Department of Economics
ISSN number 1441-5429

How time flies!

Discussion Paper no. 2021-09

Xiu Chen and Xiaojian Zhao

Abstract:

The paper identifies a potential gap between intertemporal choices and time preference: The elicited intertemporal decisions could be partly driven by a biased perception of time and thus may not completely reveal the actual time preference. To test this, we explore the causal relationship between time perception and intertemporal choices by conducting a laboratory experiment, in which cognitive load is used as a stimulating instrument to induce differences in time perception. We establish that the perceived time lengths for subjects with high cognitive load are shorter than those with low cognitive load and that individuals who underestimate time appear more patient in their intertemporal decisions. The mediation analyses show that time perception mediates a significant part of the cognitive load's effect on intertemporal choices. Our study thus demonstrates that the time preference identified by intertemporal choices might be confounded by the potentially biased perception of how time flies.

Keywords: time perception, intertemporal choice, time preference, cognitive load

JEL Classification: C91, D91

Xiu Chen: College of Business, Southern University of Science and Technology (email: cxiu506@gmail.com); Xiaojian Zhao: Monash University (email: xjzhao81@gmail.com).

How time flies![*]

Xiu Chen[†]   Xiaojian Zhao[‡]

September 2021

Abstract

The paper identifies a potential gap between intertemporal choices and time preference: The elicited intertemporal decisions could be partly driven by a biased perception of time and thus may not completely reveal the actual time preference. To test this, we explore the causal relationship between time perception and intertemporal choices by conducting a laboratory experiment, in which cognitive load is used as a stimulating instrument to induce differences in time perception. We establish that the perceived time lengths for subjects with high cognitive load are shorter than those with low cognitive load and that individuals who underestimate time appear more patient in their intertemporal decisions. The mediation analyses show that time perception mediates a significant part of the cognitive load's effect on intertemporal choices. Our study thus demonstrates that the time preference identified by intertemporal choices might be confounded by the potentially biased perception of how time flies.

Keywords: *time perception, intertemporal choice, time preference, cognitive load*

JEL Classification: C91, D91

[*]We thank Yan Chen, Zhenxing Huang, King King Li, Jane Zhang and conference participants at 2018 ESA Asian-Pacific Meeting (Brisbane), the 7th Xiamen University International Workshop on Experimental Economics and 2019 ESA Asian-Pacific Meeting (Abu Dhabi), as well as participants in many seminars for their valuable comments and suggestions. We are grateful to Lijia Wei for sharing the laboratory of the Center for Behavioral and Experimental Research of Wuhan University and Wei Huang for his supports in conducting the experiments. Meng Ji provided excellent research assistance. We acknowledge the financial supports from Chinese University of Hong Kong (Shenzhen) presidential fund, Monash Economics Department Supports and the National Natural Science Foundation of China (Grant No. 72003084).

[†]College of Business, Southern University of Science and Technology (cxiu506@gmail.com)

[‡]Monash University (xjzhao81@gmail.com).

Alice: "How long is forever?"

White Rabbit: "Sometimes, just one second."

Lewis Carroll, *Alice in Wonderland*

1 Introduction

Time preference and *time perception* are intertwined.[1] While intertemporal choices are expressions of time preference, they might also hinge upon time perception. If that is the case, can the variation in time perception (partly) account for the heterogeneity of individuals' intertemporal choices? More importantly, as the standard economics approach expects intertemporal choices to reveal time preference, do they completely represent people's actual time preference? This paper therefore formally explores the causal relationship between time perception and time discounting. Intuitively, for an individual who perceives the future as subjectively more distant, he will be more reluctant to delay consumption than another individual who perceives the same length of time as subjectively closer. If time perception partly explains the variation in the observed intertemporal choices, then the "time preference" elicited from intertemporal decisions is actually a compound of both preference and perception, resulting in an identification problem.[2]

To address our research question, we conduct an experiment in a controlled laboratory. Because the impact of *cognitive load* on time perception has been well documented in the literature (reviewed in Section 2), we use cognitive load as an external stimulating instrument to induce differences in time perception across treatments in our experiment.[3] The goal is to test at the individual level whether a change in cognitive load influences the time perception and, furthermore, whether the change in time perception affects the intertemporal choices that are supposed to reveal subjects' time preference. Our conjecture is that subjects with low cognitive load perceive objective time as subjectively longer and will thus be less willing to delay a reward and exhibit higher impatience.

In the first part of the experiment, we elicit subjects' time perception following the design in Brocas, Carillo and Tarrasó (2018a), hereafter BCT, where subjects are instructed to produce

[1] We refer to the term "time perception" as the perceived length of time following the measure in Brocas, Carrillo and Tarrasó, 2018a).

[2] Given that preferences are conventionally reckoned to be deep-rooted, recognizing this problem may help understand the puzzling findings concerning the instability of the intertemporal choices made by an individual's multiple selves over time (Banks, Blundell and Tanner, 1998; Meier and Sprenger, 2013).

[3] By *cognitive load*, we refer to the amount of mental effort demanded by a task, which is also called *cognitive busyness* (see, e.g., Moskowitz, 2005).

time intervals between 24 and 219 seconds by clicking the start and the stop button on the screen. They are simultaneously asked to complete a series of filler tasks. Each filler task has a time limit. We manipulate the cognitive load by varying the time limit across treatments, so that the time limit for the high load treatment (i.e., 10 to 15 seconds, identical to the setting in BCT's experiment) is shorter than that in the low load treatment (i.e., 18 to 23 seconds). In other words, subjects in the high load treatment need to solve a series of filler tasks with a higher intensity than subjects in the low load treatment, which generates at a higher level of cognitive busyness. In the second part, subjects' intertemporal choices are elicited using the standard Multiple Price List design, the Convex Time Budget design (Andreoni and Sprenger, 2012a; hereafter AS) or the Time Trade-off Sequence design (Attema et al., 2010; hereafter ABRW). We look into the treatment effect of cognitive load on subjects' time perception and find that subjects' perceived time is significantly longer in the low load treatment than in the high load treatment. Moreover, a significant and negative effect of time perception on the discounting choices is found: In general, subjects with shorter perceived time and those who underestimate time value the delayed reward more than the others, although the correlation between perceived time length and time discounting is affected by the cognitive load treatment. Based on these results, we conduct mediation analyses and find that time perception mediates a significant and positive indirect effect on time discounting, consistent with our conjecture, while the direct effect and the total effect of the cognitive load treatment are both negative and insignificant.

Our study suggests that an individual's perception of time affects his evaluation of intertemporal tradeoffs. While preferences may be economic primitives that are intrinsically built, intertemporal choices are partly driven by the malleable time perception. We provide evidence demonstrating the gap between the observed intertemporal choices and time preference, pointing to the potential identification problem of time preference.

The rest of the paper is organized as below. Section 2 reviews the related literature. Section 3 introduces a simple theoretical framework as our preliminaries. Section 4 presents the experimental design and procedure. Section 5 discusses the experimental results. The last section concludes.

2 Related Literature

Over the past few decades, economists have revealed individuals' time preference through intertemporal decisions: People act intertemporally as if they discount future payoffs. In the literature of time preference elicitation, the existing attempts have largely relied on experimental techniques, among which the Multiple Price List method, with a relatively simple procedure, has been widely adopted in a large number of applied problems (see, e.g., Tanaka, Camerer and Nguyen, 2010). The Convex Time Budget method designed in AS is a recently developed leading alternative. Besides, other approaches emerge in various versions of experimental designs (see, e.g., ABRW and Attema et al., 2016). However, most, if not all, of the designs depend on systematically varying the amounts and/or the dates of payoffs to infer information on decision-makers' preferences. Cohen et al. (2020) review studies that measure the time preference and report that "money earlier or later" decisions are driven in part by some factors that are distinct from the underlying time preferences. While this type of design implicitly relies on the assumption that intertemporal choices completely represent time preference, we argue that the time discounting behaviors obtained in this manner actually confound both the actual time preference and the perception of time.

The idea of segregating belief/perception and preference dates back to Savage (1954) in the context of eliciting an individual's subjective probability. He highlights the importance of isolating preference from subjective probability to identify the belief: If an individual cares more about outcomes in one state than in another, we cannot tell whether it is because his utility is more sensitive in that state or because he considers that state more likely to occur. Similar questions have also been raised in the studies on risk preference. Weber and Milliman (1997) provide empirical evidence that segregates the two driving forces of choices under risk: the relatively stable risk attitude and the perception of riskiness.[4] Andreoni and Sprenger (2012b) distinguish time preference from risk preference in intertemporal choices and suggest that present-biased behavior could originate from the uncontrolled risk of future payoffs, a contention strengthened by the experimental evidence in Miao and Zhong (2015). Across a number of choice problems, Choi et al. (2014) discuss the decision-making quality and the identification problem of distinguishing decision-making ability from preferences. Along the

[4]Relatedly, Gui, Huang and Zhao (2020) differentiate risk-seeking preferences and biased perception of risk in financial investment settings. Chew and Li (2017) empirically discriminate between social preferences and probability bias as determinants of investors' non-neutral attitude towards the perceived morality of a stock.

same line, the present paper focuses on isolating belief, i.e. perception in our context, to identify preferences by stressing the two independent driving forces of intertemporal choices: time preference and time perception.

The determinants of the underlying time preference and its associations with other human attributes and environmental factors have long attracted attention in the economics literature (Thaler and Shefrin, 1981; Loewenstein and Prelec, 1992; Becker and Mulligan, 1997). More recently, the evolution theory of time preference developed by Robson and Szentes (2008) and Robson, Szentes and Iantchev (2012) consider the evolutionary basis of time preference with an optimization process of resource allocation and show how intergenerational transfers and sex determine the patience level.[5] Galor and Özak's (2016) empirical findings suggest that the difference in time preferences across regions may be rooted in the geographical variations in the natural return to agricultural investment. Chen (2013) empirically studies whether languages grammatically associating the future and the present influences time preference and related economic behaviors.[6] This line of research explains the observed differences in time preference at a more aggregate level, as most of the discussed features (language, cultivation, etc.) are deeply ingrained constructs within a population group. We depart from this literature and test at the individual level the effect of time perception that not only varies across individuals but also wobbles among different selves of the same individual over time.

The economics and the psychophysics literature have documented the heterogeneity in time perception (Eagleman, 2008; Grondin, 2010; Bonato and Umiltà, 2014; BCT). Despite its fluid nature, a direct manipulation of time perception is not easy: Attempts have been made using noninvasive brain stimulation techniques (Mioni, Grondin and Bardi, 2019), but they either due to mechanisms still obscure or difficult to implement, especially in an economic laboratory. We draw on the findings in the literature and propose to use cognitive load as the external stimulus to implement the manipulation for two reasons: First, the effect of cognitive load on time perception has been extensively studied and proved significant. Alonso, Brocas and Carrillo (2013) theoretically explain "mistakes" in decision-making with resource allocation under neurophysi-

[5]The evolution of the time preference framework denominates time preferences in resource terms, corresponding to the marginal rate of intertemporal substitution in consumption, while time preference in our paper is denominated in utility terms, close to what Robson, Szentes and Iantchev (2012) call the "pure" rate of time preference.

[6]The effect of time-inconsistent preferences on motivated cognition comprises another part of the literature: For example, Bénabou and Tirole (2002) theoretically link time-inconsistent preferences to self-confidence; Hong, Huang and Zhao (2019) show that the sunk cost "fallacy" is a rational response to hyperbolic discounting. In both theory and experiment, Chew, Huang and Zhao (2020) establish that time preference is associated with motivational value of positive false memory.

ological limitation, suggesting that the time perception could be biased if attentional resources allocated to the time-keeping task are affected. Empirical studies also agree on the central role of attention and mental workload in temporal experience (Thomas and Weaver, 1975; Tse et al., 2004; Block, Hancock and Zakay, 2010).[7] Second, as the effect of time perception is what we desire to address, ideally the stimulating instrument for manipulation should impact time perception while exerting no influence on the discounting behavior itself. Compared to other candidates (say emotion for example, which has been proved to significantly affect intertemporal discounting, see, e.g., Ifcher and Zarghamee (2011)), evidence on whether the cognitive load affects intertemporal tradeoffs is ambiguous. Some even find little or no evidence on the effect of cognitive load on impatience, including Franco-Watkins, Rickard and Pashler (2010) and a survey study by Deck and Jahedi (2015). Nonetheless, we employ the statistical method of causal mediation analysis to encompass the potential direct effect of cognitive load and to isolate it from the indirect effect mediated by time perception which is the center of our interest.

Finally, our study is by no means the first attempt to relate time preference to time perception. There have been increasing studies discussing the relationship between time perception and intertemporal decision-making. Hornik (1984) is among the first few studies to discuss the importance of considering time perception in time-related consumer behavior studies. Wittmann and Paulus (2008) theorize a relationship between impulsivity and time experience, and Read (2001) proposes a model that explains the subadditivity in time discounting by the subjective evaluation of time. Ebert and Prelec (2007) experimentally observe the influence of time sensitivity on discounting behavior. More of the related studies focus on the relationship between time perception and time inconsistency. Starting from Roelofsma (1996) applying Weber's Law to explain time-inconsistent preferences, theoretical efforts include Takahashi's work (Takahashi, 2005; Takahashi, 2009) on the theoretical framework of dynamic inconsistency, and Cui (2011) who demonstrates that hyperbolic discounting results from the scalar property of time perception. Ray and Bossaerts (2011) propose a model linking present bias to the phenomenon of biological time, arguing that the discount function only appears to be hyperbolic against objective time but is indeed exponential against subjective time. Experimental studies also

[7]Studies in this literature mainly use two production paradigms: the *prospective paradigm,* in which participants are ex ante aware that they need to make duration judgments, and the *retrospective paradigm,* in which participants are asked to make duration judgments only after the duration ends. The literature finds different directions of the cognitive load effect under different paradigms: negative under the prospective, and positive under the retrospective paradigm (Block, Hancock and Zakay, 2010). The difference is possibly caused by the different neural processes governing subjective time, which is not the emphasis of this paper. We follow the time evaluation literature and focus on prospective timing in this paper.

reveal the correlation of hyperbolic discounting with time perception and characteristics of it (Zauberman et al., 2009; Kim and Zauberman, 2009; Han and Takahashi, 2012; Bradford, Dolan and Galizzi, 2013). Compared to the literature, the focus of this paper is not directly on the inconsistency property of time preference, but more fundamental: Will longer (shorter) time perception lead to lower (higher) valuation of the delayed reward because the future is perceived as farther (nearer)? In this regard, BCT's experiment proves the negative correlation between time perception and time discounting. Resting on the existing findings, the aim of our study is to move one step forward and provide initial evidence for the *causal* relationship between time perception and time discounting by conducting a controlled laboratory experiment with treatment manipulations.

3 Preliminaries

To explicate the relationship among time perception, time preference and intertemporal choices, and provide a premise for our experimental investigation, we present a simple framework as the theoretical preliminary elaborating on the motivation for this paper and our conjecture.

Defining time preference. While the notion of *time preference* is often considered exchangeable with the notion of *time discounting*, in this paper, we follow Frederick, Loewenstein and O'Donoghue (2002) and distinguish between the two terms to stress the distinctions among various driving forces that underlie intertemporal choices. Time preference is commonly defined as the relative valuation of a reward achieved at an earlier time point versus a reward achieved at a later time point. In this framework given an outcome stream $(\tau_1 : x_1, ..., \tau_m : x_m)$ yielding outcome x_j at a time τ_j for $j = 1, ..., m$, the discounted utility is given by:

$$DU(\tau_1 : x_1, ..., \tau_m : x_m) = \sum_{j=1}^{m} \varphi(\tau_j)U(x_j) \tag{3.1}$$

where $U(\cdot)$ is the instant utility function and $\varphi(\cdot)$ is the discount function.[8] Based on the definition, we represent the time preference by the discount function $\varphi(\cdot)$ that evaluates the delayed utility discounted against the actual time. We assume that $\varphi'(\cdot) < 0$, suggesting that instantaneous utility is discounted more when the reward is realized farther in the future.

[8]In our simple framework, we abstract from the issue of utility representation of the underlying preference relation in decision theory because the axiomatic approach is not the focus of this paper. Nonetheless, Choi et al. (2014) show that an individual's utility representation capturing his or her ability to make decisions survives under sufficient wealth, among other personal characteristics.

