Original Article

Rate maximization and hyperbolic discounting in human experiential intertemporal decision making

Maayke Suzanne Seinstra, Manuela Sellitto, and Tobias Kalenscher

Comparative Psychology, Institute of Experimental Psychology, Heinrich-Heine University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany

Decisions between differently timed outcomes are a well-studied topic in as diverse academic disciplines as economics, psychology, and behavioral ecology. Humans and other animals have been shown to make these intertemporal choices by hyperbolically devaluing rewards as a function of their delays ("delay discounting"), thus often deemed to behave myopically. In behavioral ecology, however, intertemporal choices are assumed to meet optimization principles, that is, the maximization of energy or reward rate. Thus far, it is unclear how different approaches assuming these 2 currencies, reward devaluation and reward rate maximization, could be reconciled. Here, we investigated the degree at which humans ($N = 81$) discount reward value and maximize reward rate when making intertemporal decisions. We found that both hyperbolic discounting and rate maximization well approximated the choices made in a range of different intertemporal choice design conditions. Notably, rate maximization rules provided even better fits to the choice data than hyperbolic discounting models in all conditions. Interestingly, in contrast to previous findings, rate maximization was universally observed in all choice frames, and not confined to foraging settings. Moreover, rate maximization correlated with the degree of hyperbolic discounting in all conditions. This finding is in line with the possibility that evolution has favored hyperbolic discounting because it subserves reward rate maximization by allowing for flexible adjustment of preference for smaller, sooner or larger, later rewards. Thus, rate maximization may be a universal principle that has shaped intertemporal decision making in general and across a wide range of choice problems.

Key words: **hyperbolic discounting, intertemporal choice, preference reversal, reward rate maximization.**

INTRODUCTION

In our daily life, we make countless decisions between delayed consequences. These intertemporal decisions shape important aspects of our life, such as education, housing, diet, and financial well-being. Intertemporal decision making is well studied in both humans and nonhuman animals (Kalenscher *et al.* 2005; Rosati *et al.* 2007; Kalenscher and Pennartz 2008; Sellitto *et al.* 2011) by as diverse academic disciplines as economics, psychology, and behavioral ecology. All these fields share their interest in the typical behavior, common to humans and other animals, of overweighting short-term outcomes or underweighting long-term outcomes and, by consequence, of making impulsive decisions (Kalenscher *et al.* 2005, 2008; Namboodiri and Hussain Shuler 2016). However, although trying to explain the same phenomenon—intertemporal

choice and impulsive decision—the approaches in these different disciplines came up with different accounts.

Economics and psychology literature addressed great attention to intertemporal decision making because the myopic, short-sighted choice patterns of humans and other animals represent violations of the efficiency assumptions of utility maximization and time preference in economics (Kalenscher and Pennartz 2008). In behavioral economics and psychology, intertemporal choice behavior is typically expressed as delay discounting (Samuelson 1937; Kalenscher and Pennartz 2008; Hayden 2016), according to which the subjective value of a delayed reward decreases with increasing delay of its receipt (Frederick *et al.* 2002; Sellitto *et al.* 2011).

In both humans and other animals, delay discounting is best described by hyperbolic discounting models, which reflect a decrease in the subjective value of a reward with a nonconstant decay rate, characterized by a steep decline in subjective value at initial delays, and flatter decline at longer delays (Mazur 1984; Green and Myerson 1996; Kalenscher and Pennartz 2008) Due

Address correspondence to M. Sellitto. E-mail: sellitto@uni-duesseldorf.de. M.S.S. and M.S. contributed equally to this work.

to this property, hyperbolic discounting well explains the so-called "preference reversals", which previous exponential discounting models failed to account for (e.g., Frederick *et al.* 2002; Sellitto *et al.* 2011). When choosing between smaller-sooner and larger-later rewards, humans and nonhuman animals often reverse their preference when front-end delays are added or subtracted from a choice set (Green *et al.* 1994; Kirby and Herrnstein 1995). For example, even though an individual may prefer (A) €10 today over (B) €20 in 6 months, she may prefer (B') €20 in 1 year over (A') €10 in 6 months (Frederick *et al.* 2002), despite the discounted utility theory (DUT) in economics (Samuelson 1937) prescribes that a rational agent should meet the *stationarity axiom* and choose option A' over option B' since she preferred option A before (Fishburn and Rubinstein 1982).

Next to the economics approach, humans and nonhuman animals' myopic choices have also drawn the attention of the optimal foraging theory in behavioral ecology (Stephens and Krebs 1986; Bateson and Kachelnik 1996; see also Hayden 2016 for review). Inspired by evolution theory, optimal foraging theory prescribes that a Darwinian-fitness-maximizing organism should maximize energy intake over time—principle of energy rate maximization—when foraging for food (Pyke *et al.* 1977). However, impulsive decisions that, as mentioned before, do not meet the efficiency assumptions of utility maximization and time preference in economics (Kalenscher and Pennartz 2008), also apparently fail to maximize long-term energy rate (Mcdiarmid and Rilling 1965; Kalenscher *et al.* 2005; Kalenscher and Pennartz 2008). To reconcile these findings with the assumption in optimal foraging theory that evolution should have shaped optimal intertemporal decision making, Stephens and colleagues (Stephens and Anderson 2001; Stephens *et al.* 2004) argued that short-sighted, present-biased decisions *can* result in energy rate maximization, but only in natural foraging contexts to which animals' decision systems are adapted to.

Natural foraging contexts are characterized by sequential background-foreground problems (Stephens 2008; Rosati and Stevens 2009) in which one alternative is the background to all other alternatives. For instance, a flying bird spotting a potential food source has to decide whether to put its background activity (flying) on hold to exploit the potential food source (foreground), or whether to continue the exploration of the environment to find a potentially richer/safer source later. The same happens in humans when, for instance, someone has to decide whether to accept a job offer and settle or keep searching for better opportunities. However, in most laboratory studies, intertemporal decisions are typically *not* probed with sequential choice problems—so-called *patch-designs*—that are supposed to have high ecological validity, but with binary, mutually exclusive choice tasks—so-called *self-control tasks*: "choose either A or B"—to which subjects are supposedly *not* adapted to. By consequence, individuals have been shown to apparently fail to maximize energy rate in self-control tasks (Rosati *et al.* 2007; Kalenscher and Pennartz 2008).

Why does it seem that we and other animals fail to maximize long-term energy rate in self-control tasks, although we *are* thought to maximize reward rate? One answer could be that long-term energy maximization is achieved because short-sighted decision rules that only minimize the delay to the next reward, ignoring other task features such as postreward delays (e.g., Blanchard *et al.* 2013), automatically also lead to long-term rate (LTR) maximization in ecologically valid patch designs (Stephens and Anderson 2001). Organisms may thus have evolved to implement short-sighted rules because they lead to LTR maximization in sequential choice contexts, even though they result in poor performance on binary self-control problems. This has been indeed shown in animals (Stephens and Anderson 2001; Stephens and McLinn 2003) and more recently also in humans (e.g., Schweighofer *et al.* 2006; Bixter and Luhman 2013; Zarr *et al.* 2014; Carter *et al.* 2015).

