Postreward delays and systematic biases in measures of animal temporal discounting

Tommy C. Blanchard[a,b,1], John M. Pearson[c], and Benjamin Y. Hayden[a,b]

[a]Department of Brain and Cognitive Sciences and [b]Center for Visual Science, University of Rochester, Rochester, NY 14627; and [c]Department of Neurobiology, Duke University Medical School, Durham, NC 27710

Intertemporal choice tasks, which pit smaller/sooner rewards against larger/later ones, are frequently used to study time preferences and, by extension, impulsivity and self-control. When used in animals, many trials are strung together in sequence and an adjusting buffer is added after the smaller/sooner option to hold the total duration of each trial constant. Choices of the smaller/sooner option are not reward maximizing and so are taken to indicate that the animal is discounting future rewards. However, if animals fail to correctly factor in the duration of the postreward buffers, putative discounting behavior may instead reflect constrained reward maximization. Here, we report three results consistent with this discounting-free hypothesis. We find that (*i*) monkeys are insensitive to the association between the duration of postreward delays and their choices; (*ii*) they are sensitive to the length of postreward delays, although they greatly underestimate them; and (*iii*) increasing the salience of the postreward delay biases monkeys toward the larger/later option, reducing measured discounting rates. These results are incompatible with standard discounting-based accounts but are compatible with an alternative heuristic model. Our data suggest that measured intertemporal preferences in animals may not reflect impulsivity, or even mental discounting of future options, and that standard human and animal intertemporal choice tasks measure unrelated mental processes.

delay discounting | animal models | patience | foraging | rhesus macaque

Animal decision makers in natural environments regularly face choices between smaller rewards delivered sooner and larger rewards delivered later. To study how animals make this tradeoff, psychologists often measure preferences in intertemporal choice tasks, which directly pit smaller/sooner (SS) rewards against larger/later (LL) ones (1). Animals will typically sacrifice some of the long-term rate-maximizing benefits offered by the LL option to choose SS options. These present-biased preferences are sometimes thought to reflect impulsivity and poor self-control (2–4) and are often taken as a model for human impulsivity (2–6).

In nearly all animal intertemporal choice studies, many trials are strung together in sequence. When faced with a choice between an SS and an LL option, rate-maximizing behavior may sometimes dictate choosing the SS option to begin the next trial more quickly and thus increase the overall rate of reward (7). To avoid this confound, animal psychologists normally add an adjusting buffer after the SS reward to equalize trial lengths. As a result, total time for either option is matched, and the rate-maximizing strategy is to always choose the LL option. However, this buffering strategy serves its purpose only if animals correctly incorporate postreward delays into their decisions and if they correctly associate specific postreward delays with the choices that produced them. If animals fail to do either of these things, then their preferences cannot be interpreted as reflecting subjective discounting (7).

Despite the critical importance of the postreward buffer, a good deal of evidence suggests that animals may either misestimate its duration or fail to learn the contingency between choices and specific postreward buffers (8–11). First, animals are undersensitive to the duration of postreward buffers (12–16). Second, this insensitivity is consistent with the fact that events occurring after rewards are generally less well learned than events before rewards (cf. ref. 17). Indeed, we know of no study showing that any animals accurately estimate postreward buffers in a standard intertemporal choice paradigm. Third, explicitly cueing the duration of the postreward buffer reduces discount rates substantially, suggesting that monkeys may have difficulty estimating it when uncued (11). Finally, in foraging-like experimental contexts, animals often adopt a heuristic strategy that prioritizes the ratio of delay to reward (called expectation of rates) as an approximation for the long-term benefit they should optimize (called rate of expectations) (8). This heuristic ignores postreward delays. Given these observations, we hypothesized that a specific myopia for postreward delays, rather than mental discounting, drives animals' preferences in intertemporal choice tasks.

We performed three experiments using three variants of a standard intertemporal choice task to test this hypothesis (Fig. 1). We found (*i*) that monkeys are insensitive to the contingent nature of the postreward delay even after extensive training; (*ii*) that although they are sensitive to average changes in postreward delays, they systematically underestimate postreward buffer duration; and (*iii*) that making a postreward delay more salient, by offering a second reward at the end of it, decreases measured discount factors, suggesting that failures to learn about the structure of the postreward delays contribute to choices. We propose an alternative to the standard discounting equation. This alternative predicts preferences in intertemporal choice tasks as well as the standard hyperbolic discounting equation but involves no subjective discounting. Together these results support the idea that monkeys, and possibly other animals, are biased toward immediacy,

---

Significance

Psychologists often measure impulsivity and self-control in animals using the intertemporal choice task. This task pits a delayed reward against an immediate smaller one and is repeated several times. To ensure that animals do not choose the immediate reward to progress to the next trial more quickly, an adjusting postreward buffer is usually added after the reward to equalize trial lengths. Our study suggests that monkeys (and thus possibly other animals) do not understand this element of the task, and raises the possibility that the task does not, in fact, measure temporal discounting. We propose an alternative model, which includes an explicit bounded rationality term, that fits preferences as well as traditional hyperbolic discounting models.
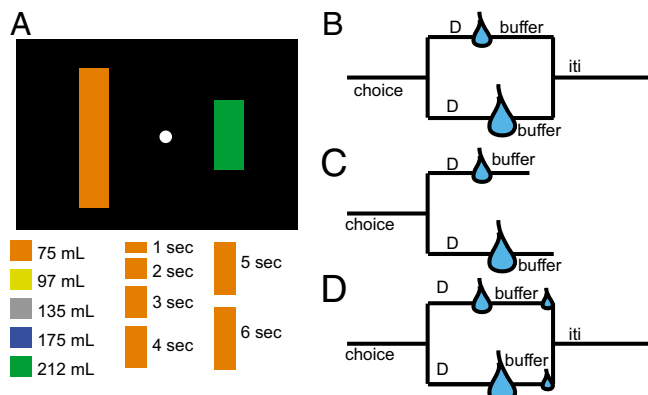

