility of not working at all is 0, but is negative for other policies, including working a little bit for one day, working a lot for one day, and working steadily.

Next we considered the case where an agent (still with $\lambda > 1$) discounts future reward ($\gamma < 1$). We found that an agent who discounts future reward to a lesser degree (larger γ) are more tolerant of time pressure, whereas an agent who discounts future reward to a greater degree (smaller γ) are sensitive to time pressure in both reward schedules (Fig 2B right panel). For example, in the delayed reward condition,

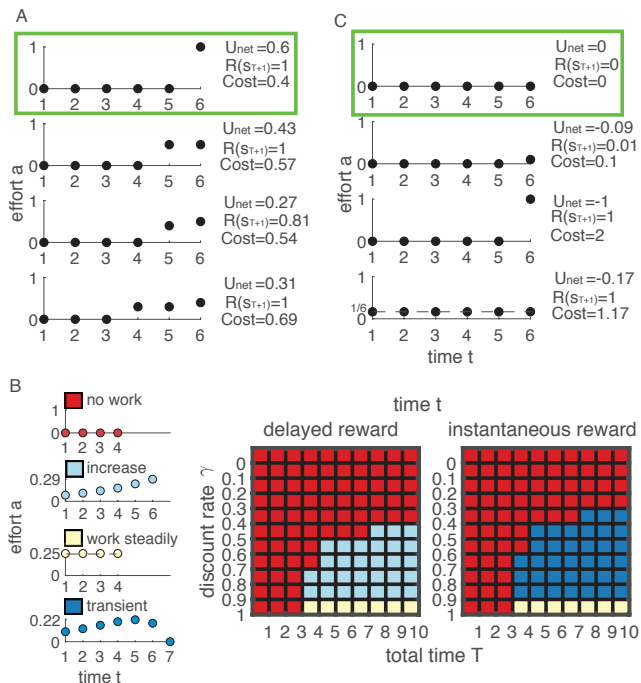


Figure 2: Model predictions. (A) First row: an example of the optimal policy (marked with green box) for a last-minute agent. Second to the fourth row: alternative suboptimal policies. (B) Top row: policy of not working at all; Second to the fourth row: alternative suboptimal policies. (C) Right panel: the effect of time pressure, temporal discounting, and reward schedule on an agent with a convex cost function as well as perfectionism. Left panel: Illustration of the policies corresponding to the different conditions in the right panel (the correspondence is marked by colors).

when $T = 5$, an agent with $\gamma > 0.5$ will choose to work, but those with $\gamma < 0.5$ will choose not to work at all. An extreme case is that when the discount rate is below some threshold, the agent chooses not to work, even when given infinite time. This discount rate threshold is lower in the instantaneous reward condition than the delayed reward condition (Fig 2B right panel, the threshold is $\gamma = 0.4$ in the delayed reward condition and $\gamma = 0.3$ in the instantaneous reward condition). Moreover, given the same γ , an agent who receives instantaneous reward are more tolerant of time pressure than an agent who receives delayed reward.

Discussion

We applied reinforcement learning theory to examine procrastination quantitatively. Two patterns of procrastination are found in a rational agent: binge working and not working at all. Moreover, our model also predicts multiple qualitatively different policies in non-procrastinators: 1) increase in the amount of effort as the deadline approaches (Fig 2B left panel second

row), 2) working steadily (Fig 2B left panel third row), and 3) transient working (a great amount of effort made in the middle and lower effort early and late (Fig 2B left panel fourth row).