A striking illustration of how organisms may implement short-sighted rules in order to achieve LTR maximization lies again in preference reversals—which, as said before, seem to indicate that individuals overweight short-term outcomes (Thaler 1981; Benzion *et al.* 1989) and, from a normative economic perspective, that they act against their own future interest. If we go back to the previous example and consider preference reversals from the perspective of reward rate maximization, choosing option A (€10 today) would yield a reward rate of €10 per day, and choosing option B (€20 in 6 months) would yield a reward rate of €0.10 per day. The rate maximization principle would prescribe choosing option A over option B because of its higher reward rate. However, if both outcomes were then shifted in time by a front-end delay of 6 months, the alternatives would now be option A': €10 in 6 months—which yields a rate of €0.05 per day—and option B': €20 in 1 year—which yields a rate of €0.11 per day. While, as said before, the DUT in economics (Samuelson 1937) prescribes that a rational agent should meet the *stationarity axiom* and choose option A' since she preferred option A before (Fishburn and Rubinstein 1982), option B' yields a higher rate in the new pair, therefore a reward-rate maximizing agent should reverse her preference, and choose B' over A'. Hence, rate maximization could only be achieved by a decision rule allowing for time-inconsistent preference reversals. Because rate maximization models have been developed to account for nonhuman animals' foraging behavior, the logic of our example may be better understood when replacing financial rewards with food rewards. Consider an animal that chooses between option A: 2 food-items in 2 s (rate: 1 item/s) and option B: 4 items in 8 s (rate: 0.5 items/s). The rate maximization principle would prescribe choosing option A because of its higher energy rate. If both outcomes were then shifted in time by 10 s, the alternatives would now yield A': 2 food-items in 10 + 2 s (rate: 0.17 items/s) and B': 4 items in 10 + 8 s (rate: 0.22 items/s). Now, even though a hypothetical, economically ideal, and time-consistent forager should choose option A' over B', rate maximization would prescribe a preference reversal, thus choosing B' over A'. Note that the logic of these examples still holds when extending them to single or repeated choice scenarios with nonexistent (in one-shot choices), fixed or variable postreward delays.

To date, it remains unclear, however, whether rate maximization and hyperbolic discounting are 2 contradicting, possibly irresoluble concepts (Stephens *et al.* 2004), or whether they are 2 sides of the same coin that, when considered together, can unravel the evolutionary mystery of short-sighted intertemporal choice. Here, we address this question through an experiential intertemporal choice task. We asked human participants to make both binary and sequential intertemporal choices between smaller-sooner and larger-later monetary rewards, with immediately experienced delays. Crucially, depending on the delay parameters, reward rate maximization required choosing the larger-later option in some trial blocks, and choosing the smaller-sooner option in other trial-blocks. Thus, an ideal optimal forager should flexibly shift her preferences between smaller and larger rewards. We adopted a repeated-measures design with 2 design conditions (self-control vs. patch), which enabled us to obtain individual discount rates and rate maximization scores for each of them to investigate how human participants maximize long-term reward rate in comparison to how they devalue future rewards.

METHODS

Participants

We recruited 93 participants (60 female) at the Heinrich-Heine University Düsseldorf. Exclusion criteria were psychiatric or psychological disorders, lack of German language proficiency, smoking more than 5 cigarettes per day, drinking more than 1 bottle of wine or 1.5 L beer a day on average, and consumption of recreational illicit drugs more than 2 times a month. These criteria were chosen to avoid drug-related effects on intertemporal decision making (Bickel *et al.* 2012). Participants were between 18 and 45 years old (M = 23.2, SD = 5.2) and were enrolled in various study programs (language studies: 22; psychology: 13; (business) economics: 9; history: 8; computer science: 6; law: 6; media and culture: 6; biology: 5; other studies [*n* < 5]: 20). Participants received a monetary reimbursement consisting of a show up fee of 3€ plus their earnings during one part of the experiment (see below), which could lead up to a total amount of 17€. Payment was received in the form of a personal cheque at the end of the session. This study was approved by the local ethical committee of the Psychology department at the Heinrich-Heine University Düsseldorf.

Materials

General task procedure

Participants made a series of choices between a smaller-sooner (SS) monetary reward and a larger-later (LL) monetary reward. The nature of the task was experiential, that is, delays and rewards were real and experienced by the participants. In a within-subject design, we manipulated the design of an intertemporal choice task (sequential "patch" condition vs. binary "self-control" condition, see below and Figure 1).

Each design condition consisted of 6 separate blocks of trials that varied in delay to the smaller-sooner reward as well as the delay to the larger-later reward (see Table 1; in our task, the delay indicates the time between the decision and the onset of the reward screen, informing the participant about the reward magnitude, see below).

Figure 1
Task structure in the self-control (A) and patch (B) condition. Choices were made between a smaller, sooner (SS) and a larger, later (LL) option. One grey circle indicates a reward of 5 cents. ITI: inter-trial interval; D = delay; R = reward.

Each of the 6 blocks was presented in the self-control as well as the patch design (see below and Figure 1). The 3 blocks with the same delay to the small reward (i.e., either 3s or 9s) within a task design were presented together in a cluster (to maintain some structure in the task for participants; note that the blocks in one cluster differed in the delay to the larger-later reward only). Within each cluster, the blocks were presented in pseudo-random fashion. Participants thus completed 2 clusters of 3 blocks each in the self-control design, and 2 clusters of 3 blocks each in the patch-design. After each cluster, participants had a short, approximately 1-min break while the next cluster was started. The clusters were presented pseudo-randomly as well.

Participants made one decision per trial; the number of trials per block was variable; trials in a block were repeated until the block duration elapsed. Block duration was fixed and determined such that participants could choose the option with the longest delay at least 7 times in each block, including a decision time of 5 s per trial.
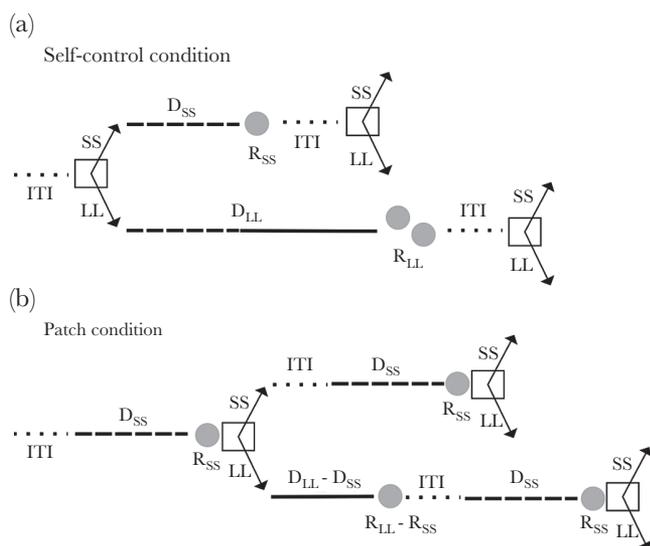
Self-control design

In the self-control design condition (see Figure 1), participants made binary, binding choices between smaller-sooner and larger-later rewards. The smaller-sooner reward consisted of 5 cents and was delayed by either 3 s or 9 s. The larger-later reward consisted of 10 cents, with a delay of 5 s, 10 s, or 15 s (with smaller-sooner delay of 3 s), or 11 s, 21 s, or 31 s (with smaller-sooner delay of 9 s). The delay of the larger-later option was varied across 3 blocks of trials in a given condition in a pseudo-random fashion so that each block yielded a new pair of options; delay/reward option pairs were kept constant across trials within a block. Trial duration was not fixed; the number of trials per block was variable and depended on block duration.

