Optimal and sub-optimal temporal decisions can explain procrastination in a real-world task

Anonymous CogSci submission

Abstract

Keywords: procrastination; computational modeling; temporal decisions; temporal discounting; naturalistic data

Introduction

Taxonomy

Definition

Mechanisms

Why commit but late?

Why defect?

Why not commit?

Methods

Task

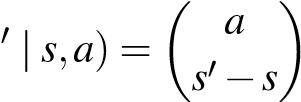
We analysed and modelled real-world data from Zhang and Ma (2023). To receive course credit, 194 bachelor students in a Psychology course had to be subjects in at least 7 hours of experiments over the course of a semester that lasted for 110 days or about 15 weeks. Each research study required a minimum time of 0.5 hours or a multiple thereof, hence defining a unit of work. Additional participation above 7 hours contributed ⅛th of a grade point per unit, up to 4 extra hours of work. Abundant research opportunities of on average 15 hours per student implied they did not compete (Zhang & Ma, 2023).

We f

Models

In order to model the temporal allocation of work in this task, we used Markov Decision Processes (MDPs). MDPs provide a normative framework to model sequential decisions where summed discounted long-term net utility must be maximized given task dynamics (Sutton & Barto, 2018; Dayan & Daw, 2008). In our case, the solution to the MDP is an optimal policy specifying the utility-maximising allocation of work through the semester. It is also straightforward to relax assumptions such as exact knowledge of the decision problem or time-consistent discounting and derive sub-optimal policies that might mimic real behavior in the task.

We formalised the task as follows. The agent has 16 weeks (0 ≤ *t* ≤ 15) to complete at least 14 units and up to 22 units of work. At each week, the agent occupies a state (0≤*st* ≤22) corresponding to the number of units so far completed. Every week, the agent can decide to complete some number of units (0 ≤ *at* ≤ 22−*st*). However, there exists a Binomial success probability or efficacy (η) governing the number of units actually completed. This is a probabilistic interpretation of effectiveness or rate of unit completion, where higher η means greater number of units can be completed per week. The Binomial transition probabilities are given by:

*P*(*s*η*s*′−*s* (1−η)*a*−*s*′+*s* (1)

Every unit of work incurs an immediate effort cost *r*work , while the remaining time not used for work (i.e., 22−*at* units) is used to ‘shirk’ giving an immediate reward *r*shirk. Shirking refers to any alternative task including other university work, relaxing, chores, etc. Finally, there is a reward associated with completing each unit *r*unit , which is only delivered if 14 units are completed. Every unit above 14 units earns an extra reward of *r*extra.

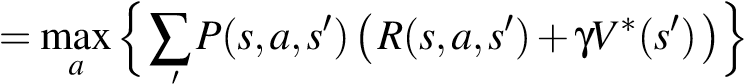
Technically, the rewards are ‘delivered’ as soon as 14 units (7 hours) are hit, since the students know immediately when they have finished the requirement. However, in the absence of explicit rewards, there is still some ambiguity about how they perceive the reward schedule. It is also possible that the reward of finishing is only perceived with formal confirmation of the credits and grades at the end of the semester. Indeed, some students indicated in verbal responses that they felt the deadline was too far to act initially. We show below how these different reward schedules contribute to different mechanisms leading to procrastination.

In all cases, we find the optimal value function by recursively maximising the (discounted) sum of current and future rewards or the returns as given by the Bellman optimality equation (Sutton & Barto, 2018):

∗ n o

*V* (*s*)= max *Q*(*a,s*) (2)

*a*

 (3)

*s*

For action selection, we use a softmax rule, giving rise to noisy decisions and ultimately, noisy trajectories of completion:

exp(β *Q*(*a,s*))

π(*a*|*s*)= (4)

∑*a* exp(β *Q*(*a,s*))

where β is the inverse temperature that controls the extent of the noise.

# Exponential discounting with delayed rewards

In the first model, there is a common exponential discount factor γ for positive and negative rewards, and we assume the rewards from completing a minimum of 14 units come at the end of the semester. We set *r*unit = 4*.*0 and *r*extra = *r*unit4 . The rewards from shirking and efforts from working were *r*shirk = 0*.*1 and *r*work = −0*.*3 and come immediately at the time of action. The softmax paramater β= 7*.*0

Temporal discounting induces a temporal preference for working later

Say there is no discounting, that is, γ= 1. This means that a reward is as valuable later in time as it is immediately. In this case, it is optimal to work and complete the requirements at the start of the semester and then shirk on other tasks later, as seen in Figure 1A.