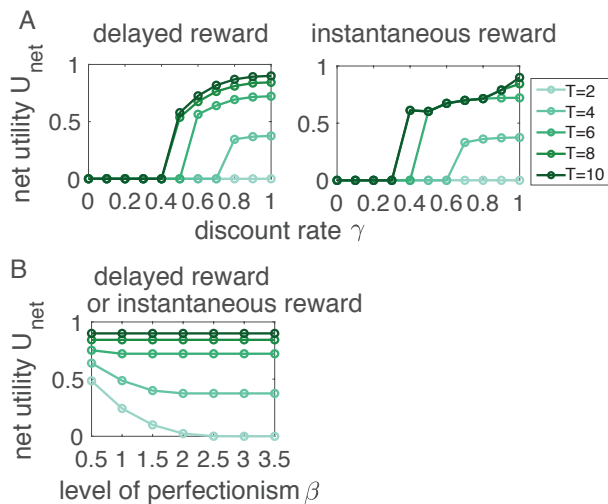


Figure 3: (A) The effect of discount rate on net utility received by an agent with a convex cost function as well as perfectionism, in delayed reward condition (left panel) and instantaneous reward condition (right panel) respectively. (B) The effect of level of perfectionism on net utility for an agent with a convex cost function and $\gamma \rightarrow 1$. The effect stays the same between the delayed reward and the instantaneous reward conditions.

The model predicts that a rational agent with certain cognitive and personality traits chooses to procrastinate under certain task conditions. Can they obtain higher net utility if they change their traits? To address this question, we tested three traits in our model separately. First, we studied how discount rate influences net utility, particularly on an agent with a convex cost function as well as perfectionism. We found that people who discount future reward less (larger γ) obtain higher net utility in both delayed and instantaneous reward conditions, given all different total time (Fig 3A). In other words, a rational agent obtains higher net utility if they extend their temporal horizon. Second, we tested how the level of perfectionism (β) affects net utility. Under low time pressure, the net utility stays the same across different levels of perfectionism, whereas, under high time pressure, people with a higher level of perfectionism obtain lower net utility (Fig 3B). This result suggests that, under time pressure, a rational agent should lower their level of perfectionism to obtain higher net utility. Finally, given enough time, an agent with a convex cost function often obtains higher net utility than an agent with a concave cost function. The reason is that, the net utility received by an agent with a concave cost function is independent of the total time, whereas an agent with a convex cost function obtains higher net utility when the total time increases (Fig 3).

Our model predicts multiple qualitatively different policies, which are in principle experimentally testable. We can change the task condition (reward schedule or total time) to see if an agent changes their policy according to the model predictions. Also, we can compare the policies across people with concave versus convex cost functions, different levels of perfectionism, or discount rates.

This preliminary model can be improved in several aspects. First, we only considered two reward schedules where reward comes either at the very end or along the way. In daily life, we often receive rewards in a mixed schedule. Also, we assumed a deterministic state transition for simplicity. This state transition could be instead stochastic. Finally, the mapping from effort to progress sometimes depends on task state, for instance, writing a paper: even though we make the same amount of effort, we might make very little progress in the beginning, but much more progress later. Nevertheless, this simple model is a basis for quantitatively examining procrastination.

References

- Akerlof, G. A. (1991). Procrastination and obedience. *The American Economic Review*, 81(2), 1–19.
- Bellman, R. (1957). A markovian decision process. *Journal of Mathematics and Mechanics*, 679–684.
- Ferrari, J. R. (2001). Procrastination as self-regulation failure of performance: effects of cognitive load, self-awareness, and time limits on working best under pressure. *European journal of Personality*, 15(5), 391–406.
- Flett, G. L., & Hewitt, P. L. (2002). Perfectionism and maladjustment: An overview of theoretical, definitional, and treatment issues.
- Frost, R. O., Marten, P., Lahart, C., & Rosenblate, R. (1990). The dimensions of perfectionism. *Cognitive therapy and research*, 14(5), 449–468.
- Glimcher, P. W., & Fehr, E. (2013). *Neuroeconomics: Decision making and the brain*. Academic Press.
- Kool, W., & Botvinick, M. (2018). Mental labour. *Nature human behaviour*, 1.
- Maule, O. S. A. J. (1993). *Time pressure and stress in human judgment and decision making*. Springer Science & Business Media.
- Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological review*, 86(3), 214.
- O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review*, 89(1), 103–124.
- Solomon, L. J., & Rothblum, E. D. (1984). Academic procrastination: Frequency and cognitive-behavioral correlates. *Journal of counseling psychology*, 31(4), 503.
- Steel, P. (2007). The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological bulletin*, 133(1), 65.
- Steel, P., & König, C. J. (2006). Integrating theories of motivation. *Academy of management review*, 31(4), 889–913.