Participants were not instructed about delay and reward magnitudes, but had to learn them by experience. A trial started with the intertrial interval (ITI), indicated by a white cross at the center of the screen, which was fixed at 5 s. The ITI was followed by the choice screen, on which 2 differently colored circles were presented on each side of the screen. The different delay/reward combinations were associated with unique circle-colors. Participants indicated their choice on a standard keyboard by pressing the "x" key for the left option, and the "m" key for the right option. Key-side assignment was also indicated on the screen below the circles for participants' convenience. Participants had unlimited time to make their decisions, but after 3 seconds they were prompted by the message "please make a choice", blinking red below the circles on the screen. After participants selected one of the colored circles, a dynamic progress bar indicated the delay length until reward presentation. After the delay, information about the reward magnitude was shown at the center of the screen for 2 s, and the cumulated earnings across past trials were additionally shown below the reward information. Following reward presentation, the next trial started immediately. Trials were repeated within a block until the block duration expired. When the block time was up in the middle of a trial, this trial was finished before the next block started.

Patch-design

The 2 clusters with a patch design were economically identical to the self-control condition in terms of delays, rewards, trial and block structure, screen composition, information format, as well as participant instructions. The only difference to the self-control condition was the sequential nature of the decision structure: while, in the self-control condition, participants made binding binary choices

Table 1

Task parameters per block

Block	R_{SS}	R_{LL}	D_{SS}	D_{LL}	ITI[a]	rr_{SS}[b]	rr_{LL}	ΔLTR[c]	Block duration
1	5 Cent	10 Cent	3 s	5 s	5 s	0.63	1.00	37.50	119 s
2	5 Cent	10 Cent	3 s	10 s	5 s	0.63	0.67	4.17	154 s
3	5 Cent	10 Cent	3 s	15 s	5 s	0.63	0.50	−12.50	189 s
4	5 Cent	10 Cent	9 s	11 s	5 s	0.36	0.63	26.79	161 s
5	5 Cent	10 Cent	9 s	21 s	5 s	0.36	0.38	2.75	231 s
6	5 Cent	10 Cent	9 s	31 s	5 s	0.36	0.28	−7.94	301 s

Blocks and parameters were identical in the self-control and patch designs.
[a]ITI = Intertrial interval.
[b]rr = Reward rate.
[c]Long-term rate (LTR) difference between the SS and LL option. A positive value indicates a higher LTR for the LL option.

between the smaller-sooner and larger-later rewards, in the patch condition they chose whether to stay in a "reward patch" for a fixed delay to obtain a large reward, or "leave the patch" and start a new trial after having obtained a small reward (see Figure 1). Sequential choice was implemented as follows: each trial started with the ITI (5 s), followed by a delay of 3 s (delays were indicated by dynamic progress bars as in the self-control condition) or 9 s. Subsequently, a reward screen (2 s) indicated that the participant had earned 5 cents (the smaller-sooner reward magnitude), after which the choice screen was presented. Participants indicated their choice on a standard keyboard by pressing the "x" key for the left option, and the "m" key for the right option. A choice of the smaller-sooner option resulted in the start of the next trial (i.e., was followed by the ITI of the next trial) and a choice of the larger-later option resulted in a further delay of 2 s, 7 s, or 12 s in the 3 s smaller-sooner delay blocks, or a further delay of 2 s, 12 s, or 32 s in the 9 s smaller-sooner delay blocks. Following the end of the delay, a further screen indicated that participants earned another 5 cents (thus, resulting in a sum of 5 + 5 = 10 cents in this trial, equivalent to the magnitude of a larger-later reward), and the next trial started. Again, the order of delay conditions was pseudo-randomized across blocks.

As mentioned, block duration, trial setup, and general design features were identical in the patch- and the self-control conditions. Also, as before, participants were not instructed about the outcome parameters, but had to learn them through experience. Note that, in the patch condition, the prechoice delays (3 s or 9 s) and default rewards (5 cents in all conditions) were identical to the smaller-sooner rewards in the self-control condition (see above and Figure 1), and the sum of pre- and post-choice delays in the patch condition (5 s, 10 s, and 15 s for blocks 1–3 and 11 s, 21 s, and 31 s for blocks 4–6) as well as the sum of rewards (10 cents) matched the larger-later parameters in the self-control condition.

All conditions were fully incentive-compatible and accumulated earnings were paid out to the participants after experiment completion. The task was programmed in Matlab (Mathworks, Inc.) using the Cogent Graphics toolbox developed by John Romaya at the LON at the Wellcome Department of Imaging Neuroscience.

Offline delay discounting task

To obtain an offline measure of the participants' hyperbolic discount rates, we used a task design similar to the one described by Kirby *et al.* (1999). This enabled us to compare participant's hyperbolic discount rates in a task structure commonly used to measure hyperbolic discounting with the hyperbolic discount rates in the general task described above. This task estimated the individual discount rates k by assuming a hyperbolic discount function underlying

choice behavior. The task consisted of 27 choices between hypothetical rewards. In each trial, participants were offered the choice between a smaller reward available now and a larger but delayed reward. The smaller rewards ranged between €11 and €80, and the larger rewards between €25 and €85. The delays ranged between 7 and 186 days. Combinations of reward amounts and delays were such that indifference between the options would yield one of 9 distinct discount rate k_{Kirby}, that is, there were 9 sets of 3 trials yielding the same k-value, one with a relatively small, medium, and large delayed reward. Trials were presented in a specific order. One option was presented on the left of the screen, while the alternative option was presented on the right side of the screen. Participants had to press "x" or "m" to choose the left or the right option, respectively. Participants had unlimited time to make their decisions. At the start of the task participants were asked to make the choices in accordance with their personal preference, and that there were no right or wrong answers. Participants were informed beforehand that this task would not be reimbursed.

Post-test questionnaire

This questionnaire consisted of questions about demographics (age, income, marital status, nationality, profession, field of study), questions regarding current physical state (known diseases, psychiatric treatment, smoking behavior, alcohol use) as well as questions regarding the decision tasks: we asked whether participants had problems focusing on the task (yes/no), how easy it was to understand the tasks (5-point Likert scale), which strategy they used when making their choices (open question), whether they calculated the total duration of choice options (yes/no), to what extent they tried to obtain the highest possible reward (5-point Likert scale), whether they always chose the same color, independent of the outcome (always, often, sometimes, or never), whether their choices reflected their personal preferences (yes/no), and whether we could trust their answers (yes/no).

Additional measures

We additionally measured self-reported impulsivity using the Quick Delay Questionnaire (QDQ) and the Barratt Impulsiveness Scale (BIS) as well as time perception using a time production task. For procedure and results, see Supplementary Materials.