Fig. 1. Schematic of tasks used in this study. (*A*) Illustration of display. On each trial, two vertically oriented colored bars appeared to the left and right of a fixation point. The height of each bar indicates the delay until reward for that option. Both prereward delays were chosen randomly from a range of 0–6 s for all versions of the task, indicated by option height (*Lower*). Color indicates the reward size (*Lower*). Two bars were chosen randomly for each trial. (*B*) In the standard intertemporal choice task, postreward buffer duration is adjusted (from 0 to 6 s) so that total trial duration is equated regardless of choice. In the random buffer task, postreward buffer duration is chosen randomly from a uniform distribution (0–6 s). (*C*) In the constant buffer task, buffers are the same regardless of choice. (*D*) In the second-reward task, buffers are adjusted (as in the standard task), but a second, small reward is given at the end of the buffer to draw attention to the delay immediately before it. Tasks were run in different weeks to reduce possible interference between tasks.

not because of impulsivity or because they discount the future, but in large part because they have difficulty learning the structure of intertemporal choice tasks.

Results

Monkeys Fail to Associate Duration of Adjusting Buffer with Their Choices.
In interpreting results of standard intertemporal choice tasks, it is assumed that the causal link between animal's choices and the duration of postreward buffers is well learned. If at the time of testing subjects have failed to learn that their choice determines the duration of the postreward delay, then buffers will seem to be random. If buffers are random, then the long-term rate-maximizing strategy is to treat all postreward delays as their average value (across all choices) and to choose the SS option in some cases (especially when the reward difference is small and the delay difference is large). The reason for this is that choosing the SS can lead to another trial and its reward sooner. Thus, failure to learn the buffer structure can lead to SS choices, which appear to reflect discounting but are actually caused by a failure to learn task structure.

We hypothesize that monkeys are insensitive to the contingent nature of the postreward delay and instead treat variability of postreward durations as uncontrollable randomness in their environment. If this hypothesis is true, then preserving the distribution of postreward delays in the task, but decoupling them from choices, should have no effect on preferences. On the other hand, discounting approaches, which assume that monkeys fully incorporate postreward delays into their choices, would predict that randomizing the duration of the buffer would lead to more frequent choice of the SS option and thus an increased discount factor. (The reason for this is that, by balancing delays, a random buffer decreases opportunity cost associated with choosing SS.)

Following extensive training (at least 10,000 trials and several days of stable behavior) on a standard version of the intertemporal choice task, we tested three monkeys in a simple variant in which postreward delays were selected at random (0–6 s) on each trial and thus decoupled from monkeys' decisions. For all versions of

the task, reward size of an option was indicated by the color of the stimulus (orange, 75 μL; yellow, 97 μL; gray, 135 μL; blue, 175 μL; and green, 212 μL), and the height indicated the delay associated with choice of that target. Prereward delays for both options were chosen from a range of 0–6 s on all trials (Fig. 1*A*). We interleaved long (300 trials) blocks of this random buffer task with blocks of the standard intertemporal choice task (Fig. 1*B*). Block changes were clearly cued by a long pause (~30 s) and change in screen color. The overall distribution of postreward delays was the same as in the standard task; the only factor that varied was the contingent link between choice and buffer duration. We found no effect of condition (random vs. standard) on the likelihood of choosing the LL option (three-way ANOVA, day × session × subject); condition, df = 1, F = 0.64, P = 0.425, sum squares 0.14; session, df = 23, F = 0.86, P = 0.651, sum squares = 4.67; subject, df = 2, F = 228.43, P < 0.0001, sum squares = 103.57).

The standard measure of the relative preference for a LL and SS option is the discount factor, which gives the half-life of a reward per unit time assuming a hyperbolic discount function (3, 18). This number can serve as a standardized measure for decision maker's relative preferences between SS and LL options regardless of task details, even if the question of whether the animal psychologically discounts has not yet been determined. Based on the monkeys' choices, we used the standard hyperbolic discounting formula to compute the animals' discount factors in the standard and random buffer task (*SI Methods*, Figs. 2 and 3*A*).

We observed no effect of condition on performance in any of the three monkeys tested (Figs. 2 and 3*A*, P > 0.5 for all three subjects, bootstrap significance test; *SI Methods*). Specifically, we observed discount factors of 0.54 ± 0.051 (subject C), 1.14 ± 0.16 (subject H), and 0.22 ± 0.027 (subject J) in the standard condition and discount factors of 0.576 ± 0.68 (subject C), 1.24 ± 0.185 (subject H), and 0.24 ± 0.02 (subject J) in the random condition. These results are consistent with the idea that monkeys do not detect any difference in these two tasks, and thus that they treat the postreward adjusting buffer as a nonpredictable random delay.

