Proposal: Using click activity from the Canvas learning management system to study procrastination and effort allocation in time  
Sahiti Chebolu and Peter Dayan  
  
Background

1. Procrastination literature:  
   Procrastination is widespread affecting some 80% of students and 20% of adults (Steel, 2007). Many suffer effects on their health (Sirois, 2007) and finances  
   (O’Donoghue & Rabin, 1998), and most procrastinators wish to reduce it (O’Brien, 2002). There is no single agreed definition of procrastination, but the myriad procrastination questionnaires include at least some of the following factors: delay of actions and work, unnecessary or unreasonable delays, delaying in spite of intending differently, irrationality in the sense of failing to maximise utility, suffering consequences like missing deadlines or stress due to rushing (Lay, 1986; Mann et al, 1997; McCown et al, 1989; Steel, 2010). Similarly, a search for mechanisms of procrastination reveals a variety of personality trait correlates like lack of conscientiousness facets such as self-control, achievement motivation, discipline etc. (Lee et al, 2006) and neuroticism facets such as impulsiveness, fear of failure low self-esteem etc. (Pychyl & Flett, 2012). In a similar vein, people procrastinate in tasks with a diverse structures and characteristics: when rewards for work are delivered immediately or after a delay tasks with and without deadlines, in the presence or absence of uncertainty about aspects of the task like reward or effort timing or magnitude, and typically tasks that are considered aversive, stressful or boring (Rozental & Carlbring, 2014; Shu & Gneezy, 2010; Zhang et al, 2019).
2. Taxonomy:  
   In Chebolu (2023), we aimed to tackle this heterogeneity by constructing a taxonomy of procrastination types and relatedly, mechanisms based on principles from decision making and reinforcement learning. We proposed three main types of procrastination based on the structure of delays and underlying temporal decisions: deliberate delays leaving insufficient time to complete it, delaying to a later time in spite of intending to act earlier, not committing to a time of action, and engaging late due to reversing a decision to abandon the task. For each of the definitions, we can ask what the reasons for the respective decisions might be, naturally giving rise to a classification of mechanisms. For example, why commit to a time of action that leaves insufficient time, why defect on a decision to act and why not commit to a time of action despite intending to engage?
3. Evidence:  
   Given this theoretical framework, we wanted to test if there is any evidence for the presence of various types and mechanisms empirically. To do this, we used data from a real-world task collected by Zhang and Ma (2023). In this study, 194 bachelor students in a Psychology course at NYU had to participate in at least 7 hours of experiments over a 16-week semester (110 days) to receive course credit, with each research study requiring a minimum time of 0.5 hours or a multiple thereof, hence defining a unit of work. The data consisted of the times at which each student completed each unit of work. In the original study, Zhang and Ma (2023) showed a correlation between students’ discount factors and the extent of their delays.   
    We clustered the data to find patterns of work allocation and identified some mechanisms that explain why people delay. Some explanations for delays are: discounting of future rewards with delay, time-inconsistent choices from non-exponential discounting or waiting for better conditions or more interesting tasks. Qualitatively, it seems like many of these explanations could explain the same pattern in the data. Quantitative fitting proves a bit more challenging due to the availability of only a single trajectory per person.  
    There are several advantages to analyzing and modelling another dataset with following characteristics in addition to this one: firstly, as in this dataset, there should be information about when (and how much) work has been done towards a task along with information about task structure like deadlines, rewards and costs for finishing or not, etc.; secondly, the presence of multiple trials per person allows to determine the best fitting models better; thirdly, the current task is easy for the students which means that each unit of work can be completed with full certainty once started. However, many real-life situations involve tough tasks where success is not guaranteed and here, other factors like fear of failure, anxiety, low self esteem can come into play.

Research questions and hypotheses

1. What data are we interested in?   
   For these reasons, the Canvas learning management system data collected under strand 2 of the UCL-MUST project is especially interesting and relevant to the goals of this project. Information about when students access course-related material, when they work on and submit their assignments coupled with information about course context like the deadlines and time available for each assignment can together inform how students distribute their efforts in time towards a deadline. Hence, data from chemistry and biology courses CHM …, CHM …, BIO … where clicks have been annotated and information about course deadlines is present is of specific interest to us in terms of analysis and modelling.
2. What do we want to do with it? Research questions?  
   Broadly, we have the following questions:   
   a. Is there any evidence for multiple types of and/ or explanations for procrastination from students’ behavior on this task?  
   b. How do students allocate work over time? Are their allocation patterns consistent across trials? What mechanisms and explanations underly these decisions?  
     
   As a first step, we propose to conduct a model-agnostic analysis of how students have allocated efforts towards a task, like an assignment or a quiz. If an assignment is done online on Canvas, then it might be possible to extract when a student has worked on the assignment. Otherwise, the click activity preceding a deadline (but after the previous deadline) can be taken as a proxy for the amount of effort applied at a given time point towards a task. Given such time courses of working per person per task, we can cluster them to characterize the different ways and styles of work allocation in the task. Further, since multiple trajectories are available per student, it will be possible to determine how stable or variable their behavior is in time across tasks in a course or across multiple courses.   
    Following such a descriptive analysis of the time courses, we will turn to understand the mechanistic basis of students’ temporal decisions through computational models. As before, we will use a Markov Decision Process framework. In addition to the explanations of discounting of delayed rewards, non-exponential discount factors, waiting for interesting tasks, we can also explore other explanations:   
   a. Delays from waiting to check if an assignment problem can be improved, or waiting to gain more information from peers, self-learning or classes.  
   b. Fear of failure on a hard task might lead students to delay engaging in order to avoid finding out that would confirm their fears.   
   c. Self-handicapping by delaying to set oneself up for failure to preserve one’s self-esteem by deflecting blame to an external cause (delay).   
    After formalizing these models, we will fit them to the data to determine which of them best explain the patterns in the data and if there are multiple models that are equally explanatory.
3. Assumptions about data:   
   a. Clicks are annotated such that it is possible to filter out clicks not related to the task or not contributing to task completion.  
   b. Click activity either directly informs how much work is done towards the task or this can be indirectly inferred from the amount of activity.

References