

Rats exhibit similar biases in foraging and intertemporal choice tasks

Gary A. Kane^{1,2*}, Aaron M. Bornstein^{1,3}, Amitai Shenhav⁴, Robert C. Wilson⁵, Nathaniel D. Daw¹, Jonathan D. Cohen¹

*For correspondence:

gkane@rowland.harvard.edu (GAK)

¹Department of Psychology and Princeton Neuroscience Institute, Princeton University; ²The Rowland Institute at Harvard, Harvard University; ³Department of Cognitive Sciences and Center for the Neurobiology of Learning and Memory, University of California Irvine; ⁴Department of Cognitive, Linguistic, and Psychological Sciences and Carney Institute for Brain Science, Brown University; ⁵Department of Psychology and Cognitive Science Program, University of Arizona

Abstract Animals, including humans, consistently exhibit myopia in two different contexts: foraging, in which they harvest locally beyond what is predicted by optimal foraging theory, and intertemporal choice, in which they exhibit a preference for immediate vs. delayed rewards beyond what is predicted by rational (exponential) discounting. Despite the similarity in behavior between these two contexts, previous efforts to reconcile these observations in terms of a consistent pattern of time preferences have failed. Here, via extensive behavioral testing and quantitative modeling, we show that rats exhibit similar time preferences in both contexts: they prefer immediate vs. delayed rewards and they are sensitive to opportunity costs of delays to future decisions. Further, a quasi-hyperbolic discounting model, a form of hyperbolic discounting with separate components for short- and long-term rewards, explains individual rats' time preferences across both contexts, providing evidence for a common mechanism for myopic behavior in foraging and intertemporal choice.

Introduction

Serial stay-or-search problems are ubiquitous across many domains, including employment, internet search, mate search, and animal foraging. For instance, in patch foraging problems, animals must choose between an immediately available opportunity for reward or the pursuit of potentially better but more distal opportunities. It is typically assumed that animals seek to maximize the long-term average reward (net of cost) rate, as a proxy for reproductive fitness. The optimal behavior for maximizing this currency in foraging tasks, described by the Marginal Value Theorem (MVT; **Charnov 1976**), is to choose the immediately available opportunity if it provides a reward rate greater than the average reward rate across all alternative options, which includes the costs of accessing those options. Animals tend to follow the basic predictions of long-term reward maximization: they are generally more likely to pursue opportunities for larger vs. smaller rewards and, if the cost of searching for alternatives is greater, they are more likely to pursue opportunities for smaller rewards (**Stephens and Krebs, 1986; Constantino and Daw, 2015; Hayden et al., 2011; Kane et al., 2017**).

Although animal behavior follows the basic predictions of optimal foraging behavior described

41 by MVT, in the majority of studies across a variety of species, including humans, non-human
 42 primates, and rodents, animals exhibit a consistent bias towards pursuing immediately available
 43 rewards relative to predictions of MVT, often referred to as "overharvesting" (*Constantino and*
Daw, 2015; Hayden et al., 2011; Kane et al., 2017; Nonacs, 2001; Kolling et al., 2012; Shenhav
et al., 2014; Wikenheiser et al., 2013; Carter and Redish, 2016). Prior studies have proposed two
 44 explanations for overharvesting: subjective costs, such as an aversion to rejecting an immediately
 45 available reward (*Wikenheiser et al., 2013; Carter and Redish, 2016*); and nonlinear reward utility or
 46 diminishing returns, by which larger rewards are not perceived as proportionally larger than smaller
 47 rewards (*Constantino and Daw, 2015*). But these hypotheses have never been systematically
 48 compared in a set of experiments designed to directly test their predictions. Furthermore, according
 49 to these rate-maximizing hypotheses, the perceived value of rewards does not differ between
 50 situations in which the delays occur before or after reward is received. In this respect, the predictions
 51 made by these hypotheses (which are still grounded in a core assumption that animals attempt to
 52 maximize the long-term reward rate) are not compatible with an otherwise seemingly similar bias
 53 that is widely observed in standard intertemporal choice tasks (also referred to as delay discounting
 54 or self-control tasks): a preference for smaller, more immediate rewards over larger, delayed
 55 rewards (*Ainslie, 1992; Kirby, 1997*).

56 The preference for more immediate rewards in intertemporal choice tasks is commonly ex-
 57 plained in one of two ways, both assuming that animals choose as though they were optimizing a
 58 different currency than long-term reward rate: temporal discounting or short-term rate maximiza-
 59 tion. According to temporal discounting, the perceived value of a future reward is discounted by the
 60 time until its receipt. Temporal discounting can arise even when maximizing the long-term reward
 61 rate, for certain environments. In particular, discounting can be adaptive in unstable environments
 62 — if the environment is likely to change before future rewards can be acquired, it is appropriate to
 63 place greater value on more predictable rewards available in the near future. Under this hypothesis,
 64 and the further assumption that expected rewards disappear at a constant rate, a long-term reward
 65 rate maximizer would discount rewards exponentially in their delay (*Gallistel and Gibbon, 2000*;
Kacelnik and Todd, 1992). However, animal preferences typically follow a hyperbolic-like form: the
 66 rate of discounting is steeper initially and decreases over time (*Gallistel and Gibbon, 2000; Kacelnik*
and Todd, 1992; Thaler, 1981). This yields inconsistent time preferences or preference reversals: an
 67 animal may prefer to wait longer for a larger reward if both options are distant, but will change their
 68 mind and prefer the smaller reward as the time to both options draws near (*Ainslie, 1992; Kirby,*
1997; Gallistel and Gibbon, 2000; Kacelnik and Todd, 1992). Recent theoretical work has shown
 69 that hyperbolic time preferences may arise from imperfect foresight — if the variance in predicting
 70 the timing of future outcomes increases with the delay to the outcome, a long-term reward rate
 71 maximizer would exhibit hyperbolic time preferences (*Gabaix and Laibson, 2017*). Similarly, short-
 72 term maximization rules predict that animals seek to maximize reward over shorter time horizons;
 73 this may also be motivated as an approximation to long-term reward maximization as it may be dif-
 74 ficult to accurately predict all future rewards (*Stephens, 2002; Stephens et al., 2004*). Along similar
 75 lines, *Namboodiri et al. (2014)* argue that, rather than maximizing long-term reward rate into the
 76 future, animals may select options that maximize reward rate up to the current point in time, or
 77 due to environmental factors (e.g. non-stationarity) or biological constraints (e.g. computational
 78 constraints), over a finite interval of time. Just as hyperbolic discounting may arise from imperfect
 79 foresight (*Gabaix and Laibson, 2017*), maximizing reward rate over shorter time horizons predicts
 80 hyperbolic time preferences (*Namboodiri et al., 2014*).