Procedure

Upon arrival, participants were asked to read and sign an informed consent form and the procedure of the session. The number of participants tested at the same time ranged from 1 to 4. Each participant was seated in his/her own cubicle that ensured privacy

throughout the session. Identical laptops were used to ensure similar processing speed. No other participants or the experimenter could see the laptop screens during task performance. Before staring the tasks, participants received written instructions. The instructions stressed, among others, that, although the 4 tasks (i.e., conditions) may look similar, they were independent of each other. In addition, participants were told that each task had a fixed duration, independent of the choices that were made, and that their earnings depended on their choices. After written and verbal instructions and an opportunity for questions and answers, participants performed the 4 task conditions in random order. After each task condition, participants saw the monetary amount they had earned in that particular condition and were prompted to ask the experimenter to start the next task. The main task was followed by Kirby's discounting task, before which the participant received short oral instructions that were also repeated on screen before the task started. This was followed by the time production task, and QDQ and BIS questionnaires (see Supplementary Materials for results at these tasks). Finally, the participants filled out the post-test questionnaire. Participants then received a show-up fee of €3 plus their earnings from the main task in the form of a personal cheque that they could cash at any bank. If requested, participants were informed about the aim of the study.

Analysis

Rate maximization scores

The choice alternatives in each trial differed in their long-term reward rate (here: the cumulative reward amount per block; note that larger, later rewards do not always yield higher reward rates; depending on the task parameters, choices of smaller, sooner rewards may produce more optimal outcomes, see Table 1 for details). To estimate to what extent individuals maximize long-term reward rate we calculated LTR scores, which reflect the proportion of choices of the alternative with the highest reward rate, averaged across all 6 blocks in each design condition, resulting in 2 rate scores per individual.

We used a softmax rule to approximate the probability of choosing the alternative with the highest reward rate:

$$p_j = 1 / \left(1 + e^{-\mu*(C)}\right) \tag{1}$$

in which p is the proportion of choices for the alternative with the highest reward rate in block j, μ is a temperature parameter indicating the sensitivity to differences in reward rates, and C is the currency to be maximized, here, reflecting the difference in reward rates. Goodness of fit was estimated using the Akaike Information Criterion (AIC).

Hyperbolic discounting

To estimate hyperbolic discounting, we used the same softmax decision rule in Equation 1 to estimate hyperbolic discount rates k from the proportion of choices for the larger-later reward p_{LL}.

For hyperbolic discounting, the currency C in Equation 1 was given by v_{LL} - v_{SS}, where v_{LL} and v_{SS} were the subjective, discounted values of the larger-later reward in block j, or smaller-sooner reward, respectively, obtained from Mazur's hyperbolic model (Mazur 1984):

$$v_i = \frac{R_i}{1 + k(D_i)} \tag{2}$$

where v_i indicates the subjective, time-discounted reward value of reward i with reward magnitude R, and delay D. k is an individual discount factor determining the steepness of the discount function.

We used all 6 blocks of each design (self-control and patch) to estimate the individual discount parameter k. We computed a single k-value per participant, pooling across trials from both design conditions. Additionally, separate k-values were estimated for each design condition, resulting in 2 different model fits for each individual. Reward magnitude R and delay D in Equation 2 was adjusted for each design (see Figure 1). Again, goodness of fit was estimated using the AIC.

Model comparisons and data analysis

All parameter estimations were performed using least squares methods in MATLAB R2011a (Mathworks, Inc). When estimates in raw form as well as their log transformations violated the normality assumption, nonparametric tests were performed.

Predictions

Table 2 shows the predicted choice preferences per block for the rate maximization and hyperbolic discounting model. The predictions of the hyperbolic model depend on the individual discount parameter k estimates.

RESULTS

Task and trial completion

Twelve participants were excluded because they indicated, in the postexperiment debriefing questionnaires, having based their choice on the option with their favorite color ($N = 4$), to be unmotivated or unwilling to maximize their payoff ($N = 2$), to deliberately choose against their preference ($N = 5$), or they indicated that their given answers were not to be trusted ($N = 1$). Together this resulted in a final sample of 81 participants (mean age = 23.2, SD = 5.0).

The number of trials per block was variable. On average, participants completed 11 trials in the first, 13 trials in the second, and 17 trials in the third block in each task design (note that the more often the smaller-sooner reward was chosen, the more trials could be completed within the fixed time). There were no notable differences in number of trials completed between the 4 conditions. All participants completed at least 7 trials in each block, except for one participant who completed only one trial in the second block of the patch condition (this block was excluded from further analysis). Therefore, for each participant, the first 7 trials per block were used in *all* subsequent analyses.

### Manipulation check: sensitivity to parameter manipulations

As a manipulation check, we tested whether participants were sensitive to the delay differences across blocks. To this end, we compared *the proportion of large reward choices* (p_{LL}) between blocks with similar smaller-sooner reward delay within each design condition (Figure 2). There was a significant difference in p_{LL} across blocks within each smaller-sooner delay (3 s and 9 s) and design (self-control and patch) condition: Friedman's chi-square test for multiple repeated measures, *all* $\chi^2 > 11.00$, *all* $P < 0.003$.

Also within each smaller-sooner delay and task design, participants were sensitive to the changes in delay to the large reward: Wilcoxon pair-wise comparisons showed significant differences in p_{LL} between consecutive blocks with similar smaller-sooner delays, *all* $Z < -3.5$, *all* $P < 0.001$, with the exception of patch-condition (3 s), block 2 versus 3: $Z = -1.09$, $P = 0.274$.

Table 2

Predicted preference for the SS or LL reward per block per decision model

Block	Maximizing LTR: both designs	Discounting: self-control design	Discounting: patch design
1	LL	LL	LL
2	LL	$k < 0.25$: LL	LL
		$k > 0.25$: SS	
3	SS	$k < 0.12$: LL	SS
		$k > 0.12$: SS	
4	LL	LL	LL
5	LL	$k < 0.35$: LL	LL
		$k > 0.35$: SS	
6	SS	$k < 0.09$: LL	SS
		$k > 0.09$: SS	

Predictions for LTR maximization were based on the calculation of reward rates using the total delay (prereward delay + ITI) and reward of each option. Predictions with regard to delay discounting were based on the discounted value of the options, which were calculated using Mazur's hyperbolic function (Mazur 1984). Only prereward delays were included when calculating the discounted value for k-values ranging from 0.0 to 1.0.
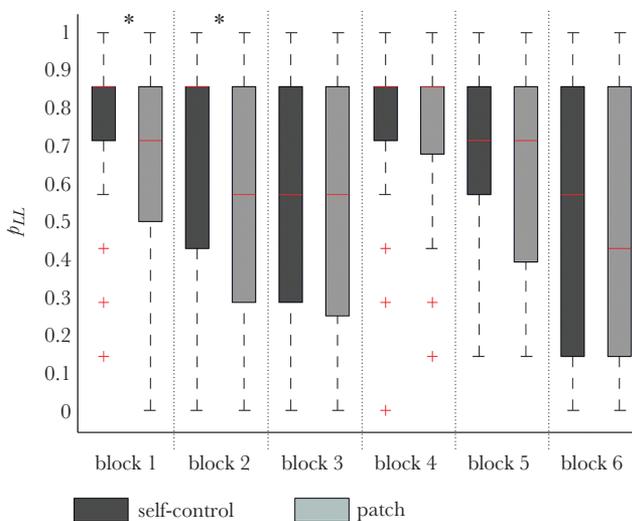
Figure 2
Boxplots of the proportion of choices for the large reward (pLL) in each blocks per condition.