Suppose that, counter to our hypothesis, measured discounting rates do reflect mental discounting of future rewards. In this case, then switching monkeys to the random task should increase their preference for the SS option beyond what it is in the standard task, because they can now move on to the next trial sooner in some cases. How much of an increase would we then expect in SS preferences? To answer this question, we next computed the discount factor that would be observed if the animal did discount but began choosing the SS option more often (optimizing-plus-discounting values). For all three monkeys, this optimizing-plus-discounting discount factor in the random condition was significantly greater than in the optimizing-plus-discounting one (k = 1.56, P = 0.002 for subject C; k = 1.79, P = 0.0090 for subject H; and k = 1.16, P < 0.0001 for subject J, bootstrap significance test).
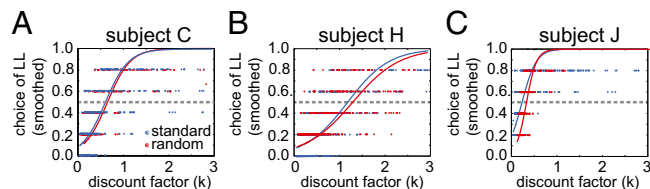
Fig. 2. Raw data. Plot of preference data (dots) and fit curves (lines) in the standard (blue) and random (red) versions of the task. Horizontal axis is the discount factor associated with preference indifference for each pair of options, and is defined as k = (v1 − v2)/(D1*v2 − D2*v1). Vertical axis indicates likelihood of choosing the LL option. Horizontal line indicates preference neutrality (*P*. = 0.5). *A*, *B*, and *C* refer to subjects C, H, and J, respectively.
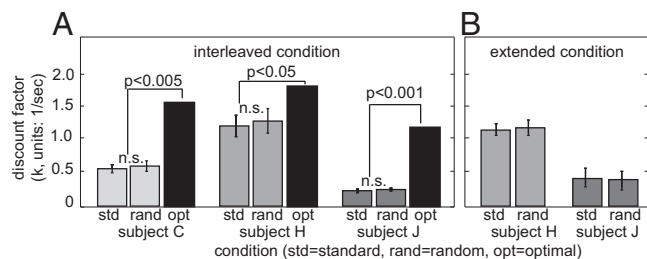
Fig. 3. Behavior of three monkeys in random buffer variant of the intertemporal choice task. (*A*) Behavior in interleaved trial set. Randomizing the length of the postreward buffer (rand) does not affect preferences, as indicated by best fit discount factor (k, *y*-axis), relative to standard task (std). Subjects that discount at the measured rates and yet fully account for the task structure should increase their choices of the SS option in the random buffer task, leading to increased measured k values (optimal, black bars); this pattern was not observed. Bars indicate SE. (*B*) Behavior in overtraining set (*Methods*).

It is possible that the random task somehow affected the monkeys' performance on the standard task and that, by interleaving the two tasks, we induced changes in the monkeys' preferences in the interleaved standard trials. To test for this possibility, we specifically examined the monkeys' preferences in the final thousand trials of training, after they had become familiar with the task and their preferences had presumably converged to a steady value. We found that these values were quite similar, and not significantly different from, values obtained on the standard task during the interleaved sessions. Specifically, we observed discount factors of 0.55 ± 0.082 (subject C), 1.28 ± 0.20 (subject H), and 0.22 ± 0.035 (subject J). These differences are not statistically different ($P > 0.5$ in all cases).

We also considered the possibility that the overtraining on the standard task made the monkeys adopt a habitual response strategy and thus unlikely to adjust preferences in the face of changes in payoffs. If monkeys are overtrained, their biases may appear only late in the blocks. To test this idea, we reexamined the data, but excluding the first half of each block, only examining the final 150 trials of each 300-trial block. We found no difference in the pattern of preferences. Specifically, we observed discount factors of 0.52 ± 0.10 (subject C), 1.08 ± 0.44 (subject H), and 0.23 ± 0.063 (subject J) in the standard condition and discount factors of 0.57 ± 0.99 (subject C), 1.2 ± 0.35 (subject H), and 0.19 ± 0.051 (subject J). This confirms that differences in behavior between the two cases did not appear later in the block.

To further test this idea, we performed an additional set of overtraining sessions in the random buffer task on two of the monkeys (subject H, 2,518 trials and subject J, 4,745 trials; Fig. 3*B*). We reasoned that this amount of overtraining would give the monkeys sufficient time to learn the new random contingencies. We analyzed only the last 1,000 trials, reasoning that these trials would be most likely to reflect stabilized behavior. Because these data were collected several months after the initial sessions, we also collected a new set of standard task data for these two monkeys ($n = 754$ for subject H and $n = 1,526$ for subject J). In this control study, we again found no effect of condition on preference. In the random task, we found values of $k = 1.16 \pm 0.10$ (subject H) and $k = 0.43 \pm 0.14$ (subject J). In the standard task, we found values of $k = 1.12 \pm 0.09$ (subject H) and $k = 0.42 \pm 0.15$ (subject J).
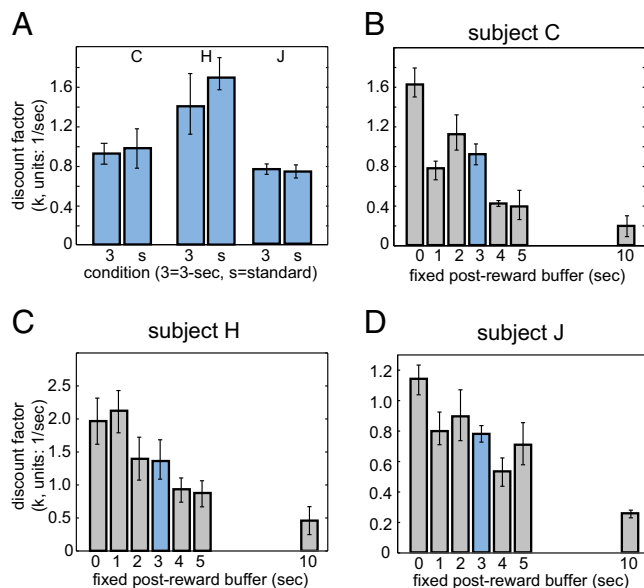
Preferences in the Constant Buffer Version of the Intertemporal Choice Task. To investigate whether monkeys incorporate any elements of postreward buffers into their decisions, we next measured preferences in a constant buffer intertemporal choice task.