Defining time discounting. If actual time preference is defined as how the delayed utility is discounted against the actual time, what intertemporal choices reveal is how the delayed utility is perceived to be discounted against the perceived time instead, which may be biased. Therefore, revisiting Eq. (3.1), the discounted utility is correctly measured if and only if the subjective time is accounted for in the discount function estimation based on the intertemporal choices:

$$\begin{aligned} DU(\tau_1 : x_1, ..., \tau_m : x_m) &= \sum_{j=1}^{m} \varphi(\tau_j) U(x_j) \\ &= \sum_{j=1}^{m} \tilde{\varphi}(p(\tau_j)) U(x_j) \end{aligned} \tag{3.2}$$

where $p(\tau_j)$ describes the perceived time p dependent on the actual time τ_j. However, the subjective time is not typically applied to the derivation; hence, what has been evaluated is the perceived discounted utility such that:

$$\tilde{DU}(\tau_1 : x_1, ..., \tau_m : x_m) = \sum_{j=1}^{m} \tilde{\varphi}(\tau_j) U(x_j) \tag{3.3}$$

Similarly we assume that $\tilde{\varphi}'(\cdot) < 0$. We define time discounting as $\tilde{\varphi}(\cdot)$ to encompass time perception as part of the reason for temporal discounting decisions.[9] Accordingly, time discounting in this paper is the discount function estimated from the observed intertemporal choices applying the objective time. Comparing Eqs. (3.2) and (3.3) implies that intertemporal choices completely reveal the time preference, i.e. $\tilde{\varphi}(\tau_j) = \varphi(\tau_j)$, if and only if time is perceived accurately, i.e. $p(\tau_j) = \tau_j$. Otherwise, the estimated time discounting $\tilde{\varphi}(\cdot)$ represents merely a compound of the time preference and the misperception of time.

Consider the following example in which we assume exponential discounting for $\varphi(\cdot)$ and $\tilde{\varphi}(\cdot)$ for the sake of simplicity. Suppose that between two options of A) receiving \$5 now and B) receiving \$10 $\bar{\tau}$ units of time later, Alice chooses Option A and Bob chooses Option B. The seemingly impatient choice of Alice compared to Bob could then be represented by $\tilde{\varphi}_A(\cdot) < \tilde{\varphi}_B(\cdot)$ where the subscript A (B) denotes Alice (Bob).

However, if we want to study the time preference, note that there is one degree of freedom in Eq. (3.2). On the one hand, time perception could be normalized as $\bar{p} = p_A(\bar{\tau}) = p_B(\bar{\tau})$. Then

[9]The approach to defining time discounting in this paper follows Frederick, Loewenstein and O'Donoghue (2002), but we adopt the belief-based model and thus focus on time perception among the driving forces behind intertemporal decisions.

by $\varphi_i(\cdot) = \tilde{\varphi}_i(p_i(\cdot)),\ i = A,\ B$, we have $\varphi_A(\cdot) < \varphi_B(\cdot)$ which suggests that Alice is inherently more impatient than Bob. On the other hand, if we normalize the discount function, i.e. let $\varphi_A(\cdot) = \varphi_B(\cdot)$, then since $\tilde{\varphi}'(\cdot) < 0$, instead of having different preferences, we have different perceptions, i.e. $p_A(\bar{\tau}) > p_B(\bar{\tau})$, which suggests that the perceived "impatience" of Alice could in fact be attributed to her longer perception of time. That is to say, had they both perceived $\bar{\tau}$ equally, we would not have observed a difference in their time preferences. Eq. (3.2) implies that even with the actual time preferences being the same, a difference in time perception could lead to different intertemporal tradeoffs if the gap between the subjective and objective time is not properly addressed.

4 Experimental Design and Procedure

Based on the preliminaries, our experiment is designed to test whether there is a causal effect of time perception on intertemporal choices. The experiment was composed of two main parts: time perception (TP), followed by time discounting (TD).

4.1 Design

4.1.1 Time perception

Building on BCT, we designed ten rounds of time perception exercises preceded by one practice round. In the official rounds, subjects were asked to finish 10 prospective *time production tasks* with intervals τ of 24, 31, 41, 53, 69, 89, 116, 151, 196 and 219 seconds, one in each round with the order randomized but the same for all subjects. At the beginning of each round, we presented subjects with the time interval τ to be produced in this round. After that, subjects produced their time estimates for this interval τ by clicking the button "START" and the button "STOP" on the screen respectively. The duration between the two clicks marked the subject's estimation of the time interval (see the details in the screenshots in Appendix B.1).

We manipulate the cognitive load to induce differences in the time perception across treatments. During their estimation of the time interval, subjects were asked to solve a series of *filler tasks*. In each filler task, a 4×6 table was shown to the subjects. As in Figure 1, each row and column of the table had a name. Subjects were instructed to click on one of the cells. The order of the row names and of the column names, as well as the bold parts of the instruction, were randomly changed from one filler task to another. Subjects were told to complete each filler task within a *time limit* that was randomly determined and unspecified to them, and the failure to complete a task within the time limit would be counted as an incorrect answer. The ongoing filler task during which the subjects clicked the "STOP" button would not be included for counting.

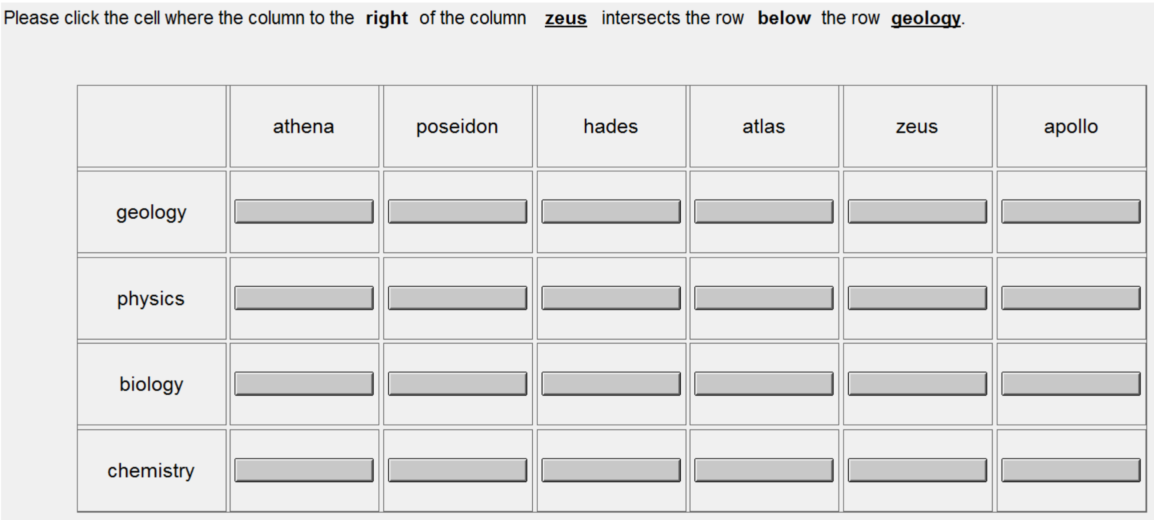


Figure 1: Example of a Filler Task

The time limit was varied to manipulate the cognitive load across treatments:

- High load treatment: The time limit is randomly decided between 10 and 15 seconds.
- Low load treatment: The time limit is randomly decided between 18 and 23 seconds.

In the high load treatment, the intensity of the filler tasks is higher than that in the low load treatment, implying a higher frequency for subjects in the high load treatment to apply mental efforts to solve the filler tasks. [10]

The amount earned depended on subject performance in both the time production tasks and the filler tasks. For each subject, one round was randomly selected to calculate the payment. A subject was required to correctly solve at least 80% of the filler tasks in this round in order to get paid. Then, the more accurate the estimation, the more he or she earned: The subject would be paid HK$120 if the time estimate produced in this round was within ±5% of the actual length of τ, HK$60 if within ±10% and HK$30 if within ±20%. A subject earned nothing if he or she failed in more than 20% of the filler tasks in this round regardless of how accurate the time estimate was.[11]

[10] As an anonymous reviewer noted, it might be more difficult for subjects in the high load treatment than subjects in the low load treatment to count seconds. The influence, however, is two-way and would not systematically lead to longer or shorter perceptions of time. Therefore, the impact, if any, should not bias the result.

[11] We adopted this incentive scheme for its ease of explanation compared to the (incentive compatible) quadratic

4.1.2 Time discounting

To prevent the results from being dependent on the setting of a specific experimental design, we adopted three well-established time discounting elicitation methods that all rely on intertemporal choices: the MPL method, the CTB method designed by AS and the Time-Tradeoff (TTO) Sequences method designed by ABRW.[12] Subjects were assigned randomly to the MPL, the CTB or the TTO group.

Multiple Price List. Subjects in this group were provided with a fixed array of paired options such that

- Option A: Receive κ in u days
- Option B: Receive γ in $u + k$ days

and asked to choose one for each of the 30 rows. The first 15 rows are proximal tasks in which $u = 0$ (i.e. today) and $k = 31$, and the last 15 rows are distal tasks in which $u = 90$ and $k = 31$. $\kappa = 100$ for all 30 rows and γ ranges from 100 to 142 for every 15 rows (see the details in the screenshots in Appendix B.2.1). We allowed subjects to switch back and forth as they wished but no individual with multiple switching was observed in our sample.[13]

Convex Time Budget. We closely followed the design and the allocation procedure in AS. Subjects were asked to allocate 100 experimental tokens either to a sooner time u with exchange rate a_u, or a later time $u+k$ with a_{u+k}. The gross interest rate $1+r$ was determined by the relative exchange rate a_{u+k}/a_u. We denote the amounts allocated at dates u and $u+k$ by x_u and x_{u+k}. The same 3×3 design as in AS was implemented in our experiment, with the starting time $u \in \{0, 7, 35\}$ in days and the delay $k \in \{35, 70, 98\}$ in days, each pair consisting of 5 different exchange rates, summing up to a total of 45 allocation tasks. The allocations were made on the computer by dragging the scrollbar on a slider (see the details in the screenshots in Appendix B.2.2).

To avoid the convenience gained by concentrating payments in one period, all subjects were informed in advance that for the subject who was randomly selected to be paid for the CTB

scoring rule.

[12] The instructions for the TD task were only presented on the computer screen without oral explanations in public by the experimental instructor. Otherwise, the difference in waiting time for the TD task would affect subjects' time perception following the TP task, and thus contaminate the treatment effect.

[13] Relatedly, Yu, Zhang and Zuo (2021) experimentally suggest that multiple switching is not as important as discussed in the literature. Subjects may simply misunderstand the tasks, and a nudge treatment can generally improve subjects' comprehension.

task, the participation fee would be paid with half at the sooner and half at the later date regardless of their choices.

Time-Tradeoff Sequences. A TTO sequence is a sequence of time points u_0,...,u_n that fulfills the following indifferences with the two outcomes $\kappa < \gamma$:

$$\begin{aligned} (u_0 : \kappa) &\sim (u_1 : \gamma) \\ (u_1 : \kappa) &\sim (u_2 : \gamma) \\ &\vdots \\ (u_{n-1} : \kappa) &\sim (u_n : \gamma) \end{aligned} \tag{4.1}$$

The design of the task took an analogous form to the standard MPL, except that the amount of payment was predetermined, and subjects were asked to choose the indifferent payment date.[14] Specifically, subjects were asked to choose between two options in each line of a table, such that:

- Option A: Receive κ in u_{n-1} days;
- Option B: Receive γ in $u_{n-1} + k$ days.

There were 30 choices to make in each table, with u_{n-1} being the same for all 30 choices and k ranging from 0 to 29 (see the details in the screenshots in Appendix B.2.3). Therefore, in contrast to the standard MPL, the future payment date in our experiment was not preset and could vary across all subjects. A unique switching-point was enforced in each table.

In our experiment, subjects were asked to produce four sets of sequences with the parameters in Table 1:

Table 1: Parameters for the Four TTO Sequences

Sequence	u_0	κ	γ
I	1 day	HK\$160	HK\$180
II	1 day	HK\$610	HK\$650
III	6 days	HK\$160	HK\$180
IV	1 day	HK\$350	HK\$380

Every sequence consisted of five steps, yielding $n = 5$ in the above indifferences. To prevent any potential order effect, we first elicited u_1 for every TTO sequence, then u_2 for every TTO sequence with the order of sequences randomized, and so forth.

[14] A similar example is in Attema et al. (2016).

In all groups, one out of every twenty-five subjects was randomly selected at the end of the experiment and one of his (her) choices in the TD part would be played out for real.

Given the preliminaries and the experimental design, we seek to test the following hypotheses:

Hypothesis 1: *The objective time is subjectively perceived to be longer in the low load treatment than in the high load treatment.*

Hypothesis 1 is formed based on the findings in the literature (Block, Hancock and Zakay, 2010), and complies with intuition: As subjects are more occupied in the high load treatment, we expect them to mentally experience more in a unit of time, and thus produce a shorter interval than subjects in the low load treatment when asked to estimate a certain duration.

Hypothesis 2: *Time perception and time discounting are negatively correlated.*

In accordance with the results in BCT, we hypothesize that when the future is perceived to be more distant (high TP), people's relative valuation of the future reward is lower (low TD) and thus may appear to be more impatient.

Hypothesis 3: *Subjects exhibit more (less) patience because of the shorter (longer) perception of time induced by the high (low) cognitive load treatment.*

Presuming Hypothesis 1 and 2, we will further test whether the change in the cognitive load indirectly leads to the change in the observed patience with time perception as the mediator. In addition to the indirect effect, cognitive load could also directly affect intertemporal choices. If the direct effect goes in the same direction, or there is no direct effect as demonstrated by the literature, then we may expect Hypothesis 4 regarding the total effect of cognitive load on the observed time discounting:

Hypothesis 4: *Subjects in the high load treatment exhibit more patience in their intertemporal choices than subjects in the low load treatment.*

4.2 Procedure

A total of 221 student subjects participated in the experiment, with 51 from Hong Kong and 170 from Mainland China.[15] The experiment was conducted using z-Tree (Fischbacher, 2007).

[15]Subjects were university students recruited from Hong Kong University of Science and Technology in Hong Kong, Wuhan University and Southern University of Science and Technology in Mainland China. The monetary incentives were adjusted such that HK$3 : RMB 1 for the sessions conducted in Mainland China. While the

The experiment was first conducted with the TP part followed by the TD part, and there was then a *last stage* where the final payment was revealed to every subject. At the beginning of the experiment, subjects were provided with an instruction for the TP part. In the instructions, subjects were informed that there would be a second part without knowing specifically what the part would be. They were also told that before reaching the last stage, they were allowed to proceed without waiting for other subjects. The instructions for the TD part were provided on the computer individually after the completion of the TP part. After completing the TD part, subjects were asked to wait for others before proceeding to the last stage.

For subjects in Mainland China, all payments were made at the appointed time through WeChat.[16] For subjects in Hong Kong, all payments for the TP task were made immediately after the experiment in cash, while the randomly selected payment for the TD task were paid by cash delivery: The selected subjects were asked to provide an address that would be convenient for them to receive cash on the payment day, and the cash was delivered to the specified address on that day. The exact amount of cash was included in an envelope and delivered in person. Subjects were provided with Professor Xiaojian Zhao's name card and were promised that if they failed to receive the payment on the payment day, they could contact the professor and the cash would be delivered in person immediately.[17] The payment method was chosen to minimize the difference between receiving payments on weekdays and weekends.

4.3 Discussion

It is worth discussing two challenges in experimental studies of time perception and time discounting.

First, since we seek to test a causal hypothesis, we rely on experimental manipulations of treatments. However, despite the multiple conjectures, the mechanisms that govern time perception are not yet fully understood, so direct manipulations of time perception via nerve cell stimulation would be unmanageable (Fontes et al., 2016; Mioni, Grondin and Bardi, 2019). Hence, we intend to influence subjects' time perception by exogenously varying the cognitive load as the stimulating instrument. The problem of such indirect manipulation is that the

payment appears to be lower in Mainland China, it can still provide sufficient motivation, as a regular meal in an university canteen in Mainland China costs only approximately US$ 1. The currency difference is included in the regression analyses in Section 5 to control for its potential effect on intertemporal choices.

[16]WeChat is the most popular online social network and payment platform in China. See, e.g., Chen, Hong and Zhao (2019) for online experimentation via WeChat.

[17]This was intended to boost subjects' confidence in the future payments to minimize the impact of risk preferences.

change in cognitive load itself would exert direct influence on subjects' intertemporal decisions. We will thus use causal mediation analysis to distinguish the indirect effect mediated by time perception from the direct effect of cognitive load, and check whether the mediation effect is significant.

Second, the temporal horizons differ in the TP part and in the TD part. Ideally, it would be preferable to match the time durations used in the two parts. While economically relevant intertemporal tradeoffs involve at least temporal delays in the range of days, it is infeasible to objectively test time perception with reliably incentivized tasks in such ranges in a laboratory. In practice, we can only elicit time estimates of multiple intervals in the range of seconds and minutes. Unavoidably, we have to extrapolate our estimates downwards for the time discounting and upwards for the time perceptions. The modeling of time discounting and time perception helps us determine the most reliable extrapolations.

5 Experimental Result

In this section, we present the statistical analysis by first summarizing the results of the TP part and of the TD part, then on the basis of this, we elaborate the causal mediation analyses conducted to distinguish the causal mediation effect of time perception on time discounting.

5.1 Time perception

Table 2 presents the summary statistics of the TP part. The average filler task performance, defined as the percentage of filler tasks that were correctly answered in a round, exceeds 80% in every round, and the performance does not contrast between treatments, implying that the time limit is long enough for a subject to solve the filler tasks even in the high load treatment; thus, the perception difference between treatments can be induced by the difference in cognitive busyness generated by the mental effort inputs required to solve filler tasks. Additionally, on average all produced time intervals are longer than the intended ones, suggesting that overestimation of the time occurs exceedingly more than underestimation. It is also noteworthy that, among all the time estimates produced by the subjects, only approximately 16% are within $\pm 5\%$ of the real length, indicating that the inaccuracy in time perception is rather common.