These results suggest that participants were sensitive to reward delays and magnitudes.

Choice behavior

Choice proportions

Choice proportions were mostly similar between design conditions: block-wise comparisons (Wilcoxon) of p_{LL} choices between self-control and patch conditions revealed no significant effect of design, all $Z > -1.13$, all $P > 0.257$, except in blocks 1 and 2, block 1: $Z = -2.60$, $P = 0.009$, $r = 0.20$; block 2: $Z = -2.71$, $P = 0.007$, $r = 0.21$. In blocks 1 and 2, the proportion of large reward choices was significantly higher in the self-control than patch design.

Rate maximization

The LTR scores indicate to what extent participants' choices produced long-term reward maximization. The median scores were

0.64 (LTR$_{self-control}$) and 0.60 (LTR$_{patch}$) (see Table 3). A comparison of LTR scores showed significantly higher scores in the self-control than patch condition, $Z = -2.08$, $P = 0.038$, $r = 0.16$, indicating that participants selected the choice alternative with the higher LTR score more often in the self-control than the patch condition. Moreover, since it is possible that participants were still learning the reward contingencies in the first 7 trials, we repeated this analysis on LTR scores for both designs on the last 5 choices of each participant. We replicated the above mentioned results on LTR scores, $Z = -2.87$, $P = 0.004$, $r = 0.23$ (LTR$_{self-control}$: $Mdn = 0.67$, $range = 0.33-1$; LTR$_{patch}$: $Mdn = 0.67$, $range = 0.27-1$). In line with this difference in LTR scores between self-control and patch conditions, LTR scores were not significantly correlated between conditions, $r_s = 0.16$, $P = 0.156$ (see Table 4), indicating that participants did not maximize long-term reward rate to the same extent across design conditions.

These results suggest that, unlike in previous animal (e.g., Stephens and Anderson 2001) and human experiments (e.g., Carter *et al.* 2015), optimal decision making was not restricted to a sequential patch design.

Hyperbolic discounting

The log-k values did not differ between the 2 design conditions, $Z = -0.25$, $P = 0.80$, $r = 0.23$ (see Table 3), and they correlated with each other, $r_s = 0.34$, $P = 0.002$ (see Table 4). Moreover, LTR scores for the self-control condition and for the patch condition were positively correlated with k-values in the self-control condition, $r_s = 0.25$, $P = 0.026$, as well as in the patch condition, $r_s = 0.30$, $P = 0.008$, respectively. This indicates that higher discount parameters k went along with higher LTR maximization in both designs, implying that more impulsivity (the higher the k, the steeper the discounting) correlated with better long-term rate maximization.

Additionally, we computed k-values for both designs by considering all participants' choices (and not the first 7 only). Here, log-k values significantly differed between the 2 design conditions, $Z = -4.87$, $P < 0.001$, $r = 0.38$ (log-$k_{self-control}$: $Mdn = 0.01$, $range = 0.00-1$; log-k_{patch}: $Mdn = 0.99$, $range = 0.00-1$), with higher discount rate in the patch setting than in the self-control one (see also Carter and Redish 2016).

We additionally run Spearman correlations between k-values of the main task with k-values of Kirby's offline discounting task (see Table 3). The estimated k-values from Kirby's discounting task ($Mdn = 0.01$, $range = 0.0002-0.16$) were not correlated with the k-values of either of the 2 designs, although it positively correlated at the trend level with the k-values in the self-control condition, $r_s = 0.21$, $P = 0.060$ (see Table 4). These results make sense considering the binary design of Kirby's task, the much larger reward magnitudes and delays, and the fact that Kirby's task does not facilitate long-term considerations due to the task structure.

Earnings

The earnings within each design condition provide an indication of economic success. A Wilcoxon Signed Ranks test showed that earnings in the self-control condition ($Mdn = 6.70$, $range = 5.50-7.20$) were significantly higher compared to the earnings in the patch condition ($Mdn = 6.13$, $range = 5.15-6.65$), $Z = -7.65$, $P < 0.001$, $r = 0.60$.

Moreover, earnings were significantly correlated with LTR measures in both designs, but not with the hyperbolic discount parameter k of both designs (see Table 4). These results were corroborated by running a hierarchical regression on the *total earnings*

ize LTR, which was the case in 77.8% ($n = 63$) of participants. In
block 3, a rate-maximizing discounter should have a k-value higher
than 0.12, which was the case in 45.7% of participants ($n = 37$). In
block 5, 80.2% of all participants ($n = 65$) had a k-value lower than
0.35, thus maximizing LTR, and, in block 6, the k-value of 51.8%
of participants ($n = 42$) was higher than 0.09, again, maximizing
LTR. A Pearson Chi-square test revealed that, across all blocks,
the proportion of participants maximizing LTR was significantly
higher than the proportion of participants not maximizing LTR
(chi-square = 25, $P < 0.001$). The only block where the propor-
tion of LTR-maximizing discounters was descriptively smaller than
the proportion of nonmaximizers was block 3. In this block, a very
high level of impulsivity would have been needed for LTR maxi-
mization, and roughly half of our participants were too patient to
meet this strong impulsivity requirement. A similar trend could be
observed in block 6 where only slightly more than half of the par-
ticipants had sufficiently high discount rates to maximize LTR. The
observation that many participants were too patient to maximize
LTR in blocks 3 and 6, where a high level of impatience would
have been optimal, is in line with the positive correlation between
discount rates and LTR scores reported above: while all our partici-
pants were patient enough to match the LL preferences predicted
by the LTR model in blocks 2 and 5, our more impulsive partic-
ipants, in contrast to the patient ones, had time preferences that
matched the LTR prescriptions for SS choices in blocks 3 and 6. In
conclusion, SS preferences in blocks 3 and 6 seem to contribute to
some extent to the positive correlation between k-values and LTR

scores. Hence, these data support the idea that, from a LTR maximization perspective, a certain level of impulsiveness is preferable over strong patience: subjects with higher k-values tended to maximize reward rate to a larger extent than flat discounters because it allowed them to flexibly shift between LL and SS preferences across blocks.

DISCUSSION

In the present study, we examined how well hyperbolic discounting and reward rate maximization explain human choice behavior in an experiential intertemporal decision making task. To this end, we compared a hyperbolic discounting model and a reward rate maximization model, using choice behavior in the "classical" binary-choice self-control design as well as in the putatively more ecologically valid patch design. The hyperbolic model explained choices in the self-control and the patch designs equally well. The same was true for the long-term rate (LTR) maximization model, which provided equally good fits to participants' choices in both design. Overall, however, the LTR maximization model provided a better fit to the data than the hyperbolic discounting model in both the self-control and patch designs combined. Moreover, LTR maximization scores were higher in the self-control design than in the patch design, while no difference in participants' degree of discounting between the 2 paradigms emerged. This finding, in contrast to previous animal and human literature showing better performance in patch than self-control designs (e.g., Stephens and Anderson 2001; Stephens et al. 2004; Schweighofer et al. 2006; Bixter and Luhman 2013; Zarr et al. 2014; Carter et al. 2015), suggests that reward rate maximization *can* be universally observed in all choice frames, and it is not necessarily confined to foraging settings only. Additionally, reward rate maximization scores correlated with the degree of hyperbolic discounting in both paradigms, indicating that the higher the discount rate, the higher the long-term reward maximization. This result went along with final earnings that were higher in the self-control task than in the patch one. The finding that steeper discounting correlated with higher rate maximization scores as well as higher earnings is counterintuitive at first sight, as steep discounting is typically associated with short-sighted, myopic decision making, and, consequently, nonoptimal choice in the economics field (Frederick et al. 2002; Kalenscher and Pennartz 2008; Sellitto et al. 2011) (see below for elaboration). Why does hyperbolic discounting, the epitome of time-inconsistent preference (Kalenscher and Pennartz 2008), go hand in hand with reward rate maximization and higher total earnings in our tasks? We maintain that individuals maximize long-term reward rate in patch and self-control designs *for the very reason* that they implement a decision rule that happens to be consistent with hyperbolic discounting. We will elaborate on this in the following.