Postreward delays were fixed to a constant duration regardless of choice and varied in long (1,000 trials) blocks (Fig. 1*C*). We used seven buffer lengths—0, 1, 2, 3, 4, 5, and 10 s (testing order was randomized by monkey; *Methods*). We also collected data from the standard task in interleaved sessions.

Let us first consider the 3-s postreward buffer (Fig. 4*A*). This condition is interesting because 3 s is the average duration of all buffers in both the standard and the random buffer variants. Thus, the 3-s condition lets us examine whether the variability in the standard task somehow affects preferences. It does not. We found that for all three monkeys, observed discount factors were not significantly different for the 3-s constant buffer and the standard task (the *P* value was >0.4 in all cases; Fig. 4*A*).

What about other buffer lengths (Fig. 4 *B–D*)? Monkeys should choose LL options more often when postreward delays grow longer (this is true regardless of whether they discount delayed rewards). We find that all three monkeys do this. Specifically, condition and subject exert a significant effect on preference rates; session does not (three-way ANOVA, day × session × subject; condition, df = 7, $F = 32.73$, $P < 0.0001$, sum squares 54.9; session, df = 41, $F = 0.82$, $P = 0.785$, sum squares 8.1; subject, df = 2, $F = 447.58$, $P < 0.0001$, sum squares 196.4).

We next computed the best fit discount factor for each of the conditions (Fig. 4 *B–D*). The correlation between buffer length and discount factors is significant for all three monkeys (Spearman rank correlation test; subject C, $r = -0.8929$, $P = 0.0123$; subject H, $r = -0.9643$, $P = 0.0028$; subject J, $r = -0.9286$, $P = 0.0067$). Note that the correlation is not driven by the 10-s delays, as it remains significant even if one removes the 10-s delays from the analysis in two of the subjects, and is close to significant in the third (Spearman rank correlation test; subject C, $r = -0.8296$, $P = 0.0582$; subject H, $r = -0.9429$, $P = 0.0167$; subject J, $r = -0.8857$, $P = 0.0333$). These results indicate that monkeys are not blind to the duration of postreward delays.

Fig. 4. Preferences depend on duration of buffer in the constant buffer variant of the standard intertemporal choice task. Each plot shows the relationship between measured discount factor (k) and duration of postreward buffer. Data for the 3-s buffer is highlighted (blue) because the expected postreward delay for the 3-s buffer trials matches that for the standard version of the task and for the random buffer version of the task. (*A*) Data for 3-s condition. (*B–D*) Data for subjects C, H, and J, respectively, for all conditions.

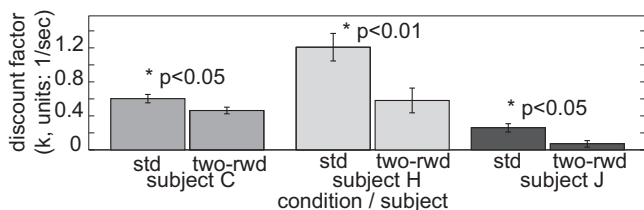**Increasing Salience of Postreward Delay Reduces Discount Factors.**
One reason why monkeys may have difficulty detecting some aspects of postreward delays is that such delays are difficult to learn (11, 19). Previous studies have found that drawing attention to the LL cue (although not the delay specifically) can increase choice of the LL option at the expense of the SS (20–22). Supporting this idea, we have previously shown that explicitly cueing the length of the postreward delays reduces measured discounting substantially (11). However, these earlier studies, including our own, did not explain why cueing the LL option affected preferences. Experiment 3 in the present study suggests that making the delay itself more salient is sufficient to bias preferences. This finding suggests that the subject's response to the delay itself is the key factor in biasing preferences to the SS option in normal conditions. To draw monkeys' attention to the delay, we measured preferences in a second-reward variant of the standard intertemporal choice task. In this version, a small second reward was given at the end of the adjusting buffer regardless of choice. The size of this second reward was set to the smallest amount used in the standard task (75 μL of water).