To compare the time perceptions of different time intervals, we derive the subjective-to-

Table 2: Summary Statistics

τ	All		Low load		High load	
	Perception	Filler task performance	Perception	Filler task performance	Perception	Filler task performance
24s	30.88	89.01	31.70	88.92	30.04	89.10
	(9.62)	(29.92)	(10.94)	(30.08)	(8.00)	(29.90)
31s	39.62	86.84	41.09	89.16	38.09	84.41
	(12.96)	(31.96)	(13.89)	(29.48)	(11.78)	(34.33)
41s	51.26	93.14	52.71	93.38	49.74	92.90
	(15.63)	(23.51)	(16.14)	(23.31)	(15.00)	(23.81)
53s	66.72	92.43	70.06	93.44	63.22	91.37
	(20.76)	(23.63)	(22.33)	(22.00)	(18.45)	(25.28)
69s	85.24	86.67	85.15	87.70	85.32	85.60
	(30.01)	(28.70)	(30.31)	(28.76)	(29.82)	(28.73)
89s	109.25	91.68	113.15	92.28	105.17	91.05
	(35.60)	(24.41)	(36.10)	(23.79)	(34.76)	(25.13)
116s	141.73	89.85	148.99	91.26	134.13	88.38
	(45.75)	(26.09)	(46.99)	(24.29)	(43.34)	(27.88)
151s	184.94	91.38	194.18	93.05	175.28	89.62
	(64.48)	(23.71)	(65.81)	(22.22)	(61.88)	(25.16)
196s	242.35	92.22	252.72	93.19	231.50	91.21
	(76.92)	(22.40)	(77.65)	(20.69)	(74.98)	(24.10)
219s	260.94	91.68	274.78	92.50	246.33	90.83
	(82.17)	(23.48)	(82.81)	(22.51)	(79.27)	(24.53)
	221		113		108	

Notes:

1) Means are presented with standard deviation in parentheses.

2) *Perceptions* are the estimated time intervals produced by subjects.

3) *Filler task performances* are the percentages of the filler tasks correctly answered in a round.

4) $t - tests$ show that there is no significant difference between the filler task performances across treatments, and $F - tests$ show that there is no significant difference across treatments in terms of the variance of the perceptions.

objective (S/O) duration ratio for each subject in each round:

$$SO_{is} = \frac{p_i(\tau_s)}{\tau_s}$$

where $p_i(\tau_s)$ is the reported time perception of subject i in round s producing time interval τ_s with one second as the unit of time. The statistical comparison of the S/O ratio shows that the reported perceptions in the low load treatment are significantly longer than those in the high load treatment (Mann-Whitney U test, $Z - statistics = 6.486$, $p - value = 0.0000$). We also probe how the direction of the perceptual bias varies across treatments by dividing the observations between the overestimated intervals (i.e., $SO_i(\tau_s) > 1$, denoted as $OVER_i(\tau_s) = 1$) and the underestimated intervals (i.e., $SO_i(\tau_s) \leq 1$, denoted as $OVER_i(\tau_s) = 0$).[18] The comparison between the treatments shows that there are significantly more intervals being overestimated in the low load treatment than in the high load treatment (Mann-Whitney U test, $Z - statistics = 6.415$, $p - value = 0.0000$).

As the time estimation tasks in the TP part elicit perceived durations in the range of seconds, whereas the intertemporal choices elicited in the TD part are in the range of days, we need to extrapolate upwards for the time estimates for temporal horizons to be comparable between time perception and time discounting. The nonlinear relationships between the magnitude of a physical stimulus and its perceived strength developed by psychophysicists, for example the Weber-Fechner law (Fechner, 1860), provide potential solutions. In this paper, we rely on Steven's power law to derive the relationship between an individual's subjective perception of time and the actual time interval.[19] For each individual i, we use the data collected from the TP part and estimate the parameters a_i and b_i in the following regression by the nonlinear least squares (NLS) estimation:

$$p_i(\tau_s) = a_i \tau_s^{b_i} + \varepsilon_{is} \tag{5.1}$$

where $s \in \{1, ..., 10\}$, $\tau_s \in \{24, 31, 41, 53, 69, 89, 116, 151, 196, 219\}$, $p_i(\tau_s)$ is subject i's reported perception for the interval τ_s, and the error term $\varepsilon_{is} \sim N(0, \sigma^2)$. The average $\hat{a}_i = 1.728$, and the average $\hat{b}_i = 0.982$. The power model fits the data remarkably well: The average R^2

[18]Strictly speaking, we do not observe any single case of perfect time perception, i.e. $SO_i(\tau) = 1$, in our sample.

[19]The Weber-Fechner law is an initial human perception modeling that posits a logarithmic relationship between the actual change in a physical stimlus and the perceived change of it, and Steven's law generalizes it by positing a power relationship (Stevens, 1957). The former performs poorly in our data: For the specification $p_i(\tau_s) = a'_i \log(\tau_s + 1) + \varepsilon_{is}$, the average adjusted R^2 is 0.799 and no individual is with an adjusted R^2 greater than 0.9.

is 0.983 and more than 95% of the subjects have an R^2 greater than 0.944. While subjects react similarly steeply to time (with the average $\hat{a}_i$ being 1.738 in the high load treatment and 1.719 in the low load treatment), subjects in the high load treatment have more concave perceptions (with the average $\hat{b}_i = 0.961$) than subjects in the low load treatment (with the average $\hat{b}_i = 1.002$), suggesting that the difference in time perception between treatments tends to amplify over longer ranges of time.

The validity of the extrapolation exercise relies on the experimental findings in the psychophysics literature. Although experimental studies on human's duration judgement in units of hours or days are scarce in the literature, the investigation of time management's role in time-based prospective memory and time estimation (Francis-Smythe and Robertson, 1999; Waldum and McDaniel, 2016), as well as the extended time perception model based on neural networks developed by Maniadakis and Trahanias (2016), implies that the time perception over both short and long periods of time are governed by a coherent underlying mechanism.[20] Therefore, the estimated time perception $\hat{p}_i$ enables us to check the extrapolated S/O ratio and the time over-/underestimation bias for longer time intervals, as we describe below.

5.2 Time discounting

We conjecture that there is an identification problem as long as the "time preference" is elicited through intertemporal decisions. Thus, we employ different time discounting elicitation designs with different estimation strategies but all depend on intertemporal choices: the MPL design, the CTB design and the TTO sequence design. All three designs provide the methodologies to estimate the *discount function* $\tilde{\varphi}_i$ for individual i whose discounted utility is evaluated as:

$$\tilde{DU}_i(u_1 : x_1, \ldots, u_m : x_m) = \sum_{j=1}^{m} \tilde{\varphi}_i(u_j) U_i(x_j)$$

where the time unit for u is a day. We let $\tau = zu$ with $z = \frac{1}{60 \times 60 \times 24}$.

For the 47 subjects in the MPL group and the 56 subjects in the CTB group, we estimated the discount function for each subject. For the 118 subjects in the TTO group, we estimated the discount functions for each subject using each of the four elicited sequences respectively. Appendix C presents the estimation details.[21]

[20]This notion is also supported by animal studies on time perception of intervals above one day (see, e.g., Crystal (2001)'s experiment on rats).

[21]While, by definition, we estimate $\tilde{\varphi}_i(\cdot)$ by applying the objective time to explore the causal relationship, the estimation results to which the subjective time is applied, i.e., $\varphi_i(u) = \tilde{\varphi}(p_i(u))$, and the comparisons between

Table 3: Distribution of the Individual Parameter Estimates

	Mean		Median		5th Percentile		95th Percentile	
	MPL	CTB	MPL	CTB	MPL	CTB	MPL	CTB
Daily discount factor: $\hat{\delta}_i$	0.997	0.997	0.997	0.997	0.993	0.988	1.000	1.000
Present bias: $\hat{\beta}_i$	0.977	0.895	1.000	0.945	0.888	0.692	1.084	1.070
CRRA Curvature: $\hat{\alpha}_i$		0.882		0.998		0.451		0.999

The MPL design is the most commonly adopted method in the literature, as it is cognitively easy for subjects to understand. Following the literature, the quasi-hyperbolic discount function and the linear utility function are assumed for subjects in the MPL group. The discount factors $\hat{\delta}_i$ in our data are close to 1 with an average of 0.997. The present bias estimators $\hat{\beta}_i$, on the other hand, exhibit substantial heterogeneity: While the average of $\hat{\beta}_i$ is 0.977, half of the estimates are above 1, suggesting future bias behavior by some subjects. The estimates we obtained are reasonably comparable to those of Andreoni, Kuhn and Sprenger (2015) with the median $\hat{\delta}_i$ being 0.999 and the median $\hat{\beta}_i$ being 0.995.

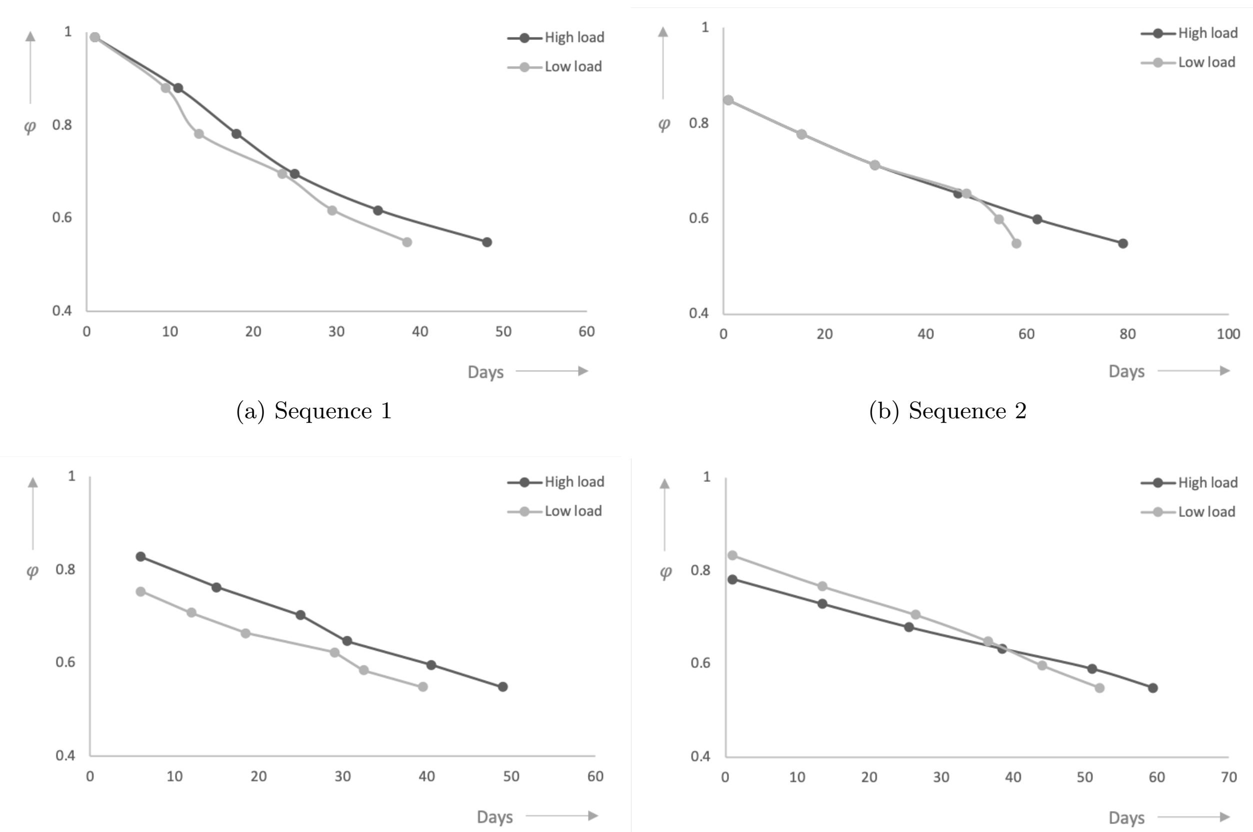
The CTB design is a leading alternative that is intended to provide subjects with an opportunity to smooth payments over time. With the quasi-hyperbolic discount function similarly assumed, the CTB design further assumes the *Constant Relative Risk Aversion* (CRRA) parameter α_i to describe the concave utility. The average discount factor $\hat{\delta}_i$ in our data is 0.997, the same as in the MPL group and conforming to the estimates in AS. The vast majority of the $\hat{\alpha}_i$ estimates are above 0.83 and below 1, also similar to those in AS. For $\hat{\beta}_i$, the majority fall between 0.9 and 1, and the average $\hat{\beta}_i$ is 0.895, smaller than that in AS. The more evident tendency of present bias in our experiment is presumably due to the immediate "today" payments through WeChat, compared to the out-lab "today" payment with several hours delay in AS's experiment. Nevertheless, again, there is a nonnegligible proportion of the estimated $\hat{\beta}_i$ above 1.

Neither the MPL nor the CTB group shows significant difference in parameters between treatments. The discount function that we use in our analysis to describe subject i's time discounting behavior is then calculated as $\tilde{\varphi}_i(u) = \hat{\beta}_i \hat{\delta}_i^u$.

While it entails parameter estimations to analyze the time discounting decisions using the first two designs, the TTO sequence design does not rely on a specified discount function that involves structural estimations for the parameters. Therefore, the discount function elicited in

$\varphi_i(\cdot)$ and $\tilde{\varphi}_i(\cdot)$ are presented in Section 5.3.3, which provides suggestive evidence for the involvement of time perception in time discounting estimation and time preference elicitation.

the TTO group needs to be studied as a whole. Figure 2 depicts the discount curves for each sequence. Note that subjects choose the delay instead of the amount of the delayed payment as in the standard MPL or CTB design; thus, the x-axis values are the median delays, or willingness to wait, over all subjects for each of the 5 elicited steps in each of the 4 sequences, while the y-axis values are the discount functions at each step in each sequence fixed across subjects. In general, we do not observe an apparent pattern of either increasing or decreasing impatience.



(a) Sequence 1

(b) Sequence 2

(c) Sequence 3

(d) Sequence 4

Figure 2: Discount Curves for the TTO group

Note: The x-axis values are the median delays over the 118 subjects for each of the 5 elicited steps in each of the 4 sequences. The y-axis values are the discount functions for each step in each sequence same across all subjects.

5.3 Analysis

As it is established in the previous section, we summarize each individual i's time perception by the parameters $(a_i,\, b_i)$ and approximate his/her time discounting by the discount function $\tilde{\varphi}_i$. In this section, we address the primary question in this paper: the causal relationship between time perception and time discounting.

5.3.1 Empirical method

Our randomized experiment allow us to investigate the total effect of cognitive load by a straightforward estimation of the following linear regression:

$$Y_i = \eta_1 + \theta_1 T_i + \xi_1 X_i + \varepsilon_{1i} \tag{5.2}$$

where T_i is the treatment dummy that equals 1 if individual i was assigned to the high load treatment and 0 otherwise, and X_i represents the controls.[22]

It is considerably more demanding to investigate how and to what extent the observed patience revealed by intertemporal choices for subjects in the high load treatment can be explained by the bias in time perception. We conjecture that the indirect effect through time perception is positive, as the future would be discounted more when it seems to be farther away for people who have a longer perception of time. Our aim is to decompose the treatment effects of cognitive load into the direct effect and the indirect effect mediated by time perception. To this end, we adopt the causal mediation analysis strategy proposed in Imai, Keele, and Tingley (2010) and Imai, Keele, and Yamamoto (2010).[23] We first define the indirect effect, also called the causal mediation effect, as

$$\text{Indirect effect}_i \equiv Y_i(t, M_i(1)) - Y_i(t, M_i(0))$$

where $M_i(1)$ and $M_i(0)$ are the potential value of the mediator should individual i be assigned to the high load treatment and the low load treatment respectively. The indirect effect is then the change in Y_i induced by the change in M_i from the low load treatment to the high load treatment, fixing individual i's treatment status at t. Because we hold the treatment constant and change only the mediator, the objective here is to isolate the treatment effect on the outcome through the mediator. Moreover, by definition, if $M_i(1) = M_i(0)$, namely the treatment has no effect on the mediator, then there is no indirect effect. We further define the average causal mediation effect (ACME) as the mean of this value over all individuals. Correspondingly, the

[22]Note that it is not necessary for the correlation between the treatment and the outcome to be significant (i.e., $\hat{\theta}_1$ does not need to be statistically distinguishable from 0) or in the same direction for the mediation effect to exist, because the direct effect and the mediation effect could be in opposite directions (Imai, Keely and Tingley, 2010).

[23]We choose this estimation strategy because of its advantages in relaxing the no-interaction assumption and applying to both continuous and discrete mediators, which are important for the statistical analyses of our data, as will be explained below.

direct effect of the treatment is defined as:

$$\text{Direct effect}_i \equiv Y_i(1, M_i(t)) - Y_i(0, M_i(t))$$

which encompasses all other mechanisms and captures the treatment effect that is irrelevant to the mediator. In this sense, the direct effect measure the portion of the treatment effect that remains after the indirect effect is taken into account. Similarly, the average direct effect is defined as the expected direct effect over all individuals. The treatment's total effect is then naturally given by the sum of its direct and indirect effects $Y_i(1, M_i(1)) - Y_i(0, M_i(0))$.