The key point is the insight that the so-called preference reversals that have led to the adoption of hyperbolic discounting models over exponential discounting models (Mazur 1984; Mazur 1987; Kalenscher and Pennartz 2008) are necessary to maximize reward rate. To explain this, we need to take a step back to normative economic DUT, which states that idealized rational decision makers should discount delayed rewards in a constant, exponential fashion, which implies stable choice preferences over time (Samuelson 1937). Time-consistent preferences can be epitomized by the *stationarity axiom*: when a subject prefers reward A at time t1 over reward B at time t2, she should also prefer reward A at t1+T over reward B at t2+T, that is, when a common time interval T, that

is, a front-end delay, is added to (or subtracted from) both delays (Fishburn and Rubinstein 1982). However, as said before, after introducing a front-end delay T (by adding or subtracting it) in the choice-set, humans and nonhuman animals often reverse their preference (Green et al. 1994; Kirby and Herrnstein 1995).

Preference reversals suggest that individuals attach disproportionally large weights to short-term outcomes (Thaler 1981; Benzion et al. 1989). This "present-bias" (also known as common difference effects or immediacy effects) is ubiquitous, yet it is an anomaly in choice because it causes violations of the *stationarity axiom* and, thus, goes along with time-inconsistent preferences. By consequence, from a normative economic perspective, it ultimately results in the tendency to act against one's own future interest.

The pervasiveness of present-biased, time-inconsistent preferences, and preference reversals is perplexing for economists, psychologists, and behavioral ecologists alike: What is the adaptive value of a choice pattern that so obviously creates nonoptimal results? One possible answer to this puzzle is, as mentioned, that natural selection has favored a decision rule that maximizes reward rate, not economic utility. Hyperbolic discounting, and the resulting propensity for preference reversals, supports reward rate maximization because, when introducing (adding or subtracting) a front-end delay T to the choice-set, the average reward rate of the 2 alternative options often reverses. Remember the example presented in the introduction: an animal chooses between option A: 2 food-items in 2 s (rate: 1 item/s) and option B: 4 items in 8 s (rate: 0.5 items/s). The rate maximization principle would prescribe choosing option A because of its higher energy rate. If both outcomes were then shifted in time by 10 s, the alternatives would now yield A': 2 food-items in 10 + 2 s (rate: 0.17 items/s) and B': 4 items in 10 + 8 s (rate: 0.22 items/s). Now, while DUT would impose time-consistent choice, that is, preference for A' over B', rate maximization would prescribe a preference reversal, thus choosing B' over A'. As mentioned in the introduction, the same logic also applies to single or repeated choices with nonexistent (in one-shot choices), fixed or variable postreward delays, and to different reward types, for example, financial rewards. Hence, rate maximization could only be achieved by a decision rule allowing for time-inconsistent preference reversals. Therefore, while DUT in economics prescribes that a rational agent should meet the *stationarity axiom*, optimal foraging theory would require the ability to flexibly shift preferences between smaller-sooner and larger-later rewards.

To understand why this example is not merely a special case, but illustrates a systematic, general requisite for flexible adjustment of preferences, one has to realize that reward rate does not drop at a constant rate with increasing front-end delays, but in a hyperbolic fashion (see Figure 3). By consequence, an optimal decision rule should systematically allow for flexible preference reversals in order to maximize reward rate in *any* choice situation with variably delayed outcomes. Or, in other words, to make optimal choices, a forager would have to do the very thing that economists stigmatize as irrational: show time-inconsistent preference reversals; were we the time-constant discounters prescribed by economic DUT, we would systematically fail to maximize reward rate when front-end delays were added to a binary choice set.

The logic illustrated in Figure 3 hinges on the natural occurrence of front-end delays. It is therefore important to note that the assumption that foraging animals very often experience such front-end delays in natural foraging scenarios, and that front-end delays matter for their foraging decisions, is realistic. Consider the quintessential choice a foraging animal has to make—whether to stay in
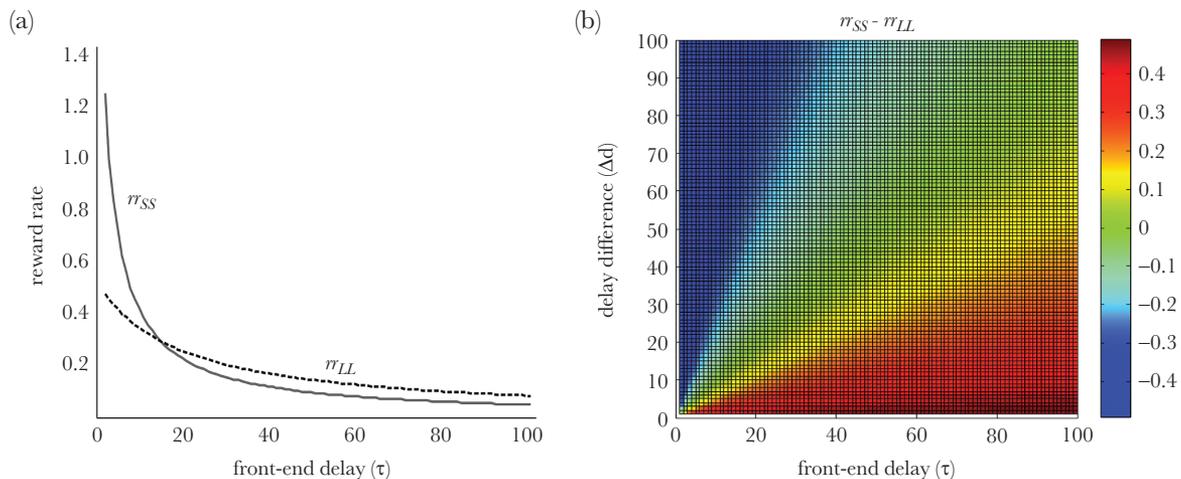
Figure 3

Rate maximization requires preference reversals. (A) Development of reward rates (rr) of a smaller, sooner and a larger, later reward with increasing front-end delay, for $rr_{SS} > rr_{LL}$ at $\tau = 0$. Reward rate decreases hyperbolically across front-end delays. Given the hyperbolic nature of the asymptotes, rr_{SS} and rr_{LL} cross over, implying optimal choice of smaller, sooner rewards left of the cross-over point, and larger, later rewards right of the cross-over point. (B) Heat plot indicating the difference in reward rate (rr_{SS} - rr_{LL}) at a range of delay differences and front-end delays, when the large to small reward ratio is 0.5. The heat plot indicates that the rate difference (in color) is determined by a linear relationship between front-end delay τ and delay difference Δd. For any delay difference Δd there is a front-end delay τ at which the rate difference rr_{SS} - rr_{LL} is 0.

its current food patch or leave the patch and move on to the next one—involves considering the travel time to the next patch. The travel time to the next patch is nothing else but a front-end delay, shifting the next foraging opportunities, in case of a leave decision, into the future by the travel time. Hence, the necessity for preference reversals, by consequence of the hyperbolic nature of reward rate decays (Figure 3), applies in a systematic way to animals making such stay-or-leave decisions.