81 An alternative explanation for the preference for immediate rewards in intertemporal choice
 82 tasks is that animals simply underestimate the duration of post-reward delays; that is, delays
 83 between receiving reward and making the next decision (*Pearson et al., 2010; Blanchard et al.,*
2013). Typically, in intertemporal choice tasks, a variable post-reward delay is added to ensure that
 84 the overall amount of time for each trial is equal, regardless of the reward size or the duration
 85 of the pre-reward delay. It has been argued that it may be difficult for animals to learn these

92 variable delays, and thus, animals may fail to consider the full duration of the delay in their decision
 93 process. Consequently, animals will perceive that it takes less time to acquire the smaller, more
 94 immediate reward and overestimate the reward rate for choosing this option. Consistent with
 95 this hypothesis, providing an explicit cue for the duration of the post-reward delay or increasing
 96 its salience by providing a small reward at the end of the post-reward delay reduces temporal
 97 discounting (*Pearson et al., 2010; Blanchard et al., 2013*).

98 Despite the similarities between overharvesting and the preference for more immediate rewards
 99 in intertemporal choice tasks, prior attempts to use temporal discounting and/or short-term rate
 100 maximization functions fit to intertemporal choice data to predict foraging behavior have failed
 101 (*Carter et al., 2015; Carter and Redish, 2016; Blanchard and Hayden, 2015*). In these studies,
 102 animals are typically closer to long-term rate maximization in foraging tasks than in intertemporal
 103 choice tasks. This has been taken as further evidence that, while animals have a good understanding
 104 of the structure of foraging tasks, they struggle to understand the structure of intertemporal
 105 choice tasks (i.e. they fail to incorporate post-reward delays into their decision process; *Blanchard*
 106 *and Hayden 2015*). However, there are two additional possibilities for why intertemporal choice
 107 models have failed to predict foraging behavior. First, models of intertemporal choice tasks usually
 108 consider rewards for the current trial and not rewards on future trials since, in these tasks, reward
 109 opportunities on future trials are often independent of the current decision. This is not true of
 110 foraging tasks, in which future opportunities for rewards depend on the current decision. Thus,
 111 this difference in decision horizon may make it difficult to explain foraging data using discounting
 112 models fit to intertemporal choice data. Second, these studies have only examined standard, single-
 113 parameter exponential and hyperbolic discounting functions. More flexible forms of temporal
 114 discounting that produce different patterns of hyperbolic time preferences have never been tested
 115 in these contexts. More flexible discounting functions include constant sensitivity discounting, by
 116 which rewards in the distant future are discounted less than rewards in the near future due to
 117 a bias in time estimation – agents become less sensitive to longer time delays (*Ebert and Prelec,*
 118 *2007; Zauberman et al., 2009*); additive-utility discounting, by which the utility of a reward, not its
 119 value, is discounted (*Killeen, 2009*); or quasi-hyperbolic discounting, which has separate terms for
 120 short- or long-term rewards or changes in the discount rate for short- and long-term rewards that
 121 correlate with activity in limbic and fronto-parietal networks respectively (*Laibson, 1997; McClure*
 122 *et al., 2004, 2007*).

123 In the present study, we found that rats exhibit similar time preferences in foraging and intertem-
 124 poral choice tasks and that time preferences in both tasks can be explained by a quasi-hyperbolic
 125 discounting model that, in both contexts, considers future rewards. Rats were tested in a series of
 126 patch foraging tasks and an intertemporal choice task. In foraging tasks, they followed the basic
 127 predictions of long-term rate maximization: they stayed longer in patches that yielded greater
 128 rewards and when the cost of searching was greater. But under certain conditions, they violated
 129 these predictions in a manner consistent with time preferences: they stayed longer in patches
 130 when given larger rewards with proportionately longer delays, and they exhibited greater sensitivity
 131 to pre- vs. post-reward delays. Similarly, in an intertemporal choice task, rats exhibited greater
 132 sensitivity to pre- vs. post-reward delays. Using these data, we tested several models to determine if
 133 temporal discounting or biases in time perception, such as insensitivity to post-reward delays, could
 134 explain rats' behavior across tasks. One model, a quasi-hyperbolic discounting model (*Laibson,*
 135 *1997; McClure et al., 2007*), provided the best fit to rat behavior across all experiments. Further-
 136 more, the quasi-hyperbolic discounting model proved to be externally valid: discounting functions
 137 fit to foraging data provided as good a fit to intertemporal choice data as discounting functions
 138 fit directly to intertemporal choice data for some rats. These findings suggest that rats exhibit
 139 similar biases in the two tasks, and quasi-hyperbolic discounting may be a common mechanism for
 140 suboptimal decision-making across tasks.

141 **Results**

142 **Rats consider long-term rewards, but exhibit a bias in processing pre- vs. post-**
 143 **reward delays**

144 Long Evans rats ($n = 8$) were tested in a series of patch foraging tasks in operant conditioning
 145 chambers (Kane et al., 2017). To harvest reward (10% sucrose water) from a patch, rats pressed a
 146 lever on one side of the front of the chamber (left or right) and reward was delivered in an adjacent
 147 port. After a post-reward delay (inter-trial interval or ITI), rats again chose to harvest a smaller
 148 reward or to leave the patch by nose poking in the back of the chamber. A nose poke to leave
 149 the patch caused the harvest lever to retract and initiated a delay to control the time to travel to
 150 the next patch. After the delay, the opposite lever extended (e.g. if the left lever was extended
 151 previously, the right lever would be extended now), and rats could then harvest from (or leave) this
 152 replenished patch (Figure 1–Figure Supplement 1).

153 In four separate experiments, we manipulated different variables of the foraging environment:
 154 i) in the “Travel Time Experiment,” a 10 s vs. 30 s delay was imposed between patches, ii) in the
 155 “Depletion Rate Experiment,” reward depleted at a rate of 8 vs. 16 μL per harvest, iii) in the “Scale
 156 Experiment,” the overall magnitude of rewards and delays was varied, such that in one condition,
 157 the size of rewards and length of delays was twice that of the other. iv) Finally, in the “Pre-vs-
 158 Post Experiment,” the placement of delays was varied, such that the total time to harvest reward
 159 remained constant, but in one condition there was no pre-reward delay and ~13 s post reward delay,
 160 and in the other there was a 3 s pre-reward delay and ~10 s post-reward delay. Parameters for each
 161 experiment are shown in Table 1. For each condition within each experiment, rats were trained for
 162 5 days and tested for an additional 5 days; all behavioral data presented is from the 5 test days. The
 163 order of conditions within each experiment was counterbalanced across rats. Every patch visit was
 164 included for analysis; mixed effects models were used to examine the effect of task condition on the
 165 number of trials spent in each patch. Random intercepts and random slopes for the effect of task
 166 condition were used to group observations within each rat. To compare rat behavior to the optimal
 167 behavior in each condition, a mixed effects model was used to test the effect of task condition on
 168 the difference between the number of trials spent in each patch and the optimal number of trials
 169 for that patch, with random intercepts and slopes for each rat. For this mixed effects model, an
 170 intercept of zero indicates optimal performance, and the slope indicates the change in behavior
 171 relative to the optimal behavior between conditions (see Methods for additional detail).