As the ACME is of the principal interest in our analysis, we now introduce the strategy of its identification. Our experimental design with random assignment enables us to estimate the total effect of the treatment, but it is insufficient to estimate the ACME because although our data gives the values of $Y_i(1, M_i(1))$ and $Y_i(0, M_i(0))$, the potential outcomes $Y_i(1, M_i(0))$ and $Y_i(0, M_i(1))$ required for the estimation of the ACME are never observable. Therefore to identify the ACME, following Imai, Keele, and Yamamoto (2010), we impose a set of Sequential Ignorability (SI) assumption that consists of two sequentially made ignorability assumptions:

SI Assumption 1. Given the baseline characteristics, the treatment assignment is ignorable. Formally, this assumption can be written as:

$$\{Y_i(t\prime, m), M_i(t)\} \perp T_i \mid X_i = x.$$

That is, the treatment assignment is assumed to be statistically independent of potential outcomes and potential mediators. In the present study, this assumption holds because both treatments are randomly assigned in our experiment.

SI Assumption 2. Given the observed value of the ignorable treatment and the baseline characteristics, the mediator is ignorableable. Formally, this statement can be written as:

$$Y_i(t\prime, m) \perp M_i(t) \mid T_i = t, \ X_i = x.$$

The second assumption requires that there are no omitted variables that simultaneously influence M_i and Y_i, i.e. the mediator is statistically independent of the potential outcome, once T_i and X_i are accounted for. Put differently, this assumption requires that the mediator can be regarded as if it were randomized among individuals who share the same treatment status and pretreatment confounders. This assumption will be violated if there are unobserved variables that confound the relationship between the outcome and the mediator variables. The assumption embodied in SI Assumption 2 is therefore quite strong and thus deserves special attention. We return to this point later.

In addition to its ease of interpretation, another advantage of making such strong assumption is that it provides a generalization to valid estimates of the ACME without being tied to specific statistical models used for the mediator and the outcome. That is, no additional distributional or functional form assumptions regarding the mediator or outcome variables, e.g., the linearity or no-interaction assumption, is required for the identification of the ACME. Under the linearity and the no-interaction assumptions, the ACME can be estimated using the following two-regression system:

$$M_i = \eta_2 + \theta_2 T_i + \xi_2 X_i + \varepsilon_{2i} \tag{5.3}$$

$$Y_i = \eta_3 + \theta_3 T_i + \omega M_i + \xi_3 X_i + \varepsilon_{3i}. \tag{5.4}$$

In this simple case, the ACME is delivered by the product of the coefficients $\hat{\theta}_2 \cdot \hat{\omega}$ estimated in Eqs. (5.3) and (5.4). The average direct effect is given by $\hat{\theta}_3$ from Eq. (5.4). Moreover, the literature (see, e.g., Judd and Kenny, 1981, Kraemer et al., 2002; Kraemer et al., 2008) also shows that the no-interaction assumption can be relaxed by replacing Eq. (5.4) with an alternative specification:

$$Y_i = \eta_3' + \theta_3' T_i + \omega' M_i + \mu T_i M_i + \xi_3' X_i + \varepsilon_{3i}'. \tag{5.5}$$

which allows the ACME to depend on the treatment status. The ACME is then delivered by the product of the coefficients $\hat{\theta}_2 \cdot (\hat{\omega}' + \hat{\mu} t),\ t = 0,\ 1$ using coefficient estimates in Eqs. (5.3) and (5.5), and the average direct effect is given by $\hat{\theta}_3'$ estimated in Eq. (5.5).

In either case, we employ the estimation algorithms proposed by Imai, Keele, and Yamamoto (2010) to test the statistical significance of the ACME, and the significance of the average direct effect. The main idea of their algorithm is to obtain the Monte Carlo draws of the potential outcomes $Y_i(1, M_i(0))$ and $Y_i(0, M_i(1))$ by sampling $M_i(t')$ from the selected mediator model, $f(M_i | T_i = t',\ X_i = x)$, and then given this draw of the mediator, sampling $Y_i(t,\ M_i(t'))$ from the outcome model, $f(Y_i | T_i = t,\ M_i(t'),\ X_i = x)$. The procedure is maintained regardless of statistical models of the mediator and the outcome and thus can accommodate various types of mediators, outcome variables and relationship models.

Because the identification of the ACME requires a strong assumption of SI, it is important to understand how our results change for different degrees of violation of it. Mathematically, the SI assumption implies that the correlation between the error terms ε_{2i} from Eq. (5.3) and ε_{3i} from Eq. (5.4) (or ε'_{3i} from Eq. (5.5) if the interactive effect is included), denoted as ρ, would be zero. Conversely, nonzero values of this correlation would imply that SI assumption has been violated. To quantify the degree to which our empirical findings are robust to a potential violation of the SI assumption, we performed the sensitivity analysis proposed by Imai, Keele, and Yamamoto (2010) and Imai, Keele and Tingley (2010). Although the SI assumption cannot be tested directly (Manski, 2007), the sensitivity analysis allows us to understand how the ACME would change when SI Assumption 2 is violated to different extents. The results of the sensitivity analyses are presented following the estimations of the ACME in Section 5.3.2.

5.3.2 Statistical results

Recall that upward extrapolations of the time estimates and downward extrapolations of time discounting are needed for the temporal horizons to be comparable. However, subjects may behave in different manners with respect to their time perception and time discounting dependent on the horizon due to the potentially nonlinear perception in actual time and the nonexponential time discount functions. Hence, the rankings of extrapolated time perception and the discount function could change with the time horizon we focus on. In this respect, we rely on the power functional form $p_i(\cdot)$ for the time estimates and the discount function $\tilde{\varphi}_i(\cdot)$ and employ the exercise performed in BCT to determine the time ranges for which the individual estimates can be steadily ranked.

To find such a range in the time perception data, we performed the following exercise for the time estimates closely following the procedure in BCT:

Step 1: We evaluated each subject's estimate of an time interval of length τ_v (i.e., $\hat{p}_i(\tau_v) = \hat{a}_i\tau_v^{\hat{b}_i}$) and then ranked all subjects by their time estimates.

Step 2: We repeated Step 1 for another interval of length $\tau_{v'}$.

Step 3: We checked by how many positions this ranking changed between τ_v and $\tau_{v'}$.

As a result, 46% of the subjects' rankings changed by more than 11 positions (5% of all subjects) between 657 seconds (three times the longest interval estimated by our subjects) and

1 hour. Then, it dropped to 33% between 1 hour and 6 hours, to 6% between 6 hours and 12 hours and to 0% thereafter. Overall, the ranking of the time perceptions between subjects can be reasonably preserved for intervals above 6 hours, and with extreme stability for intervals above 12 hours.

We similarly determine the time intervals for which we can stably rank subjects by their time discounting. We evaluated for each subject the value of the discount function $\tilde{\varphi}_i(u)$ at different u: 1 hour, 6 hours, 12 hours, 18 hours, 1 day and 7 days. Then group by group, we ordered our subjects from the maximum to the minimum in terms of their discount functions at each of these time points. We found that 6% of the subjects changed ranks by more than 5% of the positions (i.e., ranks changed by more than 2 for subjects in the MPL group, 3 in the CTB group and 6 positions in the TTO group) between 1 day and 7 days, which drops to 0.8% between 12 hours and 1 day, and no rank is changed between any of the shorter intervals. We then decided that the ranking of subjects' time discounting is stable for all horizons, but preserve extreme steadiness for intervals below 1 day.

In conclusion, this exercise suggests that the time perception and the time discounting of our subjects acquire reasonably consistent rankings for intervals between 12 hours and 1 day.

Table 4: Non-parametric comparisons of the time estimates

Dependent variable:	$u = 0.5$		$u = 1$	
	$\hat{SO}_i(u)$	$\hat{OVER}_i(u)$	$\hat{SO}_i(u)$	$\hat{OVER}_i(u)$
$Z - statistics$	2.693***	2.686***	2.461**	2.597***

Notes:

1) Mann-Whitney U tests are performed to compare the dependent variables.

2) $\hat{SO}_i(u)$ is the extrapolated S/O ratio of u day(s). $\hat{OVER}_i(u)$ equals 1 if subject i over-estimates u day(s) and 0 otherwise.

3) $Z - statistics$ are calculated based on the comparison between $T_i = 1$ for the high load treatment and 0 for others, so $Z - statistics > 0$ indicates a smaller value of the dependent variable for the high load treatment.

4) *, **, *** denote significance at the 0.1, 0.05 and 0.01 level.

We now follow the steps required in Section 5.3.1 to identify the ACME. We first check relationship between the treatment and the potential mediators, i.e., the extrapolated S/O ratios, denoted as $\hat{SO}_i(u)$ measuring the perceived time length, and the corresponding time overestimation, denoted as $\hat{OVER}_i(u)$ measuring the direction of the perceptual bias.[24] The

[24] More precisely, the extrapolated S/O ratio is defined such that $\hat{SO}_i(u) = \frac{\hat{\hat{a}}_i u^{\hat{b}_i}}{u}$, where $\hat{\hat{a}}_i = \hat{a} \cdot z^{\hat{b}_i - 1}$ with z being the unit multiplier defined in Section 5.2 and $\hat{a}_i$ and $\hat{b}_i$ being the time perception parameters estimated in Eq. (5.1). Time overestimation is then defined such that $\hat{OVER}_i(u) = 1$ if $\hat{SO}_i(u) > 1$ and 0 otherwise. Again, we do not observe perfect time perception, i.e., $\hat{SO}_i(u) = 1$, in our sample.

nonparametric comparisons displayed in Table 4 show the significant difference in subject's time perception between treatments. In the formal investigation of the treatment effect, we run the regressions according to Eq. (5.3) of which the results, presented in Table 5, resemble the findings: The effects of the treatment are significant in terms of both the S/O ratio and the time overestimation. Specifically, the S/O ratios are significantly smaller in the high load treatment, indicating shorter time perceptions in the high load treatment than in the low load treatment. Moreover, subjects in the high load treatment are significantly more likely to underestimate time than subjects in the low load treatment. in

Table 5: Regression analyses of the time estimates

	$u = 0.5$		$u = 1$	
	$\hat{SO}_i(u)$	$OV\hat{E}R_i(u)$	$\hat{SO}_i(u)$	$OV\hat{E}R_i(u)$
	(1)	(2)	(3)	(4)
T_i	-.7302**	-.6675**	-.9510**	-.7053**
	(.3177)	(.2765)	(.4407)	(.2765)
R^2	.0232	.0222	.0207	.0242

Notes:

1) Heteroskedasticity robust standard error in parentheses.

2) Column (1) and (2) use OLS regressions. Column (3) and (4) use logistic regressions.

3) $T_i = 1$ for high load treatment and 0 otherwise. $\hat{SO}_i(u)$ is the extrapolated S/O ratio of u day(s). $OV\hat{E}R_i(u)$ equals 1 if subject i overestimates u day(s) and 0 otherwise.

4) The average performance of the filler tasks is controlled for.

5) *, **, *** denote significance at the 0.1, 0.05 and 0.01 level.

Along with the comparison results of the reported perceptions we present in Section 5.1, the evidence supports Hypothesis 1:

Result 1: *Subjects' time perceptions are significantly longer in the low load treatment than in the high load treatment.*

Result 1 shows that subjects' cognitive load strongly affects their time perception and is thus qualified to serve its instrumental purpose of manipulation. The significant differences in the time estimates between treatments verify that our intention to use cognitive load to manipulate time perception is successful.

Having demonstrated the significant effect of the treatment on the potential mediator, we then examine the treatment effect on the outcome variables according to Eqs. (5.2), (5.4) and (5.5). Table 6 presents the regression analyses of the time discounting, in which Columns (1) - (5) present results for the interval of 12 hours, and Columns (6) - (10) present those for 1 day.

The total effect is insignificant, as suggested by the coefficients of T_i in Columns (1) and (6). Then we include the mediators for which we consider two candidates. Columns (2) and (7) regard the S/O ratio as the first candidate and show that the correlations between the S/O ratio and the discount function are negative with marginal significance. Columns (3) and (8) further include the interaction terms and show that such negative correlations are significantly affected by the treatment, as suggested by the coefficients of the interaction term $\hat{SO}_i(u) \cdot T_i$. The negative correlation remains significant in the high load treatment (in which the experimental condition is identical to that in BCT) but barely exists in the low load treatment. The results suggest that time perception and time discounting are negatively associated, as observed in the literature including BCT. However, such associations might be vulnerable: They could be weakened by a change in the cognitive load.[25]

Result 2a: *While in general the perceived time length and time discounting are negatively correlated, the negative correlation is significantly weakened by lowering the cognitive load.*

To consider the second candidate, we replace $\hat{SO}_i(u)$ with the discrete variable $\hat{OVER}_i(u)$. Columns (4) and (9) show that the negative correlations between time overestimation and time discounting are substantial. Moreover, Columns (5) and (10) suggest that the treatment exerts no significant influence on the correlations. The correlations are in general significant in both treatments although weaker in the low load treatment, as indicated by the coefficients of $\hat{OVER}_i(u)$, representing the impact of time overestimation in the low load treatment, and the coefficients of $\hat{OVER}_i(u) + \hat{OVER}_i(u) \cdot T_i$, representing the impact in the high load treatment. This implies that an upward (downward) bias in time perception would consistently lead to higher (lower) discounts of the future in subjects' intertemporal choices.

Result 2b: *Subjects who overestimate time discount the future significantly more than those who underestimate time.*

The treatment difference notwithstanding, the evidence supports there being a significant association between time perception and time discounting. Therefore the treatment effect of cognitive load on time perception observed in Result 1 can be passed on to subjects' intertemporal decisions. We also note that the negative treatment effect indicated by the coefficients of T_i becomes positive from Columns (1) to (3) and (6) to (8), and the size of the coefficient decreases

[25] As a potential explanation, we find by regressing $\tilde{\varphi}_i$ on $\log \hat{SO}_i$ that the negative correlation between the S/O ratio and the discount function is marginally diminishing as the perceived length of time increases. The result helps explicate the unstable correlations in Columns (3) and (8) of Table 6: As subjects tend to overestimate time in the low load treatment, the negative correlation is weakened as the perceived future becomes more distant.

Table 6: Regression analyses of the discount functions

Dependent Variable: $\tilde{\varphi}_i(u)$			$u = 0.5$					$u = 1$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
T_i	-.0166	-.0194	.0167	-.0233	-.0101	-.0160	-.0187	.0118	-.0232	-.0094
	(.0162)	(.0170)	(.0147)	(.0171)	(.0126)	(.0167)	(.0173)	(.0153)	(.0175)	(.0136)
$\hat{SO}_i(u)$		-.0039	-.0007				-.0028*	-.0007		
		(.0024)	(.0027)				(.0016)	(.0010)		
$\hat{SO}_i(u) \cdot T_i$			-.0274**					-.0222**		
			(.0110)					(.0087)		
$O\hat{VER}_i(u)$				-.0403***	-.0284*				-.0404***	-.0279
				(.0141)	(.0168)				(.0146)	(.0179)
$O\hat{VER}_i(u) \cdot T_i$					-.0235					-.0247
					(.0312)					(.0322)
Linear tests ($F - statistics$ in parenthesis)										
$\hat{SO}_i(u) + \hat{SO}_i(u) \cdot T_i$			-.0281**					-.0229***		
			(6.48)					(6.94)		
$O\hat{VER}_i(u) + O\hat{VER}_i(u) \cdot T_i$					-.0519**					-.0526**
					(4.52)					(4.45)
R^2	.1089	.1141	.1405	.1318	.1338	.0930	.0977	.1228	.1145	.1165

Notes:

1) Standard deviation in parentheses. Errors are clustered at the individual level. Weighted regressions are performed so that individuals in different groups are equally weighted despite the amount of observation.

2) $\tilde{\varphi}_i(u)$ is the estimated discount function of u day(s). $T_i = 1$ for high load treatment and 0 otherwise. $\hat{SO}_i(u)$ is the extrapolated S/O ratio of u days. $O\hat{VER}_i(u)$ equals 1 if subject i overestimates u day(s) and 0 otherwise.

3) The fixed effects of different time preference elicitation methods are included.

4) The average performance of the filler tasks and the currency difference are controlled for.

5) *, **, *** denote significance at the 0.1, 0.05 and 0.01 level.

from Columns (4) to (5) and (9) to (10), providing suggestive evidence for the existence of a positive mediation effect.

Having demonstrated the substantial treatment effect on time perception and the salient correlation between time perception and time discounting, in the following, we focus on exploring the direction and the magnitude of the mediation effect. We present the results of the causal mediation analyses in Table 7, displaying the estimates of the ACMEs, the direct effects and the proportions of the total effect mediated. As discussed above, we employ the estimation strategy proposed by Imai, Keely and Tingley (2010) because their approach relaxes the no-interaction assumption and accommodates both continuous and discrete mediators.