On the other hand, if delayed rewards are discounted, that is, γ *<* 1, there is a temporal preference for shirking and obtaining rewards immediately over working and paying effort costs to secure a distant reward of the credits. As seen in Figure 1B, work is put off until the end of the semester when γ= 0*.*9.

Efficacy affects the extent of delay

When γ = 1, the policy irrespective of efficacy, is to finish in the beginning. Figure 1A shows that lower the efficacy, longer it takes to finish the task in practice.

When γ = 0*.*9, efficacy controls how late in the semester a subject can afford to delay working, while allowing a reasonable chance of finishing as shown in Figure 1B. Therefore, higher the efficacy, the longer work can be delayed. Some students did suggest that they put off the requirements because they foresaw that it wouldn’t take them too long. Further, at very low efficacies (for example η= 0*.*3), it is no longer worth working for the low-reward extra units beyond the requirement of 7 hours.

A gap between real and assumed efficacy leads to overestimation of delay

So far, we assumed that the subjects calculate their best course of action based on perfect knowledge of their abilities. However, what if they underestimate their efficacy, that is, overestimate the average time it will take them to finish the task? Say γ = 0*.*9 and the actual efficacy is high, ηreal = 0*.*9 for subjects who have a low self-efficacy (ηassumed *<* 0*.*9) and so they plan to work earlier than if they had correctly estimated their efficacy. The greater the gap between the real and assumed efficacy, the earlier the subjects finish before the deadline. This leads to a sigmoid pattern of completion, where most of the work is finished at an intermediate point before the deadline as shown in Figure 1C.

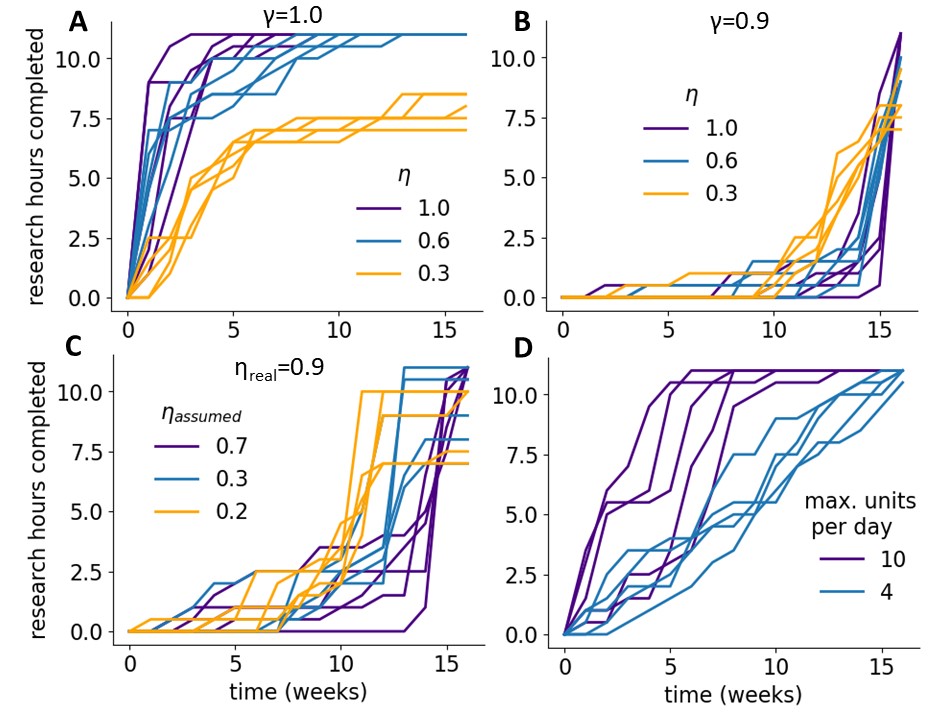


Figure 1: This is a figure.

Limits on the amount of work completed in a week could explain steady completion

Finally, each individual may have a limit on the maximum number of units that they can complete in a week. This could be from limited time available or fatigue from efforts. Even with γ = 1, this leads to a steadier course of working, since a large number of hours cannot be completed in a short time as shown in Figure 1D. The policy in this case is identical to trying to do a fixed number of units every week – something many students indicated they did.