172 The Travel Time Experiment was designed to test the two main predictions of MVT: i) that animals
 173 should stay longer in patches that yield greater rewards and ii) animals should stay longer in all
 174 patches when the cost of traveling to a new patch is greater. In this experiment, rats encountered
 175 three different patch types within sessions, which started with varying amount of reward (60, 90,
 176 or 120 μL) and depleted at the same rate (8 $\mu\text{L}/\text{harvest}$). The delay between patches was either
 177 10 s or 30 s; each travel time delay was tested in its own block of sessions and the order was
 178 counterbalanced across rats, with a range of 87–236 patches visited per condition per rat. As
 179 predicted by MVT, rats stayed for more trials in patch types that started with larger reward volume
 180 ($\beta = 118.091 \text{ trials/mL}$, SE = 1.862, $t(2490.265) = 63.423$, $p < .001$), indicating that rats considered
 181 reward across future patches. Rats also stayed longer in all patch types when time between patches
 182 was longer ($\beta = 1.893 \text{ trials}$, SE = .313, $t(118.839) = 6.040$, $p < .001$; Figure 1A), indicating sensitivity
 183 to opportunity costs. However, rats uniformly overharvested relative to predictions of MVT ($\beta_{\text{rat-MVT}}$
 184 = 3.396 trials, SE = .176, $t(6.960) = 19.269$, $p < .001$). The degree to which rats overharvested was
 185 not significantly different between the 10 s and 30 s travel conditions ($\beta_{10 \text{ s}-30 \text{ s}} = .304 \text{ trials}$, SE =
 186 .155, $t(7.3857) = 1.964$, $p = .088$).

187 The Depletion Rate Experiment tested another critical variable in foraging environments: the
 188 rate of reward depletion within a patch. Quicker reward depletion causes the local reward rate to
 189 deplete to the long-run average reward rate quicker, thus MVT predicts earlier patch leaving. Within
 190 sessions, rats encountered a single patch type (starting volume of 90 μL) that depleted at a rate of

Table 1. Parameters for each of the first 4 foraging experiments. Harvest time = time to make a decision to harvest + pre-reward delay + post-reward delay (inter-trial interval). To control reward rate in the patch, the post-reward delay was adjusted relative to the decision time to hold the harvest time constant. [†]Rats encountered all three patch types in both conditions. *One group of rats ($n = 8$) was tested on the first four experiments, but a separate group ($n = 8$) was tested on this final foraging experiment. **In this experiment, the harvest time was not held constant — the post-reward delay was always 3 s or 12 s regardless of the time to make a decision.

Experiment	Condition	Start Reward	Depletion Rate	Pre-Reward Delay	Harvest Time	Travel Time
travel time	10s	60, 90, or 120 μ L [†]	-8 μ L	0 s	10 s	10 s 30 s
	30 s					
depletion rate	-8 μ L	90 μ L	-8 μ L -16 μ L	0 s	12 s	12 s
	-16 μ L					
scale	90 μ L/10 s	90 μ L	-8 μ L -16 μ L	0 s	10 s 20 s	10 s 20 s
	180 μ L/20 s					
handling time	0 s	90 μ L	-8 μ L	0 s 3 s	15 s	15 s
	3 s					
post-reward delay*	3 s	90 μ L	-8 μ L	0 s	5-8 s** 13-16 s**	10 s
	12 s					

either 8 or 16 μ L/trial, tested in separate sessions and counterbalanced, with a range of 152-283 patches visited per condition per rat. As predicted by MVT, rats left patches earlier when patches depleted more quickly ($\beta = 2.589$ trials, SE = .155, $t(7.000) = 16.75$, $p < .001$; **Figure 1B**). But, again, rats stayed in patches longer than is predicted by MVT ($\beta_{rat-MVT} = 2.005$ trials, SE = .134, $t(7.004) = 14.97$, $p < .001$). Rats overharvested to a greater degree in the 8 μ L depletion condition than the 16 μ L depletion condition ($\beta_{8 \mu L-16 \mu L} = 1.589$ trials, SE = .155, $t(7.000) = 10.28$, $p < .001$).

These first two experiments confirm that rats qualitatively follow the predictions of MVT, but consistently overharvest. There are many possible explanations for this pattern of overharvesting, including an aversion to leaving the offer of reward within a patch and nonlinear reward utility (*Wikenheiser et al., 2013; Carter and Redish, 2016; Constantino and Daw, 2015*). The Scale Experiment was conducted in an effort to distinguish between these hypotheses by manipulating the scale of time delays and rewards. Long-term rate maximization predicts that an increase in reward size in proportion to reward delay should have no effect on the number of harvests per patch, as the reward rate across trials would be equal. But if animals' perception of reward or time is nonlinear, a manipulation of scale will affect their subjective point of equality and predict a change in behavior across the two environments. The scale of rewards and delays was manipulated in the following manner: patches started with (A) 90 or (B) 180 μ L of reward, depleted at a rate of (A) 8 or (B) 16 μ L/trial, and the duration of harvest trials and travel time between patches was (A) 10 or (B) 20 s. Rats visited a range of 60-212 patches per condition. They overharvested in both A and B conditions ($\beta_{rat-MVT} = 4.374$ trials, SE = .153, $t(6.900) = 28.597$, $p < .001$) and, contrary to predictions of MVT, they stayed in patches significantly longer and overharvested to a greater degree in the B condition that provided larger rewards but at proportionately longer delays ($\beta = 1.937$ trials, SE = .193, $t(6.972) = 9.996$, $p < .001$; **Figure 1C**). This finding suggests that a nonlinearity in the perception of reward value and/or time contributes to overharvesting.

To distinguish between biases in perception of reward, such as nonlinear reward utility, and time, such as temporal discounting or insensitivity to post-reward delays, the Pre-vs-Post Experiment directly tested rats sensitivity to time delays before vs. after reward. In this experiment, in one condition, rats received reward immediately after lever press followed by a post-reward delay of ~13

219 s before the start of the next trial. In the other condition, there was a 3 s pre-reward delay between
 220 lever press and receiving reward followed by a shorter post-reward delay of ~10 s. The total time of
 221 each trial was held constant between conditions (15 s total), so there was no difference in reward
 222 rates. Both MVT and nonlinear reward utility predict that the placement of delays is inconsequential
 223 and that rats will behave similarly in both conditions. Both temporal discounting and insensitivity
 224 to post-reward delays predict that rats will value the immediate reward more than the delayed
 225 reward and thus, would leave patches earlier in the condition with the pre-reward delay. Consistent
 226 with predictions of temporal discounting and insensitivity to post-reward delays, and contrary to
 227 predictions of MVT and nonlinear reward utility, rats left patches earlier in the environment with
 228 the pre-reward delay ($\beta = 2.345$ trials, SE = .313, $t(7.017) = 7.503$, $p < .001$; **Figure 1D**). This result
 229 suggests that a bias in rats' perception of time or the way in which they perceive delayed reward
 230 values contributes to overharvesting.