Recall that, in general, the correlation between the perceived time length and the time discounting lacks consistency across treatments. Thus, unsurprisingly, in Columns (1) and (5), we find that the S/O ratio does not serve as a mediator when no interactive effect is considered, although consistent with Hypothesis 3, the ACMEs are positive. When the interactive effects are accounted for, in Columns (2) and (6), we observe positive and statistically significant ACMEs in the high load treatment. Since the direct effects are negative, the mediators offset them by over 100% in the high load treatment. However, the significant pattern does not persist to the low load treatment according to Columns (3) and (7). The average mediation effects across treatments are significant due to the considerable size of the effect found in the high load treatment, suggesting that on average, the perceived time length mediates a nontrivial part of the impact of the cognitive load treatment on time discounting behavior.

As no treatment difference is observed in the correlation between time overestimation and time discounting according to Table 6, we confine the mediation analyses to the no-interaction condition when $O\hat{VER}_i(u)$ is regarded as the mediator variable. According to Columns (4) and (8), the ACMEs are significant and positive as we conjectured, and they offset approximately one-third of the negative direct effect. The patterns in our data suggest that a notable share of the impact of the cognitive load treatment on the intertemporal choices is channeled through the over-/underestimation of time.

Result 3: *a) A shorter time perception induced by high cognitive load increases the observed patience revealed by intertemporal choices. b) Time over-/underestimation induced by the cognitive load treatment reduces/increases the observed patience revealed by intertemporal choices.*

Table 7: Results of the causal mediation analyses

Mediator:	$u = 0.5$				$u = 1$			
	$\hat{SO}_i(u)$			$OV\hat{E}R_i(u)$	$\hat{SO}_i(u)$			$OV\hat{E}R_i(u)$
	No interaction	With interaction		No interaction	No interaction	With interaction		No interaction
		$T_i = 1$	$T_i = 0$			$T_i = 1$	$T_i = 0$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ACME		.0208**	.0005			.0219*	.0006	
	.0027	[.0010, .0503]	[-.0016, .0027]	.0072**	.0025	[-.0001, .0546]	[-.0014, .0030]	.0078**
	[-.0007,.0079]	.0106**		[.0001, .0170]	[-.0005, 0072]	.0113*		[.0005, .0177]
		[.0005, .0257]				[-.0000, .0274]		
Direct effect		-.0174	-.0377			-.0167	-.0380*	
	-.0184	[-.0510, .0152]	[-.0882, .0063]	-.0223	-.0176	[-.523, . 0161]	[-.0898, .0068]	-.0221
	[-.0517, .0157]	-.0275		[-.0557, .0120]	[-.0513, .0170]	-.0274		[-.0563, .0130]
		[-.0682,.0110]				[-.0691, .0109]		
Total effect	-.0156	-.0169		-.0150	-.0151	-.0161		-.0143
	[-.0479, .0167]	[-.0493, .0149]		[-.0466, .0161]	[-.0480, .0181]	[-.0514, ..0164]		[-.0466, .0175]
% of total effect mediated	-17.31%	-123.08%	-2.96%	-48%	-16.56%	-136.02%	-3.73%	-54.55%
		-62.72%				-70.19%		

Notes:

1) The algorithm proposed in Imai, Keele and Tingley (2010) is employed. Hicks and Tingley (2010)'s STATA package is implemented to perform the estimation.

2) Quasi-Bayesian Monte Carlo approximation with 1000 replications is used for significance testing.

3) Point estimators are replicated from the coefficient estimates in Table 5 and Table 6. For example, ACME in Column (1) is the product of the coefficient of $\hat{T}_i$ in Column (1) of Table 5 and the coefficient of $\hat{SO}_i(u)$ in Column (1) of Table 6. The point estimators do not equal exactly to the coefficient products when $OV\hat{E}R_i(u)$ is the mediator variable because nonlinear regressions are involved in such cases.

4) 95% confidence intervals are presented in brackets.

5) *, **, *** denote significance at the 0.1, 0.05 and 0.01 level.

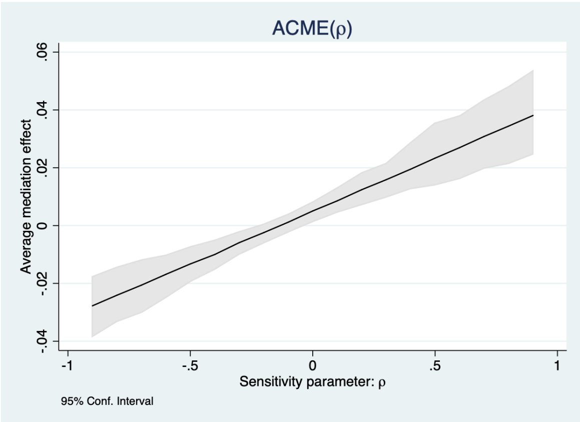
Result 3 elucidates the causal mechanisms between time perception and time discounting: Subjects with a shorter perceived time or downward bias in time perception caused by the high cognitive load are more prone to delay a reward, manifesting signs of severer impatience. This supports our conjecture that the time perception is at least partly responsible for the intertemporal decisions made in time preference elicitation tasks. If there is bias in the perception, the reported choices (partly) driven by the misperception cannot completely represent the actual preference.

While the mediation effects are positive, the direct and the total effect are negative, albeit mostly insignificant. Our results show that although higher cognitive load reduces impatience by shortening the perceived time length, as the direct effect is in the opposite direction, the total effect it produces on intertemporal choices is minor. A potential explanation for the negative direct effect is that high cognitive load reduces the willpower of a subject, which provokes a loss of self-control and thus increases the observed impatience (Loewenstein, Read and Baumeister, 2003; Ozdenoren, Salant and Silverman, 2012). This possibly cancels out the positive effect associated with time perception, resulting in the insignificant or ambiguous total effect found in our data and in the literature.

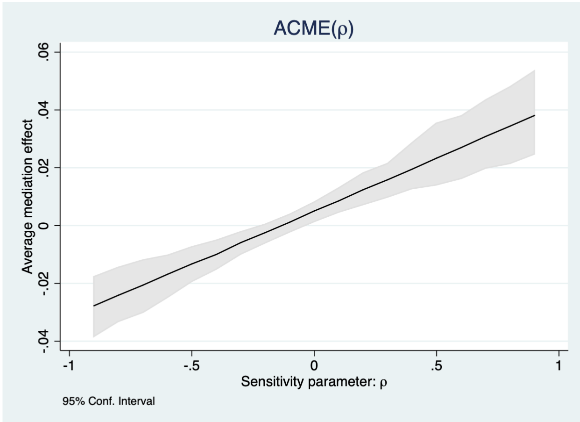
Result 4: *In general, changing cognitive load does not significantly influence the observed time discounting behaviors.*

Sensitivity analysis. Still and all, as explained in Section 5.3.1, the validity of our findings relies on the SI assumption that requires the mediators to be ignorable, conditional on the treatment and the baseline characteristics. If any unobservable factors correlate with both time perception (the mediator) and time discounting (the outcome), this assumption on the mediator is violated and the estimated ACMEs will be confounded. To understand the robustness of our results to such biases, we conduct the sensitivity analyses. The goal of the sensitivity analyses is to quantify the degree to which the key identification assumption must be violated for the original conclusion to be reversed. Under SI, the correlation between the error terms of the regression of the mediator on the treatment and of the regression of the outcomes variable on the treatment and the mediator, denoted by $\rho \equiv corr(\varepsilon_{2i}, \varepsilon_{3i})$ (or $\rho \equiv corr(\varepsilon_{2i}, \varepsilon'_{3i})$ if interactive effect is included), is zero. The question asked here is how large ρ must be to drive the mediation effect to zero, and we answer it by expressing the ACME as a function of ρ. The results from this exercise are shown in Figure 3, where we plot the ACME vs. ρ for different mediators (Figure

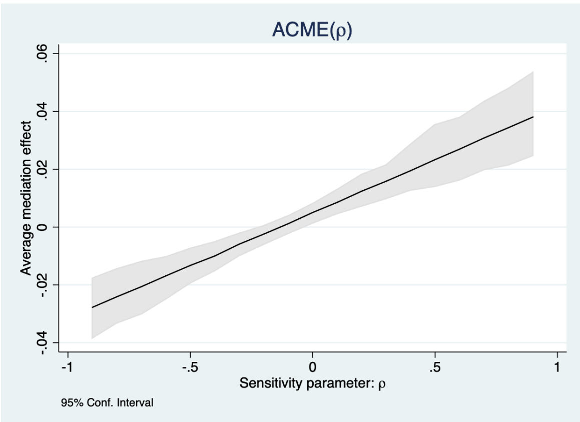
3a uses $\hat{SO}_i(0.5)$ as the mediator and Figure 3b uses $\hat{SO}_i(1)$, both with the interactive effect included; Figure 3c uses $OV\hat{E}R_i(0.5)$ as the mediator and Figure 3d uses $OV\hat{E}R_i(1)$). The curves represent the estimations of the ACME, while the shaded areas show the bootstrapped 95% confidence intervals. In Figure 3, we see that our ACME estimates appear to be somewhat sensitive to the changes in ρ, as the curves are not sufficiently flat. In addition, the values of ρ to cancel out the ACMEs are quite different between different mediator variables: Those for the S/O ratios are approximately -0.02, whereas those for time overestimation are approximately -0.2. The robustness our results regarding sensitivity should therefore be kept in mind.



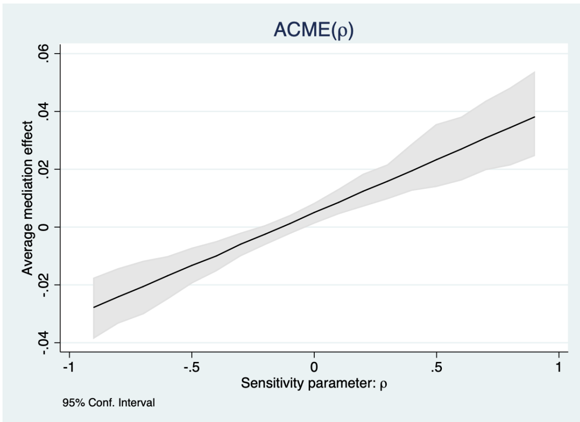
(a) S/O ratio of 12 hours with interaction effect



(b) S/O ratio of 1 day with interaction effect



(c) Time overestimation of 12 hours



(d) Time overestimation of 1 day

Figure 3: Sensitivity Analysis

Notes:
1. The curve represent the estimated ACME for the given mediator and for different values of ρ, while the shaded area represents the 95% confidence interval.
2. For Figure 3a and 3b, ρ is defined as the correlation between the error terms in Eqs. (5.3) and 5.4. For 3c and 3d, ρ is defined as the correlation between the error terms in Eqs. (5.3) and (5.5).

6 Conclusion

The essential goal of this study is to distinguish the two driving forces behind intertemporal tradeoffs: time preference and time perception. Therefore in this paper, we explore the relationship between time perception and intertemporal choices in a randomly assigned laboratory experiment, using cognitive load as an manipulation instrument to induce differences in time perception between treatments. By differentiating the cognitive load across treatments, we find that time perception is successfully manipulated by the negative impact of the treatment. Subsequently, a negative correlation of time perception and time discounting is observed: Subjects with subjectively shorter perceptions of time and those who underestimate time in general value the future reward more. The mediation analyses further imply a positive causal mediation effect of time perception on intertemporal choices: An upward (downward) bias of the time perception reduces (increases) the observed patience revealed by the discounting choices. Therefore, when the time perception is inaccurate, the time discounting estimated from the revealed intertemporal choices could deviate from the actual time preference, implying a potential identification problem for time preference. Relying solely on intertemporal decisions thus might be a beneficial yet incomplete measure to understand time preference.

Furthermore, isolating time perception from time preference may help unveil the underlying process of individuals' delay discounting, advancing our understanding of some behavioral "anomalies" and paradoxical discoveries in the time preference literature in relation to life expectancy and health-related behaviors (Banks, Blundell and Tanner, 1998; Khwaja, Silverman and Sloan, 2007; Chabris et al. 2008).

More importantly, our results shed light on the design of policy intended to improve the suboptimal time-related decisions in practice. Suboptimal time discounting has been accused of contributing to a broad range of societal problems that lead to inefficiencies such as a low savings rate or a low corporate investment rate (Loewenstein and Elster, 1992). Policies targeting the preference would be difficult to implement or ineffective since preferences are commonly considered entrenched and relatively difficult to influence. Fortunately, Brocas, Carillo and Tarrasó (2018b) find that despite their misperceptions, people have correct self-assessments of their perceptual biases. Therefore, with time perception being one of the driving forces of intertemporal decisions, policy makers may consider providing informational interventions during the intertemporal decision making process that aid the formation of a more objective time perception. While Alice in Section 3 acts as if she is more impatient than Bob, this is not necessarily true. If systematic estimation of Alice's time perception confirms that her overestimation of time is behind her "impatient" choices, then proper policies targeting the biased time perception may be effective in improving her consumption, saving and other time-related decisions. While we take an initial step in identifying the gap between people's time preference and intertemporal decisions, we leave the ambitious project of developing creative methods to estimate time preference and reassessing time-related decisions in practice for future research.

References

[1] Alonso, R., Brocas, I. and Carrillo, J. D. (2013). Resource allocation in the brain. *Review of Economic Studies,* 81(2), 501–534.

[2] Andreoni, J., Kuhn, M. A. and Sprenger, C. (2015). Measuring time preferences: A comparison of experimental methods. *Journal of Economic Behavior and Organization,* 116, 451–464.

[3] Andreoni, J. and Sprenger, C. (2012a). Estimating time preferences from convex budgets. *American Economic Review,* 102(7), 3333–3356.

[4] Andreoni, J. and Sprenger, C. (2012b). Risk preferences are not time preferences. *American Economic Review,* 102(7), 3357–3376.

[5] Attema, A. E., Bleichrodt, H., Gao, Y., Huang, Z. and Wakker, P. P. (2016). Measuring discounting without measuring utility. *American Economic Review,* 106(6), 1476–1494.

[6] Attema, A. E., Bleichrodt, H., Rohde, K. I. M. and Wakker, P. P. (2010). Time-Tradeoff sequences for analyzing discounting and time inconsistency. *Management Science,* 56(11), 2015–2030.

[7] Banks, J., Blundell, R. and Tanner, S. (1998). Is there a retirement-savings puzzle? *American Economic Review,* 88(4), 769–788.

[8] Becker, G. S. and Mulligan, C. B. (1997). The endogenous determination of time preference. *Quarterly Journal of Economics,* 112(3), 729–758.

[9] Bénabou, R. and Tirole, J. (2002). Self-confidence and personal motivation. *Quarterly Journal of Economics,* 117(3), 871–915.

[10] Block, R. A., Hancock, P. A. and Zakay, D. (2010). How cognitive load affects duration judgments: A meta-analytic review. *Acta Psychologica,* 134(3), 330–343.

[11] Bonato, M. and Umiltà. C. (2014). Heterogeneous timescales are spatially represented. *Frontiers in Psychology,* 5.

[12] Bradford, W. D., Dolan, P. and Galizzi, M. M. (2019). Looking ahead: Subjective time perception and individual discounting. *Journal of Risk and Uncertainty,* 58(1), 43–69.

[13] Brocas, I., Carrillo, J. D. and Tarrasó. (2018a). Self-awareness of biases in time perception. *Journal of Economic Behavior and Organization,* 148, 1–19.

[14] Brocas, I., Carrillo, J. D. and Tarrasó. (2018b). How long is a minute? *Games and Economic Behavior,* 111, 305–322.

[15] Chabris, C. F., Laibson, D., Morris, C. L., Schuldt, J. P. and Taubinsky, D. (2008). Individual laboratory-measured discount rates predict field behavior. *Journal of Risk and Uncertainty,* 37(2-3), 237–269.

[16] Chen, M. K. (2013). The effect of language on economic behavior: Evidence from savings rates, health behaviors, and retirement assets. *American Economic Review,* 103(2), 690–731.

[17] Chen, X., Hong, F. and Zhao, X. (2019). Concentration and variability of forecasts in artificial investment games: An online experiment on WeChat. *Experimental Economics,* 23(3), 815–847.

[18] Chew, S. H., Huang, W. and Zhao, X. (2020). Motivated false memory. *Journal of Political Economy,* 128(10), 3913–3939.

[19] Chew, S. H. and Li, K. K. (2017). Sin stock aversion in the financial market: An experimental investigation. Working Paper, available at: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2792301`.

[20] Choi, S., Kariv, S., Müller, W. and Silverman, D. (2014). Who is (more) rational? *American Economic Review,* 104(6), 1518–1550.

[21] Cohen, J., Ericson, K. M., Laibson, D. and White, J. M. (2020). Measuring time preferences. *Journal of Economic Literature,* 58(2), 299–347.

[22] Crystal, J. D. (2001). Circadian time perception. *Journal of Experimental Psychology: Animal Behavior Processes,* 27(1), 68–78.

[23] Cui, X. (2011). Hyperbolic discounting emerges from the scalar property of interval timing. *Frontiers in Integrative Neurosci,* 5.