All three monkeys showed a greater propensity to choose the LL option in this task variant. Subject C went from choosing LL 61.4% of the time to 74.4% of the time; these numbers for subject H were 56.78–64.42%, and for subject J 64.5–71.3%. Note that all three monkeys showed a default bias toward LL; this is simply an artifact of the values we chose for our task. However, it also means that for all monkeys, the second reward biased choices away from ambivalence, suggesting that the second reward does not affect preference by making the two options more difficult to discriminate. We ran a multifactorial ANOVA on the raw preference data. We found a significant effect of condition (df = 1, $F = 327.37$, $P < 0.0001$, sum squares 100.13) and of subject (df = 2, $F = 25.21$, $P = 0.0009$, sum squares 17.7) but not of session (df = 4, $F = 1.36$, $P = 0.2849$, sum squares 1.61). All three monkeys showed lower discount factors in the second reward version of the task (Fig. 5). Specifically, k went from 0.61 to 0.46, $P = 0.039$ for subject C; from 1.21 to 0.58, $P = 0.0084$ for subject H; and from 0.26 to 0.07, $P = 0.0012$ for subject J.

A Discounting-Free Model. We hypothesize that monkeys' preferences reflect a desire to maximize (nondiscounted) long-term reward rate, but that they are limited in their ability to learn about postreward delays in intertemporal choice tasks. Here we formalize this idea by providing a specific model for how monkeys evaluate delayed options:

$$v_{measured} = \frac{v_{actual}}{D + \omega}. \qquad [1]$$

The opportunity cost imposed by a delay is the same regardless of whether it comes before or after the reward. An option's presumed rate is given by its value (v_{actual}) divided by the sum of the pre- (D) and the presumed postreward (ω) delays. The postreward delay term in the equation is not the actual delay, but the monkey's biased estimate of the delay.

The critical innovation of our model is that it does not have a discount factor and does not suppose that variability comes from specific valuation processes as it does in discounting models. Instead, variability comes from the monkeys' biased estimate of postreward delay, which is not related to valuation and does not reflect value discounting of future rewards. Thus, unlike in the standard model (Eq. 2), we use $v_{measured}$ instead of $v_{discounted}$:

$$v_{discounted} = \frac{v_{actual}}{1 + k * D}. \qquad [2]$$

Both models involve only one free variable, and are mathematically translatable, aside from a rescaling factor. Thus, they must fit the data equally well despite the fact that they have entirely different implications for the underlying mental computations. In other words, empirically fit data from the standard task are consistent with both standard delay discounting and our proposed model, whereas the results of our manipulations in experiments 1–3 argue against the standard delay discounting model and for our alternative.
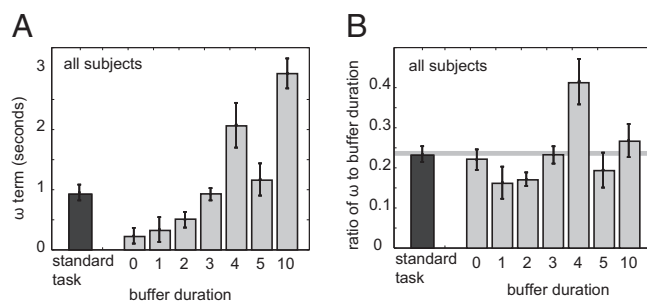
One convenient feature of our model is that the fit ω parameter provides a direct measure of the monkeys' subjective estimate of postreward delay. Fig. 6A shows the model fit values and thus the monkeys' estimate of postreward delay (i.e., ω from the model). This value rose with the postreward delay but across all conditions was consistently lower than the true postreward delay. These findings suggest that monkeys systematically underestimate postreward delay. Fig. 6B plots the ratio of ω to actual buffer duration for the standard condition and each of the constant buffer conditions for all three subjects combined. (Note, to compute this ratio, we have added in the duration of the intertrial interval, 1 s, to avoid dividing by zero.)

In the standard intertemporal choice task, this ratio was 0.242 for the group (for subject C, it was 0.268; subject H, 0.199; subject J, 0.243). Similar ratios were observed for each of the constant buffer conditions. Indeed, the ratio of ω to actual delay did not differ significantly from the value for the constant 3-s buffer (which was 0.248 for the group; subject C, 0.2688; subject H, 0.1438; J, 0.380, $P > 0.5$ for the group and for each of the three subjects). These results suggest that monkeys substantially underestimate postreward delays, and do so by a roughly constant ratio (about 0.25). We have no hypothesis about why this particular value is observed.

Discussion

Here we describe the results of three experiments examining preferences of monkeys in variants of the standard intertemporal choice task. We found (i) that monkeys are insensitive to the contingent nature of the postreward delay, (ii) that although they are sensitive to average changes in postreward delays, they systematically underestimate postreward buffer duration, and (iii) that adding a second reward at the end of the postreward delay decreases measured discount factors, suggesting that failures to learn about the postreward delays contribute to preferences. We present a model that predicts data as well as discounting models while remaining psychologically plausible and better accommodating discrepancies observed in our task variants.

This is not the first study to challenge the standard discounting approach to explaining intertemporal choice (7, 8, 19, 23, 24). Stephens and coworkers have shown that bluejay preferences in an intertemporal choice task do not predict behavior in more naturalistic contexts, such as patch-leaving tasks, even when they are mathematically equivalent (7, 9, 10, 25). Their results indicate that intertemporal choice tasks may lack external validity; failure to learn task details is one possible cause for this. Steep

Fig. 5. Enhancing salience of postreward delay decreases observed discounting. In the second-reward variant of the standard intertemporal choice task, a second squirt of juice was given at the end of the postreward buffer, regardless of choice. For both monkeys tested, discount factors declined significantly.