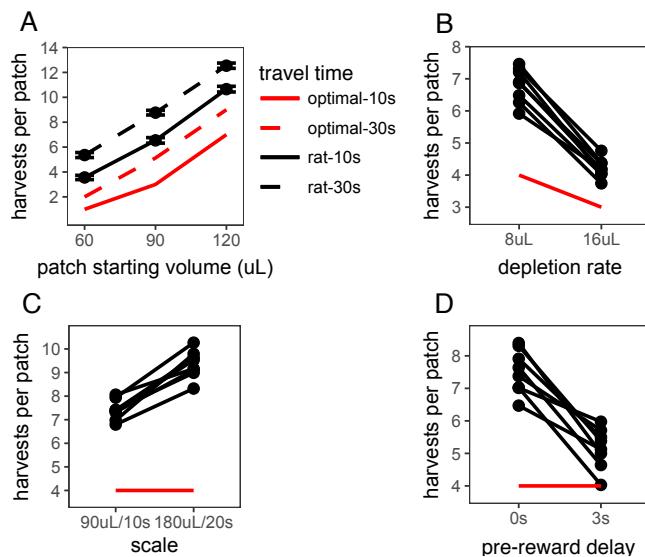


Figure 1. Rat foraging behavior in the A) Travel Time, B) Depletion Rate, C) Scale, and D) Pre-vs-Post Experiments. In A, points and error bars represent mean \pm standard error. In B-D, points and connecting lines represent behavior of each individual rat. Red lines indicate optimal behavior (per MVT).

Figure 1-Figure supplement 1. Diagram of the foraging task

Figure 1-source data 1. Trial-by-trial foraging behavior.

231 To determine whether the preference for immediate rewards can be explained by insensitivity
 232 to post-reward delays, a fifth foraging experiment, the "Post-Reward Delay Experiment," directly
 233 tested rats' sensitivity to post-reward delays. A separate cohort of rats ($n = 8$) was used for this
 234 experiment. Rats were tested in two conditions in this experiment: a short (3 s) or long (12 s)
 235 post-reward delay. The total time of harvest trials was not held constant; the longer post-reward
 236 delay increased the time to harvest from the patch. Since the longer post-reward delay increases
 237 the cost of harvesting from the patch relative to the cost of traveling to a new patch, MVT predicts
 238 that rats should leave patches earlier. Prior studies of intertemporal choice behavior have shown
 239 that animals are insensitive to post-reward delays, suggesting that they are only concerned with
 240 maximizing short-term reward rate (*Stephens and Anderson, 2001; Bateson and Kacelnik, 1996*)
 241 or that they may not have learned the duration of post-reward delays, and underestimate this
 242 duration in their decision process (*Pearson et al., 2010; Blanchard et al., 2013*). Consistent with
 243 MVT, rats were sensitive to the post-reward delay, leaving patches earlier in the 12 s delay condition
 244 ($\beta = 1.411$ trials, SE = .254, $t(6.966) = 5.546$, $p < .001$; **Figure 2A**). If rats were sensitive to the delay,
 245 but underestimated its duration, one would still expect rats to overharvest to a greater degree due
 246 to the longer delay. There was no difference in the degree to which rats overharvested between the

247 3 s and 12 s delay conditions ($\beta_{\text{rat-MVT}; 3 \text{ s}-12 \text{ s}} = .340$ trials, SE = .286, $t(6.963) = 1.188$, $p = .274$). This
 248 finding suggests that overharvesting in this experiment is not due to insensitivity to post-reward
 249 delays. However, it is possible that this finding could be explained by other forms of altered time
 250 perception that remain to be described.

251 The data from the foraging experiments described above suggest that rats exhibit time preferences
 252 in the foraging task. In a final “Intertemporal Choice Experiment,” we tested whether
 253 the same rats that participated in the Post-Reward Delay Experiment would exhibit similar time
 254 preferences in a standard intertemporal choice (i.e. a delay-discounting) task. This task consisted
 255 of a series of 20-trial episodes. On each trial, rats pressed either the left or right lever to receive a
 256 smaller-sooner (SS) reward of 40 μL after a 1 s delay or a larger-later (LL) reward of 40, 80, or 120 μL
 257 after a 1, 2, 4, or 6 s delay. For the first 10 trials of each episode, rats were forced to press either the
 258 left or right lever to learn the value and delay associated with that lever (only one lever extended on
 259 each of these trials). For the last 10 trials of an episode, both levers extended and rats were free
 260 to choose. The LL reward value and delay, and the LL lever (left or right) were randomly selected
 261 at the start of each episode. Rats were tested in two different versions of this task: one in which
 262 the post-reward delay was held constant, such that the longer pre-reward delays reduced reward
 263 rate (constant delay); and another in which the time of the trial was held constant, such that longer
 264 pre-reward delays resulted in shorter post-reward delays to keep reward rate constant (constant
 265 rate). MVT, which maximizes long-term reward rate, predicts that rats would be sensitive to the
 266 pre-reward delay in the constant delay condition but not the constant trial condition (in which the
 267 pre-reward delay does not affect reward rate).

268 Rats were given three training sessions to learn the structure of the intertemporal choice task
 269 after previously being tested in the foraging task, then they were tested for an additional 13 sessions
 270 in each condition, participating in a range of 590-2810 free choice trials per condition (constant
 271 delay vs. constant rate). Each free choice trial within each episode was counted as a separate
 272 observation. Choice data were analyzed using a generalized linear mixed-effects model (i.e. a
 273 mixed-effects logistic regression) to examine the effect of the size of the LL reward, the length of
 274 the LL delay, task condition (constant delay vs. constant rate), and their interactions on decisions
 275 to choose the LL vs. SS option, with random intercepts and random slopes for the effects of LL
 276 reward, LL delay, and task condition for each rat. Three post-hoc comparisons were used to test the
 277 effects of i) LL reward and ii) LL delay within each condition, and iii) LL delay between the constant
 278 delay and constant rate conditions (*Figure 2B*). i) In both conditions, rats were more likely to choose
 279 larger LL rewards (constant delay: $\beta = .477$, SE = .090, $\chi^2(1) = 28.320$, $p < .001$; constant rate: $\beta =$
 280 .450, SE = .089, $\chi^2(1) = 25.378$, $p < .001$), showing that they were sensitive to reward magnitude. ii)
 281 They were also sensitive to the pre-reward delay in both conditions (constant delay: $\beta = -.240$, SE
 282 = .023, $\chi^2(1) = 104.882$, $p < .001$; constant rate: $\beta = -.152$, SE = .022, $\chi^2(1) = 46.919$, $p < .001$). On
 283 average, rats were equally likely to select the LL option across conditions — the main effect of task
 284 condition was not significant ($\beta = .010$, SE = .105, $z = .092$, $p = .927$). iii) However, rats were less
 285 sensitive to increasing pre-reward delays when pre-reward delays did not affect reward rate (in the
 286 constant rate condition), indicated by a change in LL delay slope between conditions ($\beta = .088$, SE =
 287 .026, $\chi^2(1) = 11.376$, $p < .001$). Overall, rats exhibited similar time preferences in the foraging and
 288 intertemporal choice tasks: they valued rewards less with longer delays until receipt but they were
 289 sensitive to opportunity costs (e.g. time delays between receiving reward and future decisions).