[24] Deck, C. and Jahedi, S. (2015). The effect of cognitive load on economic decision making: A survey and new experiments. *European Economic Review,* 78, 97–119.

[25] Eagleman, D. M. (2008). Human time perception and its illusions. *Current Opinion in Neurobiology,* 18(2), 131–136.

[26] Ebert, J. E. J. and Prelec, D. (2007). The fragility of time: Time-insensitivity and valuation of the near and far future. *Management Science,* 53(9), 1423–1438.

[27] Fechner, G. T. (1860). Elements of psychophysics. *Readings in the History of Psychology,* 206–213.

[28] Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics,* 10(2), 171–178.

[29] Fontes, R., Ribeiro, J., Gupta, D. S., Machado, D., Lopes-Júnior, F., Magalhães, F., Bastos, V. H., Rocha, K., Marinho, V., Lima, G., Velasques, B., Ribeiro, P., Orsini, M., Pessoa, B., Leite, M. A., and Teixeira, S. (2016). Time perception mechanisms at central nervous system. *Neurology International,* 8(1), 5939.

[30] Francis-Smythe, J. A. and Robertson, I. T. (1999). On the relationship between time management and time estimation. *British Journal of Psychology,* 90(3), 333–347.

[31] Franco-Watkins, A. M., Rickard, T. C. and Pashler, H. (2010). Taxing executive processes does not necessarily increase impulsive decision making. *Experimental Psychology,* 57(3), 193–201.

[32] Frederick, S., Loewenstein, G. and O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature,* 40(2), 351–401.

[33] Galor, O. and Özak, Ö (2016). The agricultural origins of time preference. *American Economic Review,* 106 (10), 3064-3103.

[34] Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception and Psychophysics,* 72(3), 561–582.

[35] Gui, Z., Huang, Y. and Zhao, X. (2021). Whom to educate? Financial literacy and investor awareness. *China Economic Review,* 67, 101608.

[36] Han, R. and Takahashi, T. (2012). Psychophysics of time perception and valuation in temporal discounting of gain and loss. *Physica A: Statistical Mechanics and Its Applications,* 391(24), 6568–6576.

[37] Hicks, R. and Tingley, D. (2011). Causal mediation analysis. *Stata Journal,* 11(4), 605–619.

[38] Hong, F., Huang, W. and Zhao, X. (2018). Sunk cost as a self-management device. *Management Science,* 65(5).

[39] Ifcher, J. and Zarghamee, H. (2011). Happiness and time preference: The effect of positive affect in a random-assignment experiment. *American Economic Review,* 101(7), 3109–3129.

[40] Imai, K., Keele, L. and Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods,* 15(4), 309–334.

[41] Imai, K., Keele, L. and Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science,* 25(1), 51–71.

[42] Judd, C. M. and Kenny, D. A. (1981). Process analysis. *Evaluation Review,* 5(5), 602–619.

[43] Khwaja, A., Silverman, D. and Sloan, F. (2007). Time preference, time discounting, and smoking decisions. *Journal of Health Economics,* 26(5), 927–949.

[44] Kim, B. K. and Zauberman, G. (2009). Perception of anticipatory time in temporal discounting. *Journal of Neuroscience, Psychology, and Economics,* 2(2), 91–101.

[45] Kraemer, C. H., Kiernan, M., Essex, M. and Kupfer, D. J. (2008). How and why criteria defining moderators and mediators differ between the Baron and Kenny and MacArthur approaches. *Health Psychology,* 27(2, Suppl), S101–S108.

[46] Kraemer, W. J., Ratamess, N. A. and French, D. N. (2002). Resistance training for health and performance. *Current Sports Medicine Reports,* 1(3), 165–171.

[47] Loewenstein, G. and Elster, J. (1992). *Choice over time.* Russell Sage Foundation.

[48] Loewenstein, G. and Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *Quarterly Journal of Economics,* 107(2), 573–597.

[49] Loewenstein, G., Read, D. and Baumeister, R. (2004). Time and decision: Economic and psychological perspectives on intertemporal choice. *Journal of Socio-Economics,* 33(4), 519–521.

[50] Manski, C. F. (2007). *Identification for prediction and decision.* Cambridge, MA: Harvard University Press

[51] Meier, S. and Sprenger, C. D. (2013). Discounting financial literacy: Time preferences and participation in financial education programs. *Journal of Economic Behavior and Organization,* 95, 159–174.

[52] Miao, B. and Zhong, S. (2015). Comment on "risk preferences are not time preferences": Separating risk and time preference. *American Economic Review,* 105(7), 2272–2286.

[53] Mioni, G., Grondin, S., Bardi, L. and Stablum, F. (2020). Understanding time perception through non-invasive brain stimulation techniques: A review of studies. *Behavioural Brain Research,* 377, 112232.

[54] Moskowitz, G. B. (2005). On Schemas and Cognitive Misers. in *Social cognition : understanding self and others.* Guilford Press.

[55] Ozdenoren, E., Salant, S. W. and Silverman, D. (2010). Willpower and the optimal control of visceral urges. Working Paper, available at:`https://doi.org/10.2139/ssrn.1635350`

[56] Ray, D. and Bossaerts, P. (2011). Positive temporal dependence of the biological clock implies hyperbolic discounting. *Frontiers in Neuroscience,* 5.

[57] Read, D. (2001). Is time-discounting hyperbolic or subadditive? *Journal of Risk and Uncertainty,* 23(1), 5–32.

[58] Robson, A. J. and Szentes, B. (2008). Evolution of time preference by natural selection: Comment. *American Economic Review,* 98(3), 1178–1188.

[59] Robson, A. J., Szentes, B. and Iantchev, E. (2012). The evolutionary basis of time preference: Intergenerational transfers and sex. *American Economic Journal: Microeconomics,* 4(4), 172–201.

[60] Roelofsma, P. H. M. P. (1996). Modelling intertemporal choices: An anomaly approach. *Acta Psychologica,* 93(1-3), 5–22.

[61] Savage, L. J. (2006). *The foundations of statistics.* Dover Publications.

[62] Stevens, S. S. (1957). On the psychophysical law. *Psychological Review,* 64(3), 153–181.

[63] Takahashi, T. (2005). Loss of self-control in intertemporal choice may be attributable to logarithmic time-perception. *Medical Hypotheses,* 65(4), 691–693.

[64] Takahashi, T. (2009). Theoretical frameworks for neuroeconomics of intertemporal choice. *Journal of Neuroscience, Psychology, and Economics,* 2(2), 75–90.

[65] Tanaka, T., Camerer, C. F. and Nguyen, Q. (2010). Risk and time preferences: Linking experimental and household survey data from Vietnam. *American Economic Review,* 100(1), 557–571.

[66] Thaler, R. H. and Shefrin, H. M. (1981). An economic theory of self-control. *Journal of Political Economy,* 89(2), 392–406.

[67] Thomas, E. A. C. and Weaver, W. B. (1975). Cognitive processing and time perception. *Perception and Psychophysics,* 17(4), 363–367.

[68] Tse, P. U., Intriligator, J., Rivest, J. and Cavanagh, P. (2004). Attention and the subjective expansion of time. *Perception and Psychophysics,* 66(7), 1171–1189.

[69] Waldum, E. R. and McDaniel, M. A. (2016). Why are you late? Investigating the role of time management in time-based prospective memory. *Journal of Experimental Psychology: General,* 145(8), 1049–1061.

[70] Weber, E. U. and Milliman, R. A. (1997). Perceived risk attitudes: Relating risk perception to risky choice. *Management Science,* 43(2), 123–144.

[71] Wittmann, M. and Paulus, M. P. (2008). Decision making, impulsivity and time perception. *Trends in Cognitive Sciences,* 12(1), 7–12.

[72] Yu, C. W., Zhang, Y. J. and Zuo, S. X. (2021). Multiple switching and data quality in the multiple price list. *Review of Economics and Statistics,* 103 (1), 136–150.

[73] Zauberman, G., Kim, B. K., Malkoc, S. A. and Bettman, J. R. (2009). Discounting time and time discounting: Subjective time perception and intertemporal preferences. *Journal of Marketing Research,* 46(4), 543–556.

Appendix of "How time flies!"

A Instructions

Welcome to the experiment. This instruction helps you understand the experiment. The final payment for your participation will depend on how well you perform in the experiment; thus, please read the instructions carefully.

There will be two parts in the experiment. Part I involves your estimating a period of time, as instructed below. Part II asks you to choose between two options in a series of questions according to your preference, the instruction for which will be provided to you after you finish Part I. In both parts, your payment will depend on your corresponding decisions.

Your Tasks

In general, you have two parts of tasks to complete simultaneously: 1) the time estimation task, and 2) a series of filler tasks.

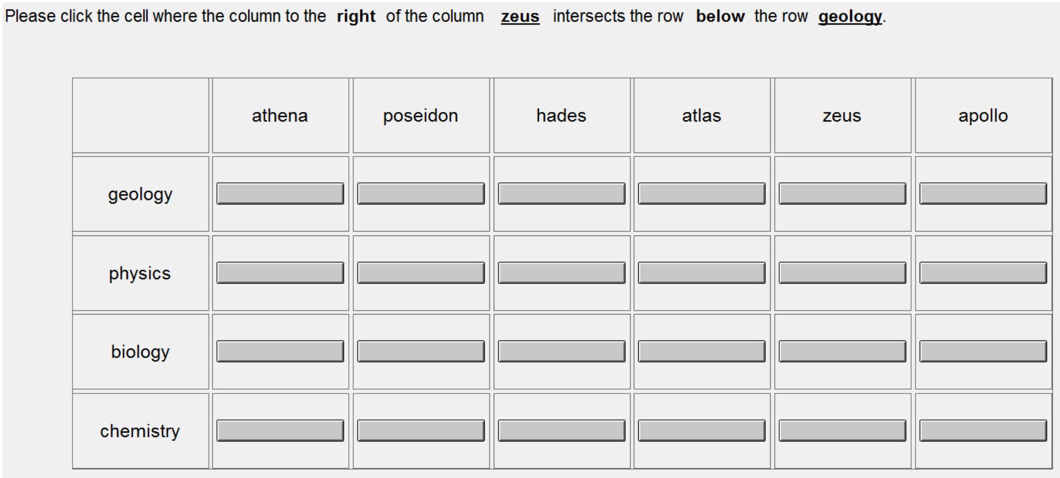
I. Time estimation task

In each round of the experiment, you will be asked to produce a time interval. You will be informed of the interval t seconds at the beginning of each round. The value of t will be changed in each round.

You will produce the time interval by controlling the stopwatch using the START and the STOP buttons on the screen: When you are ready to estimate the time interval, you may press the START button to activate the stopwatch; then there will be a STOP button at the right bottom of the screen. Press the STOP button as soon as you believe t seconds are up, and the stopwatch will be deactivated immediately. The duration between your click of the STOP button and of the START button will be your estimation of the time interval t seconds.

II. Filler tasks

You will be given a series of filler tasks while estimating the time interval.



Here is an example of the filler task. The names of the rows and columns as well as the phrasing of which cell to click on will change from task to task. By clicking the button in each cell, your answer will be recorded. If you want to change your answer, simply click on the new cell you want to change to, and the recorded answer will be replaced by the new one.

The filler tasks will be shown one on each page. At the beginning of each filler task, the computer will randomly determine a time limit p seconds, which you will not be informed of. You will need to complete each filler task within p seconds, and failure to complete it will be counted as an incorrect answer. You will not be able to skip a task before p seconds, regardless of your completion of the task. After p seconds are up, you will be automatically directed to the next task page. The value of p is independently and randomly determined for each filler task, which means that the time limits for you to complete different tasks can be different.

Rundown of the Experiment

1. At the beginning of each round, you will be informed of the time interval (t seconds) to produce in this round. You may press the START button to activate the stopwatch when you are ready.

2. The first filler task will be shown right after you press the START button. You have p seconds to complete the task, for which the value of p is randomly determined the moment you enter this task page. The next filler task will automatically be shown to you after p seconds, and value of p will be redetermined when you enter the next task page.

3. There will be a STOP button on the right bottom corner of the screen during the filler tasks. You will need to press the STOP button to deactivate the stopwatch, whenever you feel that t seconds have passed since the START button is pressed. Then the duration between your

click of the STOP button and of the START button will be your estimation.

4. When the STOP button is clicked, this round of the study is finished, and the ongoing filler task will not be counted when calculating your final payment, even if you have clicked on the cell(s). After clicking the STOP button, you will be directed to the next round with a new interval t seconds to produce.

5. There will be 1 practice round at the beginning, followed by 10 official rounds. The practice round will be the same as the official rounds, except that it has no chance to be selected to calculate your payment.

6. The history of your estimations of time and your results for the filler tasks will be provided at the end of the whole experiment only. One of the 10 official rounds will then be randomly selected to calculate your final payment, which will also be shown to you on the final page.

7. Part I and Part II proceed independently for each participant. Upon the completion of Part II, the final page of the whole experiment will determine your payment. You will be asked to sign your name to acknowledge the receipt of your cash payment. You are then free to leave after collecting the payment.

Your Payment

Your payment includes two components: 1) Payment for Part I; and 2) Payment for Part II.

Regarding your payment for Part I, we will randomly select 1 out of 10 official rounds to calculate it, with each round being equally likely to be selected, therefore it is your best interest to take every official round seriously.

Your final payment for Part I depends on 1) your accuracy on the time estimation task and 2) your performance on the filler tasks in the selected round. For the selected round, you earn money only if at least 80% of the filler tasks in this round are correctly answered. Then you will be paid HK$120 if the estimation is within ±5% of the real length of the interval t seconds, HK$60 if it is within ±10% and HK$30 if it is within ±20%. If less than 80% of the filler tasks are correctly answered, or your estimation is strictly out of ±20% of the real length of the interval t seconds, you will not earn anything in this case.

Formally, following is the formula used to calculate your payoff (in HK$):

$$\text{Your payoff} = I_{filler} \times A_{est} \times 120$$

where

- est = the duration (in seconds) between your click of the STOP button and of the START button in the selected round;
- $I_{filler} = \begin{cases} 1 & \text{if 80\% or more filler tasks in the selected round are correctly answered} \\ 0 & \text{if less than 80\% filler tasks in the selected round are correctly answered} \end{cases}$
- $A_{est} = \begin{cases} 1 & \text{if } \frac{|est-t|}{t} \leq 5\% \\ 0.5 & \text{if } 5\% < \frac{|est-t|}{t} \leq 10\% \\ 0.25 & \text{if } 10\% < \frac{|est-t|}{t} \leq 20\% \\ 0 & \text{if } \frac{|est-t|}{t} > 20\% \end{cases}$.

Comprehension Quiz

True or False:

1. The number of the filler tasks that I complete in a round depends on the time limit for each of the tasks determined by the computer, and my estimation of the time interval of the round.

2. I cannot use the number of filler tasks that I have completed in this round to calculate the time that has passed.

3. When I feel that t seconds are up, I need to click the STOP button right away even if I haven't finished the ongoing filler task.

4. I can use the information provided regarding my estimation in previous rounds to improve my accuracy in later rounds.

5. I don't need to wait for others to proceed to the next round of time estimation task, or to start Part II.

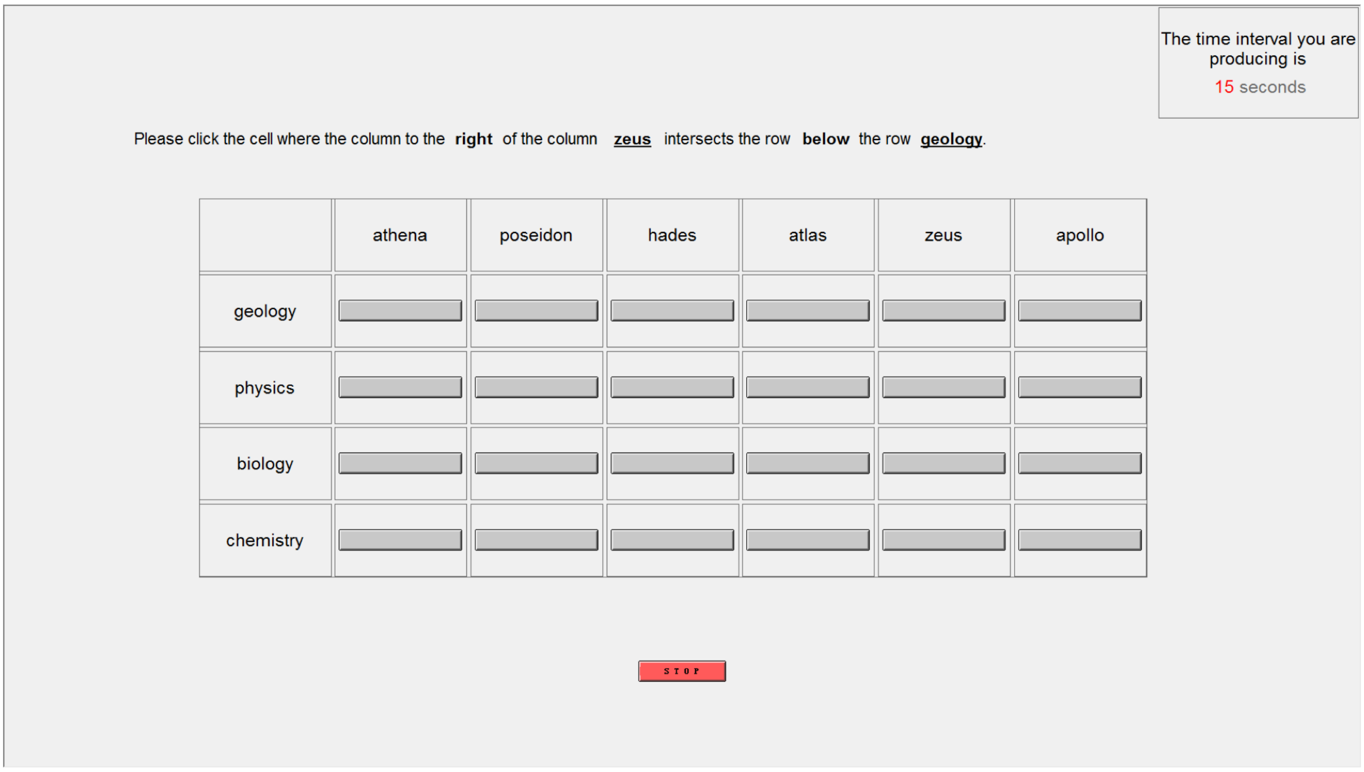
Administration

Your decisions as well as your cash payment will be kept completely confidential. From this point onwards, please turn off your mobile phone or any other electronic devices, and refrain from talking to any other participants during the experiment.