In sum, we argue that evolution has favored hyperbolic over time-consistent (or other forms of) discounting because reward-rate in ecologically valid foraging scenarios decays hyperbolically (cf. Figure 3). An optimal choice algorithm maximizing long-term reward rate should track reward rate, and thus discount hyperbolically; in other words, hyperbolic discounting is adaptive. Our results inform students of human economic decision making about the putative ultimate reasons underlying hyperbolic, time-inconsistent discounting. But because of its intellectual roots in optimal foraging theory, our ideas also shed light on the adaptive value of hyperbolic discounting in foraging animals. We therefore believe that our findings are also of relevance for scholars of behavioral ecology of nonhuman animals, too.

Clearly, our reasoning of the optimality of preference reversals is not the only explanation of intertemporal choice. Alternative accounts have put the spotlight on animals' disregard of postreward delays, that is, delays between reward delivery and the onset of the next decision, such as intertrial intervals (Pearson *et al.* 2010). Postreward delays matter for energy-rate maximization in self-control tasks, as a change in postreward delay may result in a different option having the highest long-term energy rate (Stephens and Anderson 2001). Monkeys, for instance, have been found to disregard unsignalled postreward delays during intertemporal decisions, resulting in their failure to maximize reward rate unless the salience of those delays was particularly highlighted (Blanchard *et al.* 2013). Studies focusing on the (lack of) processing of postreward delays have made very valuable contributions to our understanding of temporal aspects during foraging. However, it is important to note that postreward delay accounts and our account of the optimality

of preference reversals are not mutually exclusive, but our account offers an addition to the existing literature. Moreover, we would like to stress once again that our reasoning and logic would equally apply to tasks incorporating variable postreward delays.

It is important to note that our results are in seemingly partial disagreement with previous findings. Notably, in contrast to earlier results (e.g., Schweighofer *et al.* 2006; Bixter and Luhman 2013; Zarr *et al.* 2014; Carter *et al.* 2015) we could not replicate a patch effect as participants maximized LTR more often in the self-control than the patch design, also reflected by higher earnings in the self-control condition compared to the patch condition. Carter and colleagues (2015) suggested that different cognitive mechanisms may underlie choices in the patch and self-control conditions, which could have led to the patch-effect. However, our results suggest otherwise: in both design conditions, the LTR maximization model provides the best fit with the data. Furthermore, the estimated hyperbolic discount rates (represented by the parameter k) in both design conditions were positively correlated, and they were correlated with LTR scores in both paradigms. This hints at similar, possibly identical cognitive mechanisms in all intertemporal choice contexts under consideration.

Why did we find evidence in favor of a single cognitive mechanism underlying choices in the patch and the self-control designs, while Carter and colleagues (2015) suggested different mechanisms? The main difference between the studies is the type of dependent variable: while Carter *et al.* computed model-predicted choices across a range of LTR values, we not only quantified the extent by which individuals maximized long-term reward rate by computing LTR scores, but we also measured participants' hyperbolic discount rates for both (patch and self-control) paradigms, as well as we directly compared the maximization and the hyperbolic models within and between paradigms. This allowed us to go beyond Carter and colleagues' (2015) analysis, and perform a conceptually different examination by directly comparing the performance of LTR and discounting models in the patch and self-control paradigms.

Importantly, another difference is that, in contrast to Carter and colleagues (2015), we used a full within-subject design: while, in our experiment, all participants experienced all task manipulations, Carter and colleagues (2015) randomly assigned participants to the different ITI-, short-, and long-delay conditions. Moreover, ITIs and delay-to-reward durations in Carter and colleagues (2015) were in the range of 5 s to 90 s, whereas in the present study experienced durations varied between 3 s and 31 s (see also Carter and Redish 2016). Intertemporal choice patterns are known to be strongly modulated by the range of delays and reward magnitudes used in a given task (Read 2001). Hence, the most parsimonious explanation for the discrepancy in results is that the inference of the cognitive mechanism underlying a revealed choice pattern depends on whether the data pool comprises observations from individuals who attend to the full set of parameter manipulations, or only subsets of it.

Of additional note, participants assigned to the patch condition in Carter and colleagues' (2015) study were explicitly told how to end a trial and go to the starting point—whereas participant in the present study needed to learn by experience when and how a trial ended. Moreover, they had to actually move in one of the experiments in order to proceed with the trial in one of the conditions—whereas in the present study participants only performed the task on a computer. These differences in the experimental settings make the 2 studies not fully comparable and might have likely affected participants' performance in the tasks. Future studies need to directly compare results from designs adopting different delay ranges and instruction procedures.

Finally, it is important to acknowledge some limitations of our theory. Clearly, rate maximization is a powerful idea, but it is not the only principle guiding decision making in human and nonhuman animals. For instance, rate maximization fails to predict behaviors when animals trade-off foraging opportunities with predation risk, it often cannot explain matching behavior or spontaneous alteration between choice options, and it makes unrealistic assumptions regarding near-omniscience (animals are informed about all pertinent information), and (lack of) memory constraints (see Herrnstein 1970; Stephens and Krebs 1986; Pierce and Ollason 1987; Sih and Christensen 2001; Kalenscher et al. 2003; Stephens et al. 2007; Stevens 2010). Hence, the ideas presented in this article are only a starting point for avenues for future research to uncover the reasons for hyperbolic, time-inconsistent decision making.

In summary, we found evidence that human choice behavior in a "classic" self-control task follows long-term reward rate maximization rules as well and even better than in a patch design task. Moreover, long-term reward rate maximization correlates with the degree of hyperbolic discounting in both paradigms. We argue that natural selection may have favored the evolution of a decision rule supporting maximization of long-term energy rate, but not economic utility, that allows preference reversals over timed outcomes because time-constant discounting would result in a systematic violation of rate-optimization principles. Crucially, while the time-inconsistent preference pattern produced by the underlying decision rule seemingly resembles hyperbolic discounting, our data support the idea that the currency maximized in intertemporal choice is long-term reward rate *through* hyperbolic reward discounting. It is perhaps noteworthy that, in contrast to previous literature, we did not find an improvement in long-term rate maximization by implementing a "patch" design, which could be due to procedural and analytical differences between our and previous studies, mainly regarding differences in the dependent measures as well as training and experience of participants. Further studies should focus on how reward rate maximization may be expressed in different intertemporal choice task designs as well as in different species. For example, a study design that allows for discounters with specific discount rates to reveal a patch-effect could explain why our results differ from the results of Carter et al. (2015).