290 Quasi-hyperbolic discounting best explains behavior across all tasks

291 To test whether a common set of cognitive biases could explain time preferences in both the
 292 foraging and intertemporal choice tasks, both tasks were modeled as continuous time semi-markov
 293 processes. These models consisted of a set of states that represented the time between each
 294 event in each of the tasks (e.g. cues turning on/off, lever press, reward delivery; for state space
 295 diagrams of both tasks, see *Figure 3–Figure Supplement 1* and *Figure 4–Figure Supplement 1*).
 296 These models assumed that animals have learned the appropriate structure of the task (i.e. the

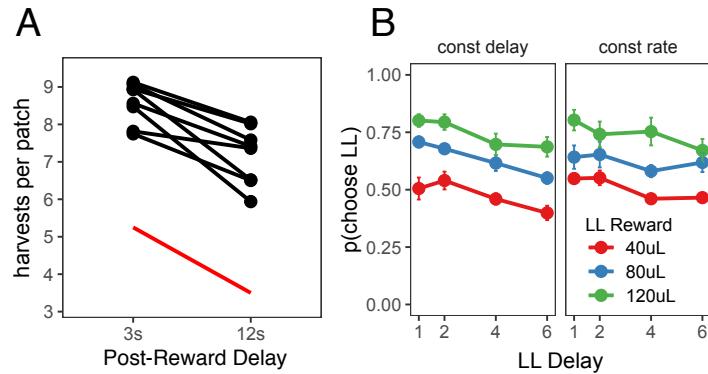


Figure 2. A) Rat behavior in the Post-Reward Delay Experiment. Points and lines represent behavior of individual rats. Red line indicates optimal behavior (per MVT). B) Rat behavior in the two-alternative intertemporal choice task. Points and error bars represent mean \pm standard error for each condition.

Figure 2-source data 1. Trial-by-trial intertemporal choice behavior

time spent and reward obtained in each state) unless otherwise noted. The value of a given state was the discounted value of all future rewards available from that state, and the agent chose the option that yielded the greatest discounted future reward via a stochastic process. As the discount factor approached 1 (i.e. no temporal discounting), this model converged to long-term reward maximization, equivalent to MVT. Additional parameters were added to the model to test four specific hypotheses for suboptimal foraging behavior: i) subjective costs associated with leaving a patch, in which the value of leaving was reduced by a “cost” term; ii) nonlinear reward utility, in which the subjective utility of a reward increased sublinearly with respect to the reward magnitude; iii) biased time perception, which assumed that animals underestimate post-reward delays, possibly due to insufficient learning of task structure (Blanchard et al., 2013; Pearson et al., 2010), or overestimate pre-reward delays; and iv) temporal discounting. A brief description of each hypothesis and its general predictions can be found in *Table 2*. For each model, group level parameters and parameters for each individual rat were fit simultaneously using an expectation-maximization algorithm (Huys et al., 2011). Parameters were fit to each experiment separately (one set of parameters for both conditions in each experiment). Model predictions were calculated separately for each rat, using the rat’s individual parameters. Full details for all models, fitting procedures, and model comparison can be found in the Methods.

Subjective costs to leave a patch and nonlinear reward utility have explained suboptimal foraging behavior in prior studies that have manipulated opportunity costs (e.g. travel time or pre-reward delays) and depletion rate (Constantino and Daw, 2015; Wikenheiser et al., 2013; Carter and Redish, 2016). However, these factors are insensitive to the placement of time delays (pre- vs. post-reward) and thus, cannot explain the preference for more immediate rewards. Consistent with prior studies, the subjective costs and nonlinear reward utility models explained overharvesting in the Travel Time, Depletion Rate, and Post-Reward Delay Experiments, but they failed to explain time preferences in the Pre-vs-Post foraging experiment (*Figure 3-Figure Supplement 2*).

We next examined whether biased time perception and temporal discounting could explain suboptimal foraging behavior across all tasks. Three implementations for biased time perception were tested: linear underestimation of post-reward delays ($postDelay = \alpha * postDelay$), nonlinear underestimation of post-reward delays ($postDelay = postDelay^\alpha$), and overestimation of pre-reward delays ($preDelay = \alpha * preDelay$). For temporal discounting, we tested the two common single-parameter discounting functions, exponential ($value = e^{-\beta*time} * reward$) and standard hyperbolic ($value = reward/(1 + k * time)$), and two more flexible discounting models: constant sensitivity discounting (Ebert and Prelec, 2007; Zauberman et al., 2009), which predicts hyperbolic time preferences due to insensitivity to longer delays ($value = e^{-\beta*time^\alpha} * reward$); and quasi-hyperbolic discounting, formalized as two competing exponential discounting systems

Table 2. Description of hypotheses for overharvesting with general, qualitative predictions for the degree of overharvesting in each experiment. Quantitative predictions depend on the exact formalization of each model and its specific parameters.

	Hypothesis	Experimental Predictions
Subjective Costs	A cost term c reduces the value of leaving a patch. Predicts greater overharvesting with higher c . Not affected by specific manipulations to reward or time.	Rats will follow qualitative predictions of the Marginal Value Theorem, but exhibit an equal degree of overharvesting across conditions in each experiment.
Nonlinear Reward Utility	Subjective value increases sublinear to reward magnitude. Predicts greater overharvesting with steeper utility functions with larger rewards.	Rats will exhibit an equal degree of overharvesting in all experiments except for the Scale experiment. In the Scale experiment, rats will overharvest more in the conditions with larger rewards.
Biased Time Perception	i) Post-reward delays perceived as shorter, ii) pre-reward delays perceived as longer, or iii) longer delays (irrespective of their placement) perceived as shorter. All three hypotheses predict greater overharvesting with longer delays.	Rats will exhibit a greater degree of overharvesting in the condition with longer delays in the Scale environment, in the condition with the longer post-reward delay in the Pre-vs-Post experiment, and in the condition with longer post-reward delay in the Post-Reward Delay experiment
Temporal Discounting	Value of future rewards discounted due to delay to receive them. Predicts greater overharvesting with greater levels of discounting and with longer delays	Rats will overharvest to a greater degree in the conditions with longer delays in the Scale and Post-Reward Delay experiments and they will leave patches earlier due to the longer pre-reward delay in the Pre-vs-Post experiment.