Fig. 6. Subjective estimates of postreward buffer for all subjects. (*A*) Subjective estimate of postreward delay, as estimated by Eq. **1**, for all three subjects combined. Subjects' estimates of postreward buffer rose with actual value but was systematically lower than the actual value. (*B*) Ratio of subjects' estimate of postreward delay to actual postreward delay, average for all three subjects. Ratio was roughly constant. Average value (0.24) is shown with horizontal line.

discounting is also inconsistent with generally patient preferences observed in delay of gratification tasks in animals (26–29). Thus, the present results are part of an emerging group of studies challenging the psychological reality of discounting as measured by intertemporal choice tasks. Although our model is not the first to include subjective time estimates to explain discounting behavior, previous models only include estimates of pre- and not postreward delays (30). More broadly, our results are consistent with the idea that animals, like humans, use "fast and frugal" mental shortcuts that provide approximately optimal choices in most but not all circumstances (31–33). The present study extends these earlier studies by identifying and quantifying one specific bias, and showing that it produces a specific pattern of behavior in the intertemporal choice task.

Although many of the ideas expressed here have been raised before (7, 8, 10, 24), the present study makes four unique contributions to the existing challenges to intertemporal choice. First, our random buffer experiment shows that animals fail to associate postreward delays with their choice. Second, our constant buffer experiment shows that choices do depend on postreward delays, but are systematically underestimated. Third, our second-reward experiment shows that cueing and salience affect discount factors greatly—consistent with our theory, but inconsistent with discounting (cf. ref. 11). Finally, our quantitative model predicts behavior as well as hyperbolic discounting without requiring any mental discounting, and allows us to measure putative subjective and inaccurate estimates of postreward delay.

Because our model emphasizes bounds to learning as a source of apparent discounting, we refer to it as a "bounded rationality model" of intertemporal choice. Bounded rationality models offer numerous advantages over the standard discounting models of choice in animals. Most importantly, they explain (not simply describe) the hyperbolic shape of the discount curve observed in many studies. Opportunity cost is linearly proportional to delay in intertemporal choices, and so an item's value is inversely proportional to delay (34). Thus, the normative nondiscounting value of a delayed option is a reciprocal (i.e., hyperbolic) function of the sum of delays. Second, they provide an explanation for the quantitative incongruity between human and animal discounting factors by arguing that they arise from unrelated psychological mechanisms, discounting and bounded rationality, respectively (7, 8, 19). Third, they provide an explanation for the frustrating lability of discount factors across conditions and studies. Our model holds that discount factors reflect variations in attention and/or learning, which are themselves quite labile, rather than in valuation functions, which ought to be stable. Finally, they are not subject to the important criticism that steep discounting should be evolu-

tionarily disfavored because it produces such suboptimal behavior, and is inconsistent with caching and other future-oriented behaviors in animals (7, 35).

Previous studies using intertemporal choice tasks in animals have sometimes come to conflicting conclusions. For example, orbitofrontal lesions lesions have been found to decrease impulsive choice in some studies (e.g., ref. 36) and increase impulsive choice in others (e.g., ref. 37). One previous study has shown that the effects of these lesions depend on whether prereward delays are cued (38). Our results extend this idea by suggesting that differences in discounting rates across studies may be influenced by differences in postreward delay length or salience.

The present criticisms do not apply to data from standard human intertemporal choice tasks, because they do not use postreward delays. Indeed, as we note, our results suggest that animal and human preferences in analogous tasks may reflect unrelated psychological processes. However, they do suggest reinterpretation of human studies in which multiple trials are strung together and adjusting postreward delays are used, including one of our own (39–43). It is possible that serial and one-shot human discounting paradigms may reflect unrelated mental processes. Consistent with this idea, other groups have proposed that miscalculation may underlie steep discounting observed in these kinds of studies (44, 45).

Despite its importance, we lack a single clear definition of impulsivity (46). Much evidence suggests that it may comprise multiple distinct elements. The closely related concept of self-control is equally important and equally nebulous. Intertemporal choice tasks frequently serve as a measure for impulsivity or self-control animal models of human self-control, including for diseases associated with poor self-control, such as addiction, obsessive compulsive disorder, and attention deficit hyperactivity disorder (ADHD) (47–49). Our results are consistent with the idea that monkeys are strict rate maximizers subject to two constraints, underestimation of postreward delays and failure to associate postreward delays with specific choices. Neither of these would seem to fit into any of the definitions of poor self-control or impulsivity. More broadly, these results serve as a reminder that animals do not necessarily solve tasks in the way in which we expect. Understanding animal behavior requires a critical attitude and benefits from applying methods of foraging theory and reinforcement learning.

Methods

All procedures were approved by the University of Rochester Institutional Animal Care and Use Committee and were designed and conducted in compliance with the Public Health Service's Guide for the Care and Use of Animals. Three male rhesus monkeys (*Macaca mulatta*) served as subjects. We used Psychtoolbox for behavioral tasks (50) and Eyelink Toolbox for gaze measurement (51).

On each trial of the task, a fixation spot first appeared in the center of the monitor. Once the subject acquired fixation and maintained it for 100 ms, two targets appeared, 275 pixels to the left and right of the central spot. Both stimuli were vertically oriented rectangles (80 pixels wide, 1–300 pixels tall). Stimuli were one of five colors (orange, yellow, gray, blue, and green). These colors indicated the size of the reward that would be obtained for the choice of this target (orange, 75; yellow, 97; gray, 135; blue, 175; and green, 212 μL). The height indicated the delay associated with choice of that target (0–6 s). The subject then selected a target by shifting the gaze toward it. Following the choice, the cued delay occurred, during which the bar shrunk at a constant rate (50 pixels per second). This shrinking thus provided a reminder of the linkage between cue length and delay. Once a choice was made, subjects could move their eyes freely without penalty. Once the cue shrunk entirely, the reward was given and the postreward buffer began. The screen was blank throughout the postreward buffer and the intertrial interval. Once this postreward buffer ended, an intertrial interval (1 s in all cases) began. No cue signaled the change from buffer period to intertrial interval time.