Immediate rewards at threshold

Together, combinations of varying discount factors and efficacies along with delayed rewards were able to explain the different completion patterns found in this real-world task. However, the rewards for completion of the minimum requirement of 7 hours do not actually come at the end of the semester (although they may be perceived as such) but can be realised once the threshold of 14 units is hit. For the next set of simulations, all the parameter settings remain the same, except the timing of rewards. We assume that rewards come immediately after the threshold of 14 units. We set η= 0*.*8.

Immediate rewards at threshold eliminate delays in working due to discounting

As before, in the absence of discounting (γ= 1), present rewards are as valuable as future rewards and hence there is a drive to finish work early. Now, even with discounting (γ *<* 1), there are no delays in working since the rewards are not delayed till the end of the semester. If anything, there is an opposite tendency to work and obtain rewards as soon as possible, expediting work even more than the no discounting case as shown in Figure 2A.

Vigour costs can lead to delays in working

We assumed so far that the effort costs of working scale linearly with the amount of work. However, effort may seem more tedious with greater amount of work akin to the costs of vigour (?, ?). One way to operationalise this is to introduce a convexity that raises the effort costs at a greater rate with more units of work (Zhang, 2024):

*r*effort(*a*)= *r*effort(*a* = 1) *ak* (5)

where *a* is the number of units of work, *k* is the convexity of the effort function, and *r*effort(*a* = 1) is the effort required for a unit of work.

Such costs may make it prohibitive to do a lot of work all at once, inducing a preference for spreading the work out and doing a little bit every week. This means that it will take several weeks until the requirement of 14 units can be completed. This leads to an interesting effect where it becomes optimal to do a small number of units in the beginning due to discounting of these temporarily delayed rewards. The effect can be seen in the progress lines that are curved upwards when discount factors are steep (γ= 0*.*6) and effort functions are relatively more convex (*k* = 2*.*5) as seen in Figure 2B. For shallower discount factors and convexities, this effect disappears.

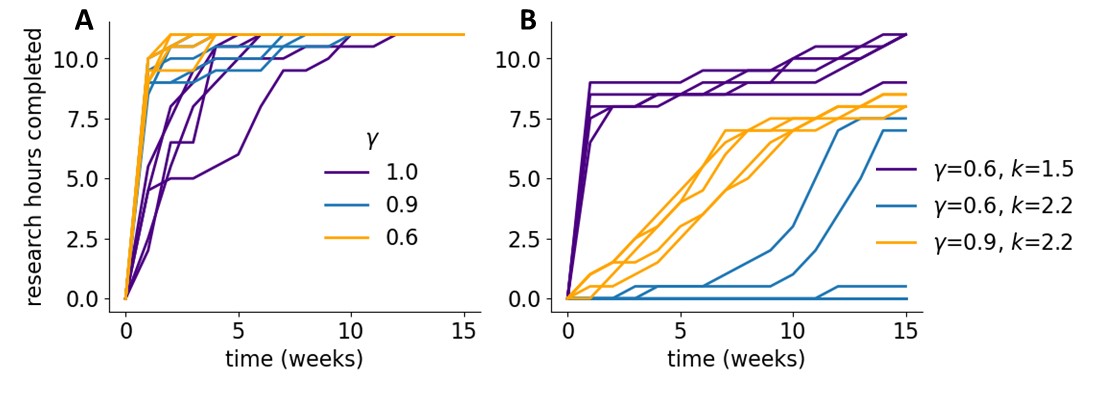


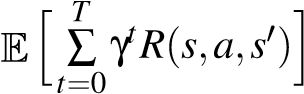
Figure 2: This is a figure.

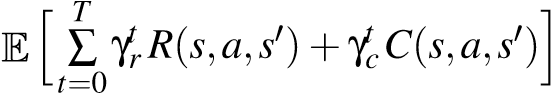
# Differential discounting of efforts and rewards

While discounting along with convex costs can explain some of the delays and patterns of working with immediate rewards, they cannot explain why someone would still delay most work to the end of the semester even if the rewards are immediate. Here, we discuss another possible route to procrastination from time-inconsistent decisions stemming from non-exponential discounting. Others have discussed preference reversals for example due to hyperbolic discounting, where a choice between distant outcomes can reverse closer to the events contrary to initial intentions (Ainslie & Haslam, 1992). Such effects have also been used to explain procrastination (Fischer, 1999; O’Donoghue & Rabin, 2001). Here, we consider different exponential discount factors for rewards and costs (negative rewards) and show how this could lead to defections and delays to the end of the semester in the immediate reward condition (Le Bouc & Pessiglione, 2022). We set the reward for each unit completed *r*unit = 1*.*0 and η= 0*.*8. The other parameter settings are as before and the rewards come immediately on reaching a threshold.