332 $value = [\omega * e^{-\beta*time} + (1 - \omega) * e^{-\delta*time}] * reward$; *Laibson 1997; McClure et al. 2007*). All of these
 333 models qualitatively predicted rat behavior across foraging experiments (*Figure 3, Figure 3–Figure*
 334 *Supplement 3, Figure 3–Figure Supplement 4*).

335 To determine which model provided the best quantitative fit, we compared the group-level
 336 Bayesian Information Criterion (integrated BIC or iBIC; *Huys et al. 2011, 2012*) of all models in each
 337 of the foraging tasks. To compare across tasks, we took the sum of iBIC for each model. The quasi-
 338 hyperbolic discounting model had the lowest sum of iBIC, the constant sensitivity discounting model
 339 the second lowest, and the hyperbolic discounting model third (*Figure 3*). These three models were
 340 also among the lowest iBIC values for each individual experiment (*Figure 3–Figure Supplement 5*).
 341 All three of these models predict that animals will exhibit hyperbolic time preferences, suggesting
 342 that suboptimal foraging behavior observed in these experiments is due to time preferences.

343 Next, we tested whether the quasi-hyperbolic discounting model that provided the best fit to
 344 foraging behavior could also explain behavior in the intertemporal choice task. As in the model of
 345 the foraging task, the model of the intertemporal choice task took into account all future rewards,
 346 including rewards from future episodes (see abbreviated state space diagram in *Figure 4–Figure*
 347 *Supplement 1*). We tested the nonlinear reward utility, biased time perception and temporal

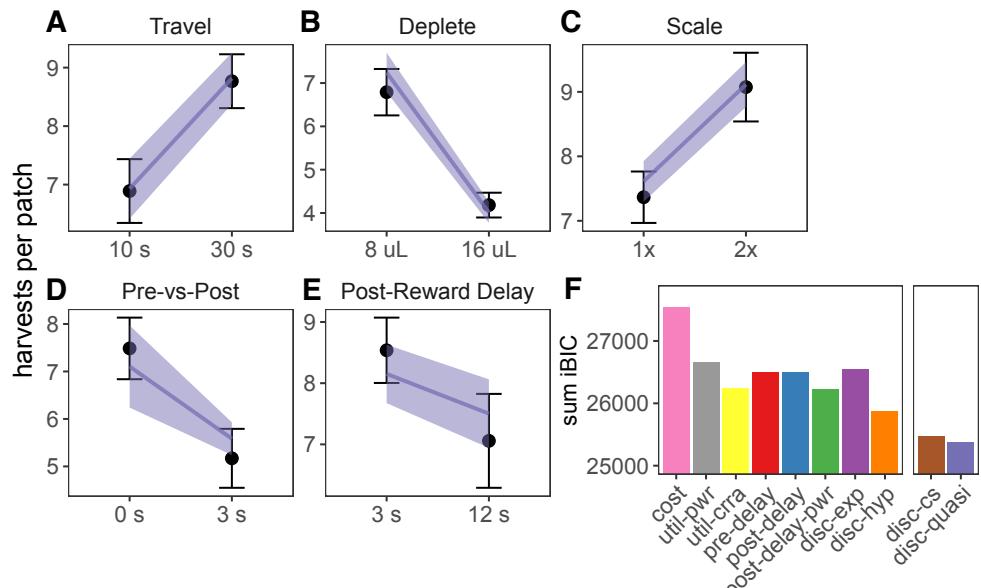


Figure 3. A-E) Predictions of the best fit quasi-hyperbolic discounting model to all foraging tasks. Points and error bars represent mean \pm standard deviation of the means for each individual rat; lines and ribbon represent the mean \pm standard deviation of the means of the model-predicted behavior for each individual rat. F) The sum of iBIC scores across all foraging tasks for each model. Cost = subjective cost model, util-pwr and util-crra = nonlinear reward utility with power and CRRA function respectively, pre-del = linear overestimation of pre-reward delays, post-del = linear underestimation of post-reward delays, post-del-pwr = underestimation of post-reward delays according to a power function, disc-exp = exponential discounting, disc-hyp = hyperbolic discounting, disc-CS = constant sensitivity discounting, disc-quasi = quasi-hyperbolic discounting.

Figure 3-Figure supplement 1. State space diagram of the foraging task.

Figure 3-Figure supplement 2. Predictions of the best fit subjective cost and nonlinear reward utility models

Figure 3-Figure supplement 3. Predictions of the best fit biased time perception models.

Figure 3-Figure supplement 4. Predictions of the best fit discounting models.

Figure 3-Figure supplement 5. iBIC for each model for each foraging experiment.

348 discounting models in this task (the subjective cost model does not apply to this task). Again, the
 349 quasi-hyperbolic discounting model had the lowest iBIC and hyperbolic discounting model the
 350 second lowest, but the constant sensitivity model had a higher iBIC than the biased time perception
 351 models (**Figure 4**). As the constant sensitivity model produces hyperbolic time preferences via
 352 insensitivity to longer delays, these results suggest that hyperbolic time preferences without
 353 insensitivity to delays is the best explanation for rat intertemporal choice behavior.

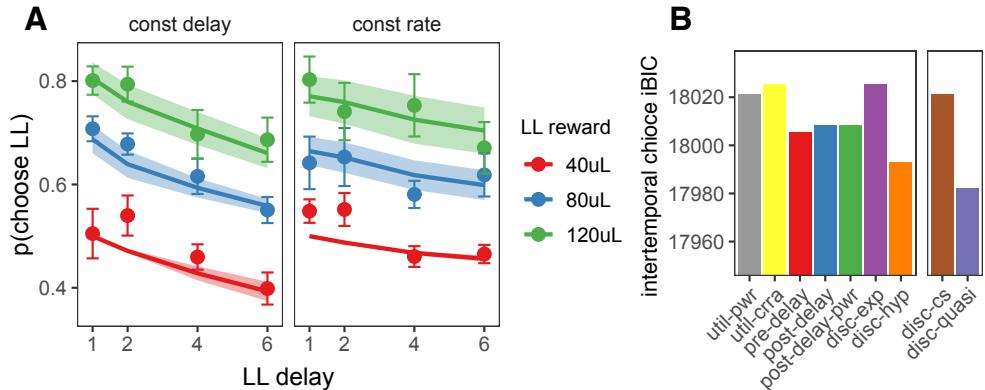


Figure 4. A) Quasi-hyperbolic model predictions for the intertemporal choice task. Points and error bars represent the mean \pm standard error of individual rat behavior; lines and ribbon represent mean \pm standard error of model predicted behavior for each individual rat. B) The iBIC score for each model for the delay discounting experiment. Util-pwr and util-crra = nonlinear reward utility with power and CRRA function respectively, pre-del = linear overestimation of pre-reward delays, post-del = linear underestimation of post-reward delays, post-del-pwr = underestimation of post-reward delays according to a power function, disc-exp = exponential discounting, disc-hyp = hyperbolic discounting, disc-CS = constant sensitivity discounting, disc-quasi = quasi-hyperbolic discounting.