If you have any questions, please raise your hand now. We will answer questions individually. If there are no questions, we will begin with the experiment.

B Screenshots of the z-tree program

B.1 Screenshots of the TP task

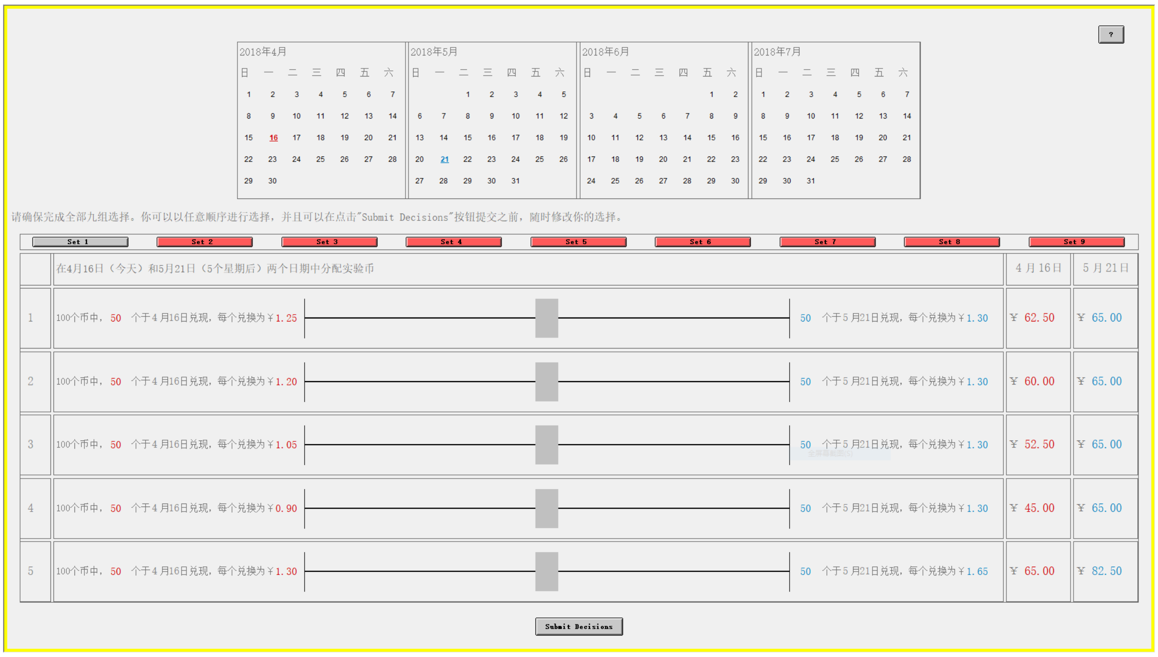


B.2 Screenshots of the TD task

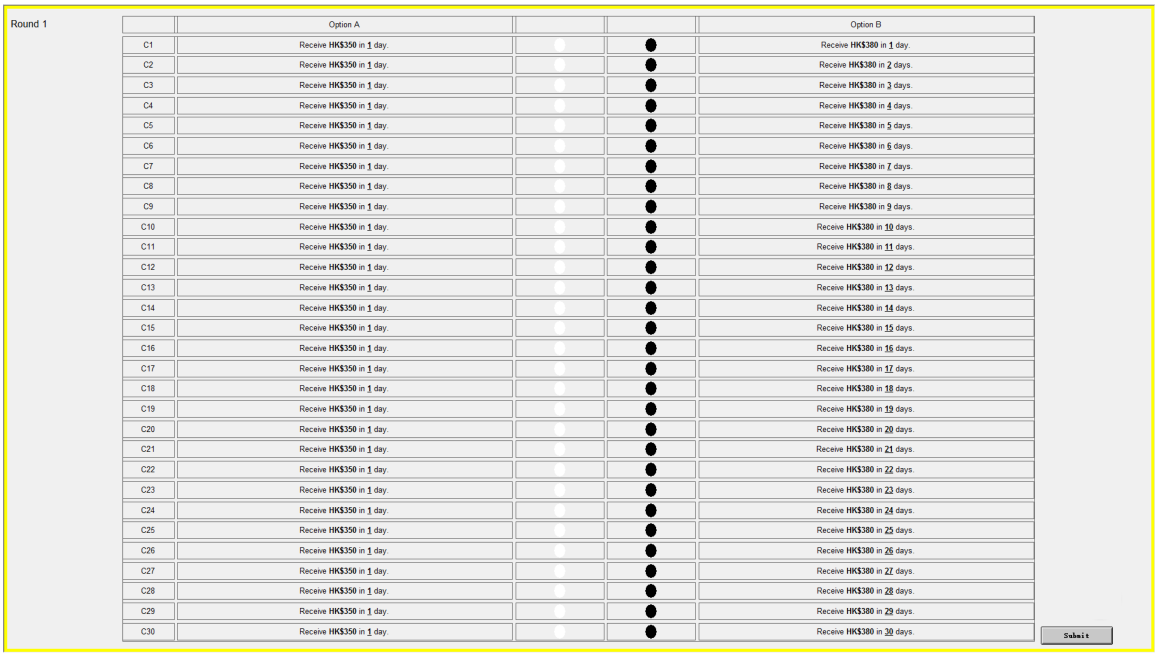
B.2.1 Multiple price list task

	A	今天	B	31 天后
1	❑	￥100	❑	￥100
2	❑	￥100	❑	￥103
3	❑	￥100	❑	￥106
4	❑	￥100	❑	￥109
5	❑	￥100	❑	￥112
6	❑	￥100	❑	￥115
7	❑	￥100	❑	￥118
8	❑	￥100	❑	￥121
9	❑	￥100	❑	￥124
10	❑	￥100	❑	￥127
11	❑	￥100	❑	￥130
12	❑	￥100	❑	￥133
13	❑	￥100	❑	￥136
14	❑	￥100	❑	￥139
15	❑	￥100	❑	￥142
	A	**90 天后**	**B**	**121 天后**
16	❑	￥100	❑	￥100
17	❑	￥100	❑	￥103
18	❑	￥100	❑	￥106
19	❑	￥100	❑	￥109
20	❑	￥100	❑	￥112
21	❑	￥100	❑	￥115
22	❑	￥100	❑	￥118
23	❑	￥100	❑	￥121
24	❑	￥100	❑	￥124
25	❑	￥100	❑	￥127
26	❑	￥100	❑	￥130
27	❑	￥100	❑	￥133
28	❑	￥100	❑	￥136
29	❑	￥100	❑	￥139
30	❑	￥100	❑	￥142

B.2.2 Convex time budget task



B.2.3 Time-tradeoff sequence task



C Details of time discounting estimations

C.1 Multiple price list method

The time discount function are elicited from comparisons between their tradeoffs in a proximal task (today versus 31 days later) and a distal task (90 days versus 121 days later). The ratio of the earlier payment to the future payment (defined as the average value of the amount in Option B at the switch point, where subjects' choices switch from Option A to Option B, and the amount in Option B right before the switch point) in distal task delivers normal discount factor δ, while the ratio of two values in proximal task delivers the near term discount factor $\beta\delta$. The ratio of two discount factors elicits the present bias parameter β.

C.2 Convex time budget method

The TD part in the CTB group asked subject i to chooses at time u to allocate a budget d between consumption x_u at time u and consumption x_{u+k} at time $u+k$. Then

$$(1+r)x_u + x_{u+k} = d \tag{C.1}$$

where $(1+r)$ is the gross interest rate. Then the discounted utility at time 0 is:

$$\tilde{DU}(u : x_u, u+k : x_{u+k}) = \tilde{\varphi}(u)U(x_u) + \tilde{\varphi}(u+k)U(x_{u+k}) \tag{C.2}$$

where $U(x_u)$ and $U(x_{u+k})$ are the instant utility at time u and $u+k$ respectively. The instant utility is assumed to be a CRRA utility, so:

$$U(x_u) = \frac{1}{\alpha}x_u^{\alpha}$$

where $\alpha > 0$ is the curvature parameter.

Following the practice in AS, we focus our attention to the quasi-hyperbolic discount functions with the following form:

$$\tilde{\varphi}(u) = \begin{cases} \beta\delta^u & u > 0 \\ 1 & u = 0 \end{cases}$$

where β is the time inconsistency parameter and $\delta \in (0,1)$ is the one period discount. By maximizing Eq. (C.2) subject to Eq. (C.1), the optimal consumption is:

$$x_u = \begin{cases} \frac{(\delta^k(1+r))^{(\frac{1}{\alpha-1})}d}{(1+(1+r)(\delta^k(1+r))^{(\frac{1}{\alpha-1})}} & u > 0 \\ \frac{(\beta\delta^k(1+r))^{(\frac{1}{\alpha-1})}d}{(1+(1+r)(\beta\delta^k(1+r))^{(\frac{1}{\alpha-1})}} & u = 0 \end{cases} \tag{C.3}$$

With Eq. (C.3), we can now fit the model to the 45 data points for each individual by the NLS approach to estimate the parameter α, β and δ.

C.3 Time-tradeoff sequence method

In the TTO group, subject i were asked to choose switch points that implied the indifference in Eq. (4.1) in the TD part. For the parameters in consecutive steps, the questions in one step were adapted to the selection in the previous step. For example, suppose in step 1, the switching point was reported to be $\tilde{u}_1$, then the indifference point elicited would be $u_1 = \tilde{u}_1 - \frac{1}{2}$, and the sooner time point in step 2 was then rounded to $\tilde{u}_1$ because we only presented integer unit of days to subjects. This would cause the indifference point to be overestimated by $\frac{1}{2} + \frac{1}{2} = 1$, so we corrected the underestimation by letting $u_2 = \tilde{u}_2 - 1$, which was an integer thus would be used as the sooner time point in step 3. Therefore, we corrected for roundings by calculating the indifference point such that $u_n = \tilde{u}_n - \frac{1}{2}$ if $n = 1, 3, 5$ and $u_n = \tilde{u}_n - 1$ if $n = 2, 4$. Then following ABRW, in order to estimate the discount function $\tilde{\varphi}(u)$ such that

$$\tilde{DU}(u_1 : x_1, \ldots, u_m : x_m) = \sum_{i=1}^{m} \tilde{\varphi}(u_i)U(x_i)$$

for a TTO sequence, we have:

$$\frac{\tilde{\varphi}(u_0)}{\tilde{\varphi}(u_1)} = \frac{\tilde{\varphi}(u_1)}{\tilde{\varphi}(u_2)} = \ldots = \frac{\tilde{\varphi}(u_{n-1})}{\tilde{\varphi}(u_n)} = \frac{U(\gamma)}{U(\kappa)}$$

Hence,

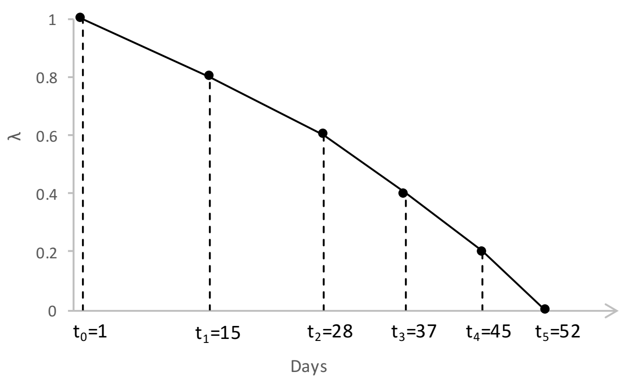
$$\ln(\tilde{\varphi}(u_0)) - \ln(\tilde{\varphi}(u_1)) = \ln(\tilde{\varphi}(u_1)) - \ln(\tilde{\varphi}(u_2)) = \ldots = \ln(\tilde{\varphi}(u_{n-1})) - \ln(\tilde{\varphi}(u_n))$$

that is, a TTO sequence is equally spaced in $\ln(\tilde{\varphi})$ units. We then are able to construct a normalized TTO curve as follows:

$$\lambda(u) = \frac{\ln(\tilde{\varphi}(u)) - l}{q}, \text{ with } l = \ln(\tilde{\varphi}(u_n)) \text{ and } q = \ln(\tilde{\varphi}(u_0)) - \ln(\tilde{\varphi}(u_n))$$

By this construction, we have $\lambda = 1$ at u_0 and 0 at u_n, with $\lambda(u_{j-1}) - \lambda(u_j) = 1/n$. Here in our study, $n = 5$ throughout as we elicit subjects' time preference up to 5 steps. Figure C.1 depicts the TTO curve constructed using the median of λ for Sequence IV observed in the low load treatment.

Figure C.1: TTO Curve of Sequence IV in the Low Load Treatment



We calculate q by assuming linear utility[26], so

$$q = \frac{\ln(U(\gamma)/U(\kappa))}{\lambda(t_0) - \lambda(t_1)} = n \cdot \ln(U(\gamma)/U(\kappa))$$

where n is the total steps in the elicitation task. Then, the discounting function is calculated as:

$$\tilde{\varphi}(t) = e^{q \cdot \lambda(t) + l}$$

where l can be chosen arbitrarily without affecting the preference.

[26]While linear utility is assumed mostly in the literature, we also try the alternative concave utility assumption by assuming the *constant relative risk-aversion* (CRRA) parameter to be 0.88, the average value of $\hat{\alpha}_i$ estimated from the CTB group. The main results are generally the same.

D Additional analyses

D.1 Discounting against subjective time.