In the standard intertemporal choice task, postreward delays were buffered so that total trial length was equal regardless of choice. Because prereward delays varied randomly from 0 to 6 s, postreward delays varied, in a wholly conjugate manner, from 0 to 6 s as well. We collected at least 10,000 trials of the standard intertemporal choice task before moving on to any variants. All variants were tested in the same order for all monkeys.

In the constant buffer task, postreward delays were the same, regardless of the monkey's choice. Postreward buffers were either 0, 1, 2, 3, 4, 5, or 10 s and were run in blocks of the same type so as to give the monkey the opportunity to learn the postreward delay. In all other aspects the task was the same as the standard intertemporal choice task. At least 1,000 trials were collected per condition over multiple sessions. Trials were not interleaved with other tasks. Order of collection for specific postreward buffers was randomized separately for each monkey to reduce any possible order effects. Constant buffer data were compared with standard intertemporal choice data collected immediately beforehand to reduce possible long-term learning effects.

In the random buffer task, postreward delays were chosen randomly from 0 to 6 s, without regard to the monkey's choice. In all other aspects the task was the same as the standard intertemporal choice task. The random buffer was run in conjunction with the standard intertemporal choice task in interleaved blocks of 300–500 trials, on multiple days. At least 3,000 trials total were collected in each of the two (random and standard) conditions. For analysis purposes, random buffer data were compared with standard intertemporal choice data collected in the same behavioral sessions to mitigate any possible variations in preferences. Blocks were signaled to the subject in the following way. Following the completion of a block, the monitor went blank and returned to the default task screen. Then the next session started. This entire process took about 30 s and validly predicted a change in task rules.

In the two-reward task, a second small reward (75 µL) was given at the end of the buffer period, before the 1-s intertrial interval. In all other aspects the task was the same as the standard intertemporal choice task.