SUPPLEMENTARY MATERIAL

Supplementary data are available at *Behavioral Ecology* online.

REFERENCES

Bateson M, Kacelnik A. 1996. Rate currencies and the foraging starling: the fallacy of the averages revisited. Behavioral Ecology. 7:341–352.

Benzion U, Rapoport A, Yagil J. 1989. Discount rates inferred from decisions - an experimental-study. Management Science. 35:270–284.

Bickel WK, Jarmolowicz DP, Mueller ET, Koffarnus MN, Gatchalian KM. 2012. Excessive discounting of delayed reinforcers as a trans-disease process contributing to addiction and other disease-related vulnerabilities: emerging evidence. Pharmacol Ther. 134:287–297.

Bixter MT, Luhmann CC. 2013. Adaptive intertemporal preferences in foraging-style environments. Front Neurosci. 7:93.

Blanchard TC, Pearson JM, Hayden BY. 2013. Postreward delays and systematic biases in measures of animal temporal discounting. Proceedings of the National Academy of Sciences. 110:15491–15496.

Carter EC, Pedersen EJ, McCullough ME. 2015. Reassessing intertemporal choice: human decision-making is more optimal in a foraging task than in a self-control task. Front Psychol. 6:95.

Carter EC, Redish AD. 2016. Rats value time differently on equivalent foraging and delay-discounting tasks. J Exp Psychol Gen. 145:1093–1101.

Fishburn PC, Rubinstein A, 1982. Time preference. International Economics Review. 23:677–694.

Frederick S, Loewenstein G, O ' Donoghue T. 2002. Time discounting and time preference: a critical review. Journal of Economic Literature. 40:351–401.

Green L, Fristoe N, Myerson J. 1994. Temporal discounting and preference reversals in choice between delayed outcomes. Psychon Bull Rev. 1:383–389.

Green L. Myerson J. 1996. Exponential versus hyperbolic discounting of delayed outcomes: risk and waiting time. American Zoologist. 36:496–505.

Hayden BY. 2016. Time discounting and time preference in animals: a critical review. Psychon Bull Rev. 23:39–53.

Herrnstein RJ. 1970. On the law of effect. J Exp Anal Behav. 13:243–266.

Kalenscher T, Diekamp B, Güntürkün O. 2003. Neural architecture of choice behaviour in a concurrent interval schedule. Eur J Neurosci. 18:2627–2637.

Kalenscher T, Pennartz CM. 2008. Is a bird in the hand worth two in the future? the neuroeconomics of intertemporal decision-making. Prog Neurobiol. 84:284–315.

Kalenscher T, Windmann S, Diekamp B, Rose J, Güntürkün O, Colombo M. 2005. Single units in the pigeon brain integrate reward amount and time-to-reward in an impulsive choice task. Curr Biol. 15:594–602.

Kirby KN. Herrnstein RJ. 1995. Preference reversals due to myopic discounting of delayed reward. Psychological Science. 6:83–89.

Kirby KN, Petry NM, Bickel WK. 1999. Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. J Exp Psychol Gen. 128:78–87.

Mazur JE. 1984. Tests of an equivalence rule for fixed and variable reinforcer delays. Journal of Experimental Psychology-Animal Behavior Processes. 10:426–436.

Mazur JE. 1987. An adjusting procedure for studying delayed reinforcement. In: Rachlin H, editor. Quantitative analyses of behavior: the effect of delay and of intervening events on reinforcement value. Hillsdale (NJ): Lawrence Erlbaum Associates. p. 19.

Mcdiarmid CG. Rilling ME. 1965. Reinforcement delay and reinforcement rate as determinants of schedule preference. Psychonomic Science. 2:195–196.

Namboodiri VM, Hussain Shuler MG. 2016. The hunt for the perfect discounting function and a reckoning of time perception. Curr Opin Neurobiol. 40:135–141.

Pearson JM, Hayden BY, Platt ML. 2010. Explicit information reduces discounting behavior in monkeys. Front Psychol. 1:237.

Pierce GJ, Ollason JG. 1987. Eight reasons why optimal foraging theory is a complete waste of time. Oikos. 49:111–117.

Pyke GH, Pulliam HR, Charnov EL. 1977. Optimal foraging - selective review of theory and tests. Quarterly Review of Biology. 52:137–154.

Read D. 2001. Is time-discounting hyperbolic or subadditive? J Risk Uncertainty. 23:5–32.

Rosati AG, Stevens JR. 2009. Rational decisions: the adaptive nature of context-dependent choice. In: Watanabe S, Blaisdell AP, Huber L, Young A, editors. Rational animals, irrational humans. Tokyo: Keio University Press. p. 101–117.

Rosati AG, Stevens JR, Hare B, Hauser MD. 2007. The evolutionary origins of human patience: temporal preferences in chimpanzees, bonobos, and human adults. Curr Biol. 17:1663–1668.

Samuelson PA. 1937. A note on measurement of utility. Rev Econ Stud. 4:7.

Schweighofer N, Shishida K, Han CE, Okamoto Y, Tanaka SC, Yamawaki S, Doya K. 2006. Humans can adopt optimal discounting strategy under real-time constraints. PLoS Comput Biol. 2:e152.

Seinstra MS, Sellitto M, Kalenscher T. 2017. Data from: rate maximization and hyperbolic discounting in human experiential intertemporal decision making. Dryad Digital Repository. http://dx.doi.org/10.5061/dryad.5040n

Sellitto M, Ciaramelli E, di Pellegrino G. 2011. The neurobiology of intertemporal choice: insight from imaging and lesion studies. Rev Neurosci. 22:565–574.

Sih A, Christensen B. 2001. Optimal diet theory: when does it work, and when and why does it fail? Anim Behav. 61:379–390.

Stephens DW. 2008. Decision ecology: foraging and the ecology of animal decision making. Cogn Affect Behav Neurosci. 8:475–484.

Stephens DW, Anderson D. 2001. The adaptive value of preference for immediacy: when shortsighted rules have farsighted consequences. Behav Ecol. 12:330–339.

Stephens D, Brown JS, Ydenberg RC. 2007. *Foraging: behavior and ecology*. Chicago: University of Chicago Press.

Stephens DW, Kerr B, Fernandez-Juricic E. 2004. Impulsiveness without discounting: the ecological rationality hypothesis. Proc Biol Sci. 271:2459–2465.

Stephens DW, Krebs JR. 1986. *Foraging theory. Monographs in behavior and ecology*. Princeton (NJ): University Press.

Stephens DW, McLinn CM. 2003. Choice and context: testing a simple short-term choice rule. Animal Behaviour. 66:59–70.

Stevens JR. 2010. Rational decision making in primates: the bounded and the ecological. Faculty Publications, Department of Psychology. 531:98–116.

Stevens JR, Stephens DW. 2010. The adaptive nature of impulsivity. Faculty Publications, Department of Psychology. 519:361–387.

Thaler R. 1981. Some empirical-evidence on dynamic inconsistency. Economics Letters. 8:201–207.

Zarr N, Alexander WH, Brown JW. 2014. Discounting of reward sequences: a test of competing formal models of hyperbolic discounting. Front Psychol. 5:178.