We have demonstrated that individual i's time perception can be well summarized by $p_i(\tau) = a_i\tau^{b_i}$ in Section 5.1 and individual i's discount function by $\tilde{\varphi}_i(u)$ in Section 5.2. In this section, we will perform proper estimations of time discounting while explicitly accounting for the individual heterogeneity in time perception. To this end, we map perceived time onto discount functions by explicitly estimating the discount functions using individual subjective, rather than objective, time such that $\varphi_i(\tau) = \tilde{\varphi}_i(p_i(\tau))$. More precisely, for subjects in the MPL and the CTB groups, the discount function is given by:

$$\varphi_i(\tau) = \beta_i' \tilde{\delta}_i'^{a_i(\tau)^{b_i}} \tag{D.1}$$

where the time unit for τ is a second, meaning that $\tilde{\delta}_i'$ can be interpreted as the discount per second. By setting $\delta_i' \equiv \tilde{\delta}_i'^{z^{b_i}}$ where $z = \frac{\tau}{u}$, we may then rewrite Eq. (D.1) as follows:

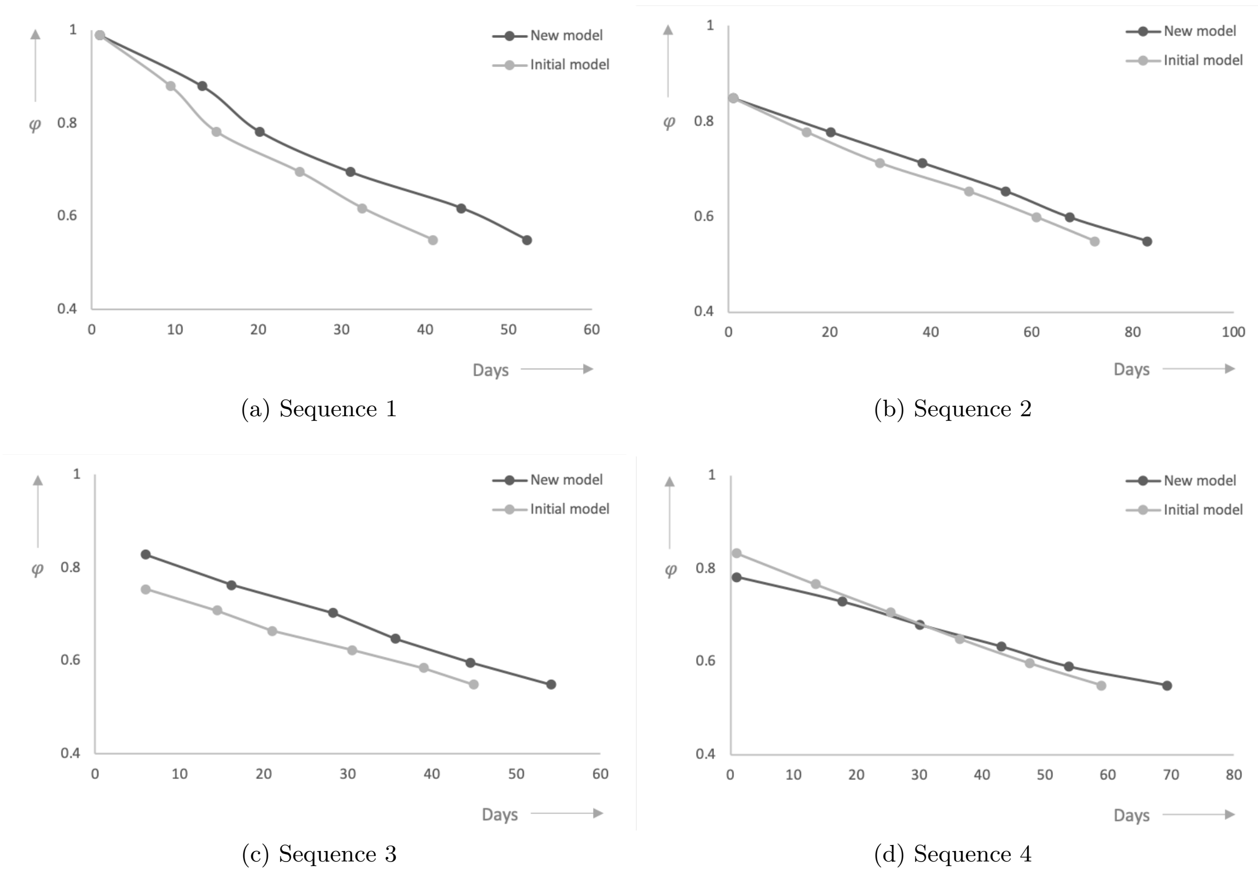
$$\varphi_i(u) = \beta_i' \delta_i'^{a_i(u)^{b_i}} \tag{D.2}$$

using one day as the time unit for u.

Table D.1: Discounting, Present Bias, and Curvature Parameter Estimates

	Mean		Median		5th Percentile		95th Percentile	
	MPL	CTB	MPL	CTB	MPL	CTB	MPL	CTB
Daily discount factor: $\hat{\delta}_i'$	0.997	0.998	0.998	0.999	0.994	0.992	1.000	1.000
Present bias: $\hat{\beta}_i'$	0.977	0.893	1.000	0.945	0.888	0.692	1.084	1.071
CRRA Curvature: $\hat{\alpha}_i'$		0.879		0.948		0.381		1.000

For the estimates to be comparable with those in Table 3, instead of estimating all 5 parameters, we import the time perception parameters $\hat{a}_i$ and $\hat{b}_i$ estimated in Section 5.1 and estimate the remaining three parameters ($\hat{\beta}_i'$, $\hat{\delta}_i'$ $\hat{\alpha}_i'$) as we did in Section 5.2. Table D.1 summarizes the distributions of the new estimators, which are extremely similar to those in Table 3. We also find that $\hat{\beta}_i$ and $\hat{\beta}_i'$ are positively correlated (pairwise correlation coefficient = 0.9997, $p-value$ = 0.0000 with 103 observations) as are $\hat{\alpha}_i$ and $\hat{\alpha}_i'$ (pairwise correlation coefficient = 0.9990, $p-value$ = 0.0000 with 56 observations). $\hat{\delta}_i$ and $\hat{\delta}_i'$ are also tightly correlated, but with slightly lesser extent (pairwise correlation coefficient = 0.8751, $p-value$ = 0.0000 with 103 observations). This is not surprising, as $\hat{\delta}'_i$, by definition, depends on the perception parameter $\hat{b}_i$ is now applied to subjective time and therefore adjusted for the individuals' perception biases. Since such biases have less effect on people's feelings about the present and future, their effect on the present-bias tendency represented by $\hat{\beta}'_i$ should be inconspicuous.[27] More interesting, when we examine the structural estimations in the CTB group, both the AIC and the BIC are lower for the new model than the initial model ($average\ AIC = 281.76$ for the new model and $average\ AIC = 293.12$ for the initial model; $average\ BIC = 287.11$ for the new model and $average\ BIC = 298.48$ for the initial model), suggesting that the new model performs better than the initial model. This result supports the notion that the information regarding time perception is useful for understanding people's intertemporal decisions.



(a) Sequence 1

(b) Sequence 2

(c) Sequence 3

(d) Sequence 4

Figure D.2: Discount Curves for the TTO group: Comparing the New and Initial Models
Note: The x-axis values are the median delays over the 118 subjects for each of the 5 elicited steps in each sequence. The y-axis values are the discount functions for each step in each sequence same across all subjects.

As intertemporal discounting observed in the TTO group cannot be compared between subjective time and objective time using parameters, we provide a figure as in Section 5.2 to roughly illustrate the discounting pattern. Figure D.2 shows that because subjects tend

[27]In fact, the derivation of $\hat{\beta}'_i$ for subjects in the MPL group does not rely on the time span at all.

to overestimate time, when subjective time is accounted for, the discount curves for the new models (i.e., the curves for $\varphi_i(\cdot)$) are generally above those for the initial model (i.e., the curves for $\tilde{\varphi}_i(\cdot)$), implying that people are actually more patient than they appear to be. However, as the curves only display the median values over all subjects in the TTO group, the large heterogeneity in time perception across individuals is always worth attention.

In summary, the main information from this exercise is that models reflecting the impact of time perception on intertemporal valuations are plausible. While the standard model based on objective duration is a reasonable approximation, the new model provides more complete information about an individual's time preference than the initial model.

D.2 Robustness to a stricter definition of time overestimation.

In the discussion above, we investigated whether and how time overestimation bias affect subjects' intertemporal decisions by dividing our sample between those who overestimate time and those who underestimate it according to the extrapolated time estimates. A possible issue regarding such a classification is that for the extrapolated time estimates that are in the neighborhood of the objective time points, a small measurement error or disturbance in the estimation would lead to the contrary mediator, as the assignment of over- or underestimation of time is binary but the time estimates are continuous.

To avoid this potential problem, we use a stricter definition and consider only the observations for which the absolute perceptual bias is above 20% (i.e., time overestimation is now defined such that $O\hat{VER}_i(u) = 1$ if $\hat{SO}_i(u) - 1 > 0.2$, and $O\hat{VER}_i(u) = 0$ if $\hat{SO}_i(u) - 1 < -0.2$); 58 observations are dropped as a result. The results in Table D.2 that reports the statistical analyses for the subsample conform with those in the previous subsection for the full sample: Time overestimation induced by low cognitive load decreases in the valuation of the delayed reward. Despite a modest reduction in significance, the correlations among the treatment, the mediator and the outcome variables as well as the positive mediation effects are preserved. The results suggest that the substantive conclusion drawn from the mediation analyses using $O\hat{VER}_i(u)$ as the mediator variable is robust to a stricter definition of time over-/underestimation.

Table D.2: Analyses on the Subsample with a Stricter Definition of Time Overestimation

	$u = 0.5$			$u = 1$
Panel A. Regressions of the mediator variable: $O\hat{VER}_i(u)$				
T_i		-.9133**		-1.1626**
		(.4157)		(.5628)
R^2		0.0275		0.0236
Panel B. Regressions of the outcome variables: $\tilde{\varphi}_i(u)$				
T_i	-.0312*	-.0353*	-.0290	-.0326*
	(.0188)	(.0193)	(.0185)	(.0189)
$O\hat{VER}_i(u)$		-.0046**		-.0032**
		(.0023)		(.0015)
R^2	0.1182	0.1279	0.1016	0.1103
Panel C. Results of the mediation analyses				
ACME		.0040*		.0036*
		[-.0005, .0105]		[-.0004, .0094]
Direct effect		-.0341		-.0315
		[-.0718, .0046]		[-.0683, .0064]
Total effect		-.0301		-.0279
		[-.0674, .0071]		[-.0645, .0090]
% of total effect mediated		-13.29%		-12.90%

Notes:

1) $O\hat{VER}_i(u)$ equals 1 (0) if subject i overestimates (underestimate) u day(s) by 20% or above. $T_i = 1$ for high load treatment and 0 otherwise. $\tilde{\varphi}_i(u)$ is the estimated discount function of u day(s).

2) For Panel A and B, coefficient estimates are presented with standard deviations in parentheses.

3) For Panel C, point estimators are presented with 95% confidence intervals in brackets. The algorithm proposed in Imai, Keele and Tingley (2010) is employed. Hicks and Tingley (2010)'s STATA package is implemented to perform the estimation. The point estimators do not equal exactly to the coefficient products because nonlinear regressions are involved.

4) Errors are clustered at the individual level. The fixed effects of different time preference elicitation methods are included. The average performance of the filler tasks and the currency difference are controlled for.

5) Weighted regressions are performed so that each individual is equally weighted despite the different amount of observations in different groups.

6) *, **, *** denote significance at the 0.1, 0.05 and 0.01 level.

D.3 Comparing across different time discounting elicitation methods.

We adopted three time discounting elicitation methods well-established in the literature to prevent the results from being dependent on the setting of a specific experimental design. In the above analyses, we include the fixed effects to control for the differences caused by the experimental and estimation methods. While the pooled results suggest that in general a positive and significant mediation effect does exist, in this section, we perform the following analyses to investigate whether and how such effect maintains across the three elicitation methods. As the MPL and CTB groups have insufficient sample sizes, and thus do not allow us to conclude valid statistical results from separated analyses, we include the interactive effects between the mediators and the time discounting elicitation methods in the pooled analyses to reach this goal.

Table D.3 and D.4 presents the results of analyses in which the interactive effects with time discounting elicitation methods are involved. Panel A repeats the results presented in Table 5 and Column (1) and (3) in Panel B repeats that in Table 6. Column (2) and (4) in Panel B include the interactive effect of time overestimation with different time discounting elicitation methods. As the main analyses already show that the mediation effect does not preserve in the low load treatment when the S/O ratio serves as the mediator variable, we focus on Table D.3 in which time overestimation is regarded as the mediator, which shows that while the significance of the effect is rather close between the MPL and the TTO group, suggested by both the insignificant coefficient of $O\hat{VER}_i(u) \cdot TTO$ in Panel B and the results in Panel C, the bias in the time perception seems to have the strongest effects on the discounting behaviors of subjects in the CTB group.[28] Nevertheless, the results of the mediation analyses displayed in Panel C imply that though weaker, the positive mediation effects maintain at least marginal significances in all three groups.

[28] A possible explanation is the experimental design: the MPL and the TTO share the similar design of choosing between two options, whereas subjects in the CTB group need to allocate a budget to smooth the payment between sooner and later.

Table D.3: Mediation Effects under Different Time Discounting Elicitation Methods

Panel A. Regressions of the mediator variable: $O\hat{VER}_i(u)$

	$u = 0.5$	$u = 1$
	(1)	(2)
T_i	-.6675**	-.7053**
	(.2765)	(.2765)
R^2	.0222	.0242

Panel B. Regressions of the outcome variables: $\tilde{\varphi}_i(u)$

	$u = 0.5$		$u = 1$	
	(1)	(2)	(3)	(4)
T_i	-.0166	-.0210	-.0160	-.0207
	(.0162)	(.0165)	(.0167)	(.0168)
$O\hat{VER}_i(u)$		-.0087†		-.0127†
		(.0058)		(.0089)
$O\hat{VER}_i(u) \cdot CTB$		-.1253**		-.1209**
		(.0479)		(.0484)
$O\hat{VER}_i(u) \cdot TTO$		-.0062		.0073
		(.0174)		(.0190)
R^2	.1089	0.1175	.0930	.1233

Panel C. Results of the mediation analyses

	$u = 0.5$			$u = 1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	MPL	CTB	TTO	MPL	CTB	TTO
ACME	.0028†	.0248**	.0016*	.0012	.0265**	.0026†
	[-.0036, .0115]	[.0013, .0569]	[-.0006, .0047]	[-.0063, .0096]	[.0025, .0600]	[-.0010, .0076]
		.0077*			.0084†	
		[-.0008, .0194]			[-.0012, .0213]	
Direct effect	-.0085	-.0777**	.0120	-.0031	-.0752**	.0063
	[-.0419, .0255]	[-.1395, -.0166]	[-.0237, .0486]	[-.0384, .0335]	[-.1366, -.0150]	[-.0298, .0429]
		-.0151			-.0164	
		[-.0569, .0272]			[-0.0587, .0262]	
Total effect		-.0074			-.0080	
		[-.0289, .0233]			[-.0300, .0238]	
% of total effect mediated	-37.84%	-335.14%	-21.62%	-15.00%	-331.25%	-32.50%
		-104.05%			-105.00%	

Notes:

1) $O\hat{VER}_i(u)$ equals 1 if subject i overestimates u day(s) and 0 otherwise. $T_i = 1$ for high load treatment and 0 otherwise. $\tilde{\varphi}_i(u)$ is the estimated discount function of u day(s).

2) For Panel A and B, coefficient estimates are presented with standard deviations in parentheses.

3) For Panel C, point estimators are presented with 95% confidence intervals in brackets. The algorithm proposed in Imai, Keele and Tingley (2010) is employed. Hicks and Tingley (2010)'s STATA package is implemented to perform the estimation. The point estimators do not equal exactly to the coefficient products because nonlinear regressions are involved.

4) Errors are clustered at the individual level. The fixed effects of different time preference elicitation methods are included. The average performance of the filler tasks and the currency difference are controlled for.

5) Weighted regressions are performed so that each individual is equally weighted despite the different amount of observations in different groups.

6) †, *, **, *** denote significance at the 0.15, 0.1, 0.05 and 0.01 level.

Table D.4: Mediation Effects under Different Time Discounting Elicitation Methods: S/O Ratio as the Mediator

	$u = 0.5$			$u = 1$		
Panel A. Regressions of the mediator variable: $\hat{SO}_i(u)$						
T_i		-.7302**			-.9510**	
		(.3177)			(.4407)	
R^2		.0232			.0207	
Panel B. Regressions of the outcome variables: $\tilde{\varphi}_i(u)$						
T_i	-.0166	-.0190		-.0160	-.0192	
	(.0162)	(.0167)		(.0167)	(.0172)	
$\hat{SO}_i(u)$		-.0015			-.0113	
		(.0103)			(.0114)	
$\hat{SO}_i(u) \cdot CTB$		-.0173			-.0031	
		(.0197)			(.0132)	
$\hat{SO}_i(u) \cdot TTO$		.0010			.0103	
		(.0147)			(.0114)	
R^2	.1089	0.1175		.0930	.1233	
Panel C. Results of the mediation analyses						
	MPL	CTB	TTO	MPL	CTB	TTO
ACME	.0099	.0138*	.0008†	.0111	.0136*	.0010†
	[-.0079, .0324]	[-.0014, .0364]	[-.0002, .0022]	[-.0113, .0384]	[-.0013, .0358]	[-.0004, .0030]
		.0060			.0063	
		[-.0021, .0173]			[-.0029, .0188]	
Direct effect	-.0315	-.0408	.0025	-.0302	-.0351	-.0014
	[-.0792, .0141]	[-.0896, .0072]	[-.0278, .0331]	[-.0779, .0159]	[-.0779, .0075]	[-.0333, .0309]
		-.0157			-.0161	
		[-.0544, .0225]			[-0.0541, .0218]	
Total effect		-.0097			-.0098	
		[-.0283, .0199]			[-.0285, .0203]	
% of total effect	-102.06%	-142.27%	-8.25%	-113,27%	-138.78%	-10.20%
mediated		-61.86%			-64.29%	

Notes:

1) $\hat{SO}_i(u)$ is the extrapolated S/O ratio of u days. $T_i = 1$ for high load treatment and 0 otherwise. $\tilde{\varphi}_i(u)$ is the estimated discount function of u day(s).

2) For Panel A and B, coefficient estimates are presented with standard deviations in parentheses.

3) For Panel C, point estimators are presented with 95% confidence intervals in brackets. The algorithm proposed in Imai, Keele and Tingley (2010) is employed. Hicks and Tingley (2010)'s STATA package is implemented to perform the estimation. The point estimators do not equal exactly to the coefficient products because nonlinear regressions are involved.

4) Errors are clustered at the individual level. The fixed effects of different time preference elicitation methods are included. The average performance of the filler tasks and the currency difference are controlled for.

5) Weighted regressions are performed so that each individual is equally weighted despite the different amount of observations in different groups.

6) †, *, **, *** denote significance at the 0.15, 0.1, 0.05 and 0.01 level.