1. Stevens JR, Stephens DN (2010) The adaptive nature of impulsivity. *Impulsivity: The Behavioral and Neurological Science of Discounting*, eds Madden GJ, Bickel WK (American Psychological Association, Washington, DC).
2. Ainslie G (1975) Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychol Bull* 82(4):463–496.
3. Rachlin H (2000) *The Science of Self-Control* (Harvard Univ Press, Cambridge, MA).
4. Rachlin H, Green L (1972) Commitment, choice and self-control. *J Exp Anal Behav* 17(1):15–22.
5. Tobin H, Logue AW (1994) Self-control across species (Columba livia, Homo sapiens, and Rattus norvegicus). *J Comp Psychol* 108(2):126–133.
6. Stevens JR, Stephens DW (2008) Patience. *Curr Biol* 18(1):R11–R12.
7. Stephens DW, Anderson D (2001) The adaptive value of preference for immediacy: When shortsighted rules have farsighted consequences. *Behav Ecol* 12(3):330–339.
8. Bateson M, Kacelnik A (1996) Rate currencies and the foraging starling: The fallacy of the averages revisited. *Behav Ecol* 7(3):341–352.
9. Stephens DW (2002) Discrimination, discounting and impulsivity: A role for an informational constraint. *Philos Trans R Soc Lond B Biol Sci* 357(1427):1527–1537.
10. Stephens DW, Kerr B, Fernández-Juricic E (2004) Impulsiveness without discounting: The ecological rationality hypothesis. *Proc Biol Sci* 271(1556):2459–2465.
11. Pearson JM, Hayden BY, Platt ML (2010) Explicit information reduces discounting behavior in monkeys. *Front Psychol* 1:237.
12. Snyderman M (1983) Delay and amount of reward in a concurrent chain. *J Exp Anal Behav* 39(3):437–447.
13. Logue AW, Smith ME, Rachlin H (1985) Sensitivity of pigeons to pre-reinforcer and postreinforcer delays. *Anim Learn Behav* 13:181–186.
14. Goldshmidt JN, Lattal KM, Fantino E (1998) Context effects on choice. *J Exp Anal Behav* 70(3):301–320.
15. Mazur JE (1989) Theories of probabilistic reinforcement. *J Exp Anal Behav* 51(1):87–99.
16. Mazur JE, Romano A (1992) Choice with delayed and probabilistic reinforcers: Effects of variability, time between trials, and conditioned reinforcers. *J Exp Anal Behav* 58(3):513–525.
17. Mackintosh NJ (1975) A theory of attention: Variations in the associablity of stimuli with reinforcement. *Psychol Rev* 82:276–298.
18. Mazur JE (1987) An adjusting procedure for studying delayed reinforcement. *Quantitative Analyses of Behavior, Vol 5. The Effect of Delay and Intervening Events on Reinforcement Value*, eds Commons ML, Mazur JE, Nevin JA, Rachlin H (Erlbaum, Mahwah, NJ).
19. Kacelnik A (2003) The evolution of patience. *Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice*, eds Loewenstein G, Read D, Baumeister RF. (Russell Sage Foundation, New York).
20. Siegel E, Rachlin H (1995) Soft commitment: Self-control achieved by response persistence. *J Exp Anal Behav* 64(2):117–128.
21. Monterosso J, Ainslie G (1999) Beyond discounting: Possible experimental models of impulse control. *Psychopharmacology (Berl)* 146(4):339–347.
22. Calvert AL, Green L, Myerson J (2011) Discounting in pigeons when the choice is between two delayed rewards: Implications for species comparisons. *Front Neurosci* 5:96.
23. Fawcett TW, McNamara JM, Houston AI (2012) When is it adaptive to be patient? A general framework for evaluating delayed rewards. *Behav Processes* 89(2):128–136.
24. Pavlic TP, Passino KM (2010) When rate maximization is impulsive. *Behav Ecol Sociobiol* 64:1255–1265.
25. Stephens DW, McLinn CM (2003) Choice and context: Testing a simple short-term choice rule. *Anim Behav* 66:59–70.
26. Reynolds B, Schiffbauer R (2005) Delay of gratification and delay discounting: A unifying feedback model of delay-related impulsive behavior. *Psychol Rec* 55:439–460.
27. Reynolds B, de Wit H, Richards JB (2002) Delay of gratification and delay discounting in rats. *Behav Processes* 59(3):157.
28. Evans TA, Beran MJ (2007) Delay of gratification and delay maintenance by rhesus macaques (Macaca mulatta). *J Gen Psychol* 134(2):199–216.
29. Bramlett JL, Perdue BM, Evans TA, Beran MJ (2012) Capuchin monkeys (Cebus apella) let lesser rewards pass them by to get better rewards. *Anim Cogn* 15(5):963–969.
30. Namboodiri VMK, Mihalas S, Marton T, Shuler MGH (2013) *A General Theory of Intertemporal Decision-Making and the Perception of Time*. Available at http://arxiv.org/abs/1306.1200. Accessed July 19, 2013.
31. Simon H (1955) A behavioral model of rational choice. *Q J Econ* 69:99–118.
32. Gigerenzer G, Selten R (2002) Bounded rationality: The adaptive toolbox. *The Adaptive Toolbox*, eds Gigerenzer G, Selten R (MIT Press, Cambridge, MA).
33. Gigerenzer G, Todd PM, Group A (1999) *Simple Heuristics That Make Us Smart* (Oxford Univ Press, Oxford).
34. Stephens DW, Krebs JR (1986) *Foraging Theory* (Princeton Univ Press, Princeton, NJ).
35. Raby CR, Alexis DM, Dickinson A, Clayton NS (2007) Planning for the future by western scrub-jays. *Nature* 445(7130):919–921.
36. Winstanley CA, Theobald DEH, Cardinal RN, Robbins TW (2004) Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *J Neurosci* 24(20): 4718–4722.
37. Rudebeck PH, Walton ME, Smyth AN, Bannerman DM, Rushworth MFS (2006) Separate neural pathways process different decision costs. *Nat Neurosci* 9(9):1161–1168.
38. Zeeb FD, Floresco SB, Winstanley CA (2010) Contributions of the orbitofrontal cortex to impulsive choice: Interactions with basal levels of impulsivity, dopamine signalling, and reward-related cues. *Psychopharmacology (Berl)* 211(1):87–98.
39. McClure SM, Ericson KM, Laibson DI, Loewenstein G, Cohen JD (2007) Time discounting for primary rewards. *J Neurosci* 27(21):5796–5804.
40. Luhmann CC, Chun MM, Yi DJ, Lee D, Wang XJ (2008) Neural dissociation of delay and uncertainty in intertemporal choice. *J Neurosci* 28(53):14459–14466.
41. Jimura K, Myerson J, Hilgard J, Braver TS, Green L (2009) Are people really more patient than other animals? Evidence from human discounting of real liquid rewards. *Psychon Bull Rev* 16(6):1071–1075.
42. Smith DV, et al. (2010) Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *J Neurosci* 30(7):2490–2495.
43. Hayden BY, Parikh PC, Deaner RO, Platt ML (2007) Economic principles motivating social attention in humans. *Proc Biol Sci* 274(1619):1751–1756.
44. Sonuga-Barke EJ, Lea SEG, Webley P (1989) An account of human "impulsivity" on self-control tasks. *Q J Exp Psychol B* 41:161–179.
45. Schweighofer N, et al. (2006) Humans can adopt optimal discounting strategy under real-time constraints. *PLOS Comput Biol* 2(11):e152.
46. Evenden JL (1999) Varieties of impulsivity. *Psychopharmacology (Berl)* 146(4):348–361.
47. Winstanley CA, Eagle DM, Robbins TW (2006) Behavioral models of impulsivity in relation to ADHD: Translation between clinical and preclinical studies. *Clin Psychol Rev* 26(4):379–395.
48. Dalley JW, Everitt BJ, Robbins TW (2011) Impulsivity, compulsivity, and top-down cognitive control. *Neuron* 69(4):680–694.
49. Dalley JW, Mar AC, Economidou D, Robbins TW (2008) Neurobehavioral mechanisms of impulsivity: Fronto-striatal systems and functional neurochemistry. *Pharmacol Biochem Behav* 90(2):250–260.
50. Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10(4):433–436.
51. Cornelissen FW, Peters EM, Palmer J (2002) The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behav Res Methods Instrum Comput* 34(4): 613–617.