Optimisation problem with different discount factors

Previously, with a single discount factor, the goal was to

maximise. However, now the objective

changes to the following: 

since there are different discount factors (γ*r* and γ*c*) for rewards and costs (*R*(*s,a,s*′) and *C*(*s,a,s*′)). This separation eliminates the structure that allowed us to solve the optimisation problem recursively before. We can still find the best sequence of actions to take from each point in time, but this policy may not be the same in the future timesteps. Hence, we must repeat this process for each horizon.

Repeated delays from temporal inconsistencies lead to procrastination when efforts are discounted more steeply than rewards

Say γ*r* = 0*.*9 and γ*c* = 0*.*5, that is, future efforts are discounted more steeply than rewards. This leads to time-inconsistent policies in the task. As shown in Figure 3, for *s* = 0, at *t* = 0 or horizon = 15, the best policy is to start working at *t* = 2, but come *t* = 2 or horizon= 13, it makes sense to delay again to *t* = 4 and so on. In some sense, there is a constant underestimation of how much effort it takes to do the task in the future. In Figure 3A, we simply show the deterministic policy without any softmax noise.

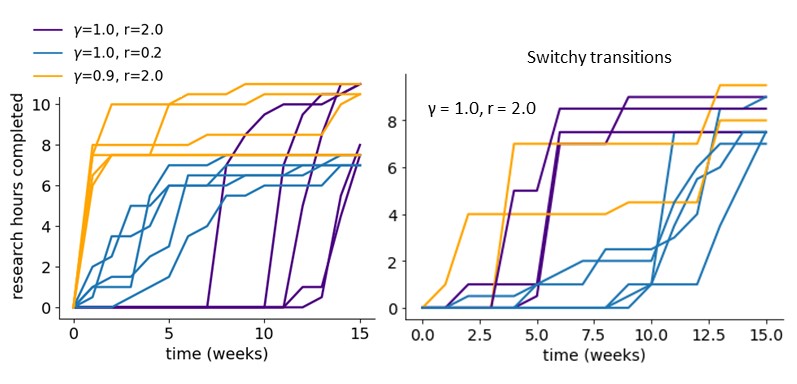


Figure 4: This is a figure.

Possible experiments to disambiguate hypotheses Discussion

References

Ainslie, G., & Haslam, N. (1992). Hyperbolic discounting. In *Choice over time* (p. 57-92). New York, NY, US: Russell Sage Foundation.

Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, Behavioral Neuroscience*, *8*, 429-453.

Fischer, C. (1999). *Read This Paper Even Later: Procrastination with Time-Inconsistent Preferences* (Tech. Rep.). Resources for the Future.

Gevers, J. M., Rutte, C. G., & Van Eerde, W. (2006). Meeting Deadlines in Work Groups: Implicit and Explicit Mechanisms. *Applied Psychology*, *55*(1), 52–72. doi: 10.1111/j.1464-0597.2006.00228.x

Konradt, U., Ellwart, T., & Gevers, J. (2021). Wasting effort or wasting time? A longitudinal study of pacing styles as a predictor of academic performance. *Learning and Individual Differences*, *88*, 102003. doi: https://doi.org/10.1016/j.lindif.2021.102003

Le Bouc, R., & Pessiglione, M. (2022). A neurocomputational account of procrastination behavior. *Nature Communications*, *13*(1), 5639. doi: 10.1038/s41467-02233119-w

O’Donoghue, T., & Rabin, M. (2001). Choice and Procrastination. *The Quarterly Journal of Economics*, *116*(1), 121–160. Retrieved 2023-09-03, from https://www.jstor.org/stable/2696445 (Publisher: Oxford University Press)

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning, second edition: An Introduction*. Cambridge, MA: MIT Press.

Zhang, P. (2024). *The dynamic nature of procrastination*. Doctoral dissertation, Department of Neural Science New York University, New York.

Zhang, P., & Ma, W. J. (2023). temporal discounting predicts procrastination in a real-world task. *PsyArXiv*.