Figure 4-Figure supplement 1. State space diagram of the intertemporal choice task.

Figure 4-Figure supplement 2. Comparison of all-future horizon and one-trial horizon discounting models.

354 If hyperbolic time preferences reflect a common explanation for suboptimal decision-making,
 355 then it might be expected that a model of behavior fit to one task could predict a rat's behavior in
 356 the other task. To test the external validity of this hypothesis, data from each task were separated
 357 into three subsets. The best fitting model from both tasks, the quasi-hyperbolic discounting model,
 358 was fit to two subsets of data from one task, then the negative log likelihood (-LL) of the data was
 359 assessed on the left out sample from both tasks. This process was repeated such that each subset
 360 served as the left out sample. To determine which discount function provided the better fit to data
 361 from each task, we calculated the difference in -LL of the left out sample between the model fit to
 362 intertemporal choice data and the model fit to foraging data (-LL difference = $-\text{LL}_{\text{itc}} - \text{LL}_{\text{forage}}$). Since
 363 smaller -LL indicates a better fit, a positive -LL difference indicates that the discount function fit to
 364 foraging data provided a better fit (i.e. the foraging -LL was lower than the intertemporal choice -LL).
 365 For the foraging task, discounting functions fit to foraging data provided a better fit than discounting
 366 functions fit to intertemporal choice data for all 8 rats. Interestingly, for the intertemporal choice
 367 task, discounting functions fit to foraging data provided a better fit than discounting functions fit to
 368 intertemporal choice data for 3 of 8 rats (**Figure 5**). The quasi-hyperbolic model fit to the foraging
 369 task generalized well to the intertemporal choice task, providing support for the idea that foraging
 370 and intertemporal choice can be described by a common discount function.

371 With temporal discounting models that consider all future rewards, the more flexible quasi-
 372 hyperbolic discounting function provided the best fit to behavior across tasks. We next directly
 373 tested whether considering future rewards affects the fit of discounting models to intertemporal
 374 choice data and the estimates of discount factors compared to temporal discounting models that
 375 only consider the next reward (one-trial horizon models). We fit one-trial horizon models for all
 376 of the previously tested discounting functions — exponential, hyperbolic, constant sensitivity, and

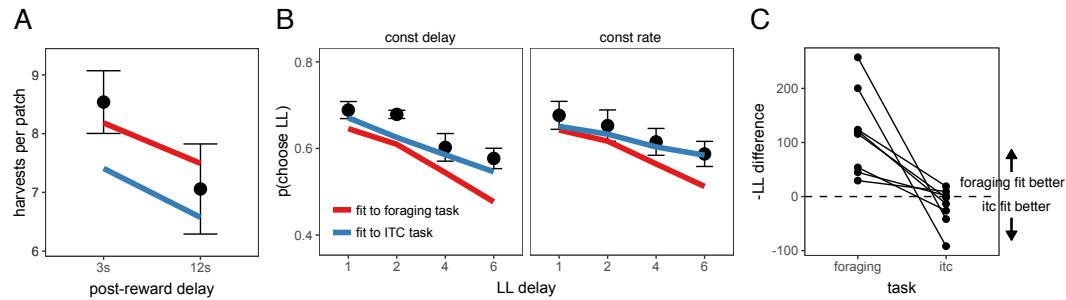


Figure 5. A) Predicted foraging behavior for quasi-hyperbolic model parameters fit to either the foraging task (red line) or delay discounting task (DD; blue line). Black points and error bars represent mean \pm standard error of rat data. B) Predicted intertemporal choice behavior for quasi-hyperbolic model parameters fit to data from either the foraging or delay discounting task, plotted against rat behavior. C) The difference in negative log likelihood of the left out sample of foraging data (left) or intertemporal choice data (right) between parameters fit to the intertemporal choice task and parameters fit to the foraging task. A negative -LL difference indicates the negative log likelihood of the data for parameters fit to the intertemporal choice task was lower than for parameters fit to the foraging task. Each point and line represents data from individual rats.

377 quasi-hyperbolic discounting — and compared them to the discounting models that considered
 378 all future rewards (all-future horizon models). For all discounting functions, the all-future horizon
 379 models had lower iBIC than one-trial horizon models (**Figure 4–Figure Supplement 2**). To compare
 380 the discount factors of each model (for the quasi-hyperbolic function that has two discount factors,
 381 we used the slow discounting β), we performed paired t-tests between log transformed discount
 382 factors measured by the all-future horizon models vs. discount factors measured by the one-trial
 383 horizon models. Measured discount factors were lower for the all-future models for all discounting
 384 functions (exponential: $t(7) = 22.439$, $p < .001$; hyperbolic: $t(7) = 7.000$, $p < .001$; constant sensitivity:
 385 $t(7) = 15.497$, $p < .001$; quasi-hyperbolic: $t(7) = 25.322$, $p < .001$; p-values adjusted using Bonferroni
 386 correction). Lastly, we tested whether the all-future horizon quasi-hyperbolic discounting model fit
 387 to the intertemporal choice data would predict foraging behavior better than the one-trial horizon
 388 quasi-hyperbolic discounting model fit to intertemporal choice data. For 6 of 8 rats, parameters fit
 389 to the one-trial horizon model produced a better fit to foraging behavior than parameters fit to the
 390 all-future horizon model. Overall, using full horizon temporal discounting models explained more
 391 of the intertemporal choice data, produced smaller estimates of discounting factors, but in the
 392 present study, it did not improve the ability of a model fit to intertemporal choice data to predict
 393 foraging behavior.

394 Discussion

395 In foraging studies, animals exhibit behavior that conforms qualitatively to predictions made by
 396 optimal foraging theory (i.e., the MVT), choosing to leave a patch when its value falls below that
 397 of the average expected value of other(s) available in the environment. However, an almost
 398 ubiquitous finding is that they overharvest, leaving a patch when its value falls to a value lower
 399 than the one predicted by MVT. Given that the rewards available within the current patch are
 400 generally available sooner than those at other patches due to travel time, one interpretation of
 401 overharvesting is that this reflects a similarly prevalent bias observed in intertemporal choice tasks,
 402 in which animals consistently show a greater preference for smaller more immediate rewards
 403 over later delayed rewards than would be predicted by optimal (i.e., exponential) discounting of
 404 future values. However, in prior studies, models of intertemporal choice behavior have been poor
 405 predictors of foraging behavior (**Blanchard and Hayden, 2015; Carter and Redish, 2016**). Here, we
 406 show that in a carefully designed series of experiments, rats exhibit similar time preferences in
 407 foraging and intertemporal choice tasks, and that a quasi-hyperbolic discounting model can explain
 408 the rich pattern of behaviors observed in both tasks.

409 The foraging behavior we observed was consistent with previous studies of foraging behavior

410 in rats, monkeys, and humans, while also revealing novel aspects of overharvesting behavior.
 411 Consistent with prior studies, rats stayed longer in patches that yielded greater rewards, stayed
 412 longer in all patch types when the cost of traveling to a new patch was greater, left patches
 413 earlier when rewards depleted more quickly, and consistently overharvested (*Constantino and*
414 Daw, 2015; Hayden et al., 2011; Kane et al., 2017). Our experiments also demonstrated that in
 415 certain environments rats violate qualitative predictions of MVT. Rats overharvested more when
 416 reward amount and delay were increased, even though reward rate was held constant, and they
 417 were differentially sensitive to whether the delay was before the receipt of the proximal reward
 418 or following its delivery. These findings supported the conjecture that overharvesting is related to
 419 time preferences.

420 A number of studies have found that the preference for smaller, more immediate rewards
 421 can be explained by insensitivity to post-reward delays (*Bateson and Kacelnik, 1996; Blanchard*
422 et al., 2013; Pearson et al., 2010; Stephens and Anderson, 2001; Mazur, 1991). One hypothesis for
 423 why animals fail to incorporate post-reward delays into decisions is that they haven't learned the
 424 structure of the task well, and thus cannot accurately predict future post-reward delays. Accordingly,
 425 providing explicit cues for the post-reward delays or increasing the salience of post-reward delays
 426 helps animals incorporate these delays into their decisions, reducing the bias towards selecting
 427 smaller, more immediate rewards over larger, delayed ones (*Pearson et al., 2010; Blanchard et al.,*
428 2013). However, in the present study, rats were sensitive to post-reward delays in both the foraging
 429 and intertemporal choice task, providing further evidence that the preference for smaller, more
 430 immediate rewards in both tasks is due to time preferences and not a poor understanding of the
 431 task structure. Furthermore, quantitative modeling supported the hypothesis that suboptimal
 432 behavior was driven by time preferences rather than insensitivity to delays.

433 The idea that animals exhibit similar decision biases in foraging and intertemporal choice
 434 paradigms, and that these biases can be explained by a common model of discounting, is in conflict
 435 with prior studies that found that animals are better at maximizing long-term reward rate in foraging
 436 than in intertemporal choice tasks, and that delay discounting models of intertemporal choice tasks
 437 are poor predictors of foraging behavior (*Stephens, 2008; Blanchard and Hayden, 2015; Carter*
438 et al., 2015; Carter and Redish, 2016). It has been argued that animals may perform better in
 439 foraging tasks because decision-making systems have evolved to solve foraging problems rather
 440 than two-alternative intertemporal choice problems (*Blanchard and Hayden, 2015; Stephens et al.,*
441 2004; Stephens, 2008). This idea has been challenged by a recent study of human decision-making
 442 in foraging and intertemporal choice tasks, finding that a long-term rate maximization model
 443 explained both foraging and intertemporal choice behavior better than a standard hyperbolic
 444 discounting model (*Seinstra et al., 2018*). Results from the present study support the interpretation
 445 that foraging and intertemporal choice behavior can be explained via a common model, but suggest
 446 that this model is quasi-hyperbolic discounting. We found that a quasi-hyperbolic discounting model
 447 provided the best explanation to rat behavior across multiple foraging tasks and an intertemporal
 448 choice task, and that a quasi-hyperbolic discounting model fit to individual rat foraging behavior
 449 can predict their intertemporal choice behavior.

450 Two potential explanations for why temporal discounting models have failed to predict foraging
 451 behavior in prior studies are that i) prior studies have only tested single-parameter exponential
 452 and hyperbolic discounting functions, whereas the present study also tested the more flexible
 453 quasi-hyperbolic discounting function; and ii) in most of these studies, models of intertemporal
 454 choice tasks have only considered the most proximal reward (the reward received as a consequence
 455 of the decision at hand). This assumption seems appropriate as, in most intertemporal choice tasks,
 456 opportunities for future rewards do not depend on the current decision, so the value of rewards
 457 received for future decisions are equal for both the SS and LL rewards. But in foraging tasks, future
 458 opportunities for reward depend on current decisions, so it is critical for foraging models to include
 459 all future rewards into estimates of reward value. For this reason, comparing discount functions fit
 460 to intertemporal choice models that consider all future reward may provide better estimates of

461 foraging behavior than discount functions fit to intertemporal choice models that only consider
 462 rewards from the most proximal decision. Consistent with this hypothesis, we found that adding
 463 the value of future rewards to intertemporal choice models reduces estimates of discount factors.
 464 However, with our data, the quasi-hyperbolic discounting model fit to the intertemporal choice
 465 task that included all future rewards did not predict foraging behavior better than an equivalent
 466 model that only considered the most proximal reward. One reason why including all future rewards
 467 may not have improved cross-task predictions is that, in the present study, the quasi-hyperbolic
 468 discounting model fit to the intertemporal choice task predicted less overharvesting than was
 469 exhibited by rats in the foraging task. Reducing estimates of the discount factor with a model
 470 that considers all future rewards predicts even less overharvesting (i.e. behavior that is closer to
 471 long-term reward maximization). But in other studies, temporal discounting models typically predict
 472 greater overharvesting than is exhibited by animals (*Blanchard and Hayden, 2015; Carter et al.,*
 473 *2015*). In these cases, obtaining smaller, potentially more accurate estimates of discount factors by
 474 including all future rewards into intertemporal choice models may improve cross-task predictions.

475 Although quasi-hyperbolic discounting provided the best singular explanation for rat behav-
 476 ior across our tasks, many of the other models tested were capable of explaining some of the
 477 biases exhibited by rats. Thus, we cannot exclude the possibility that subjective costs, diminish-
 478 ing marginal utility, and/or biased estimation of time intervals may independently contribute to
 479 suboptimal decision-making. Furthermore, additional hypotheses or additional variants of the
 480 above-mentioned hypotheses that have not been tested in the present study may provide alter-
 481 native explanations for suboptimal decision making in foraging and intertemporal choice tasks.
 482 Importantly, our data indicate that quasi-hyperbolic discounting may provide a link between for-
 483 aging and intertemporal choice tasks, and it highlights the importance of future work considering
 484 the source of time preferences. These observations are buttressed by recent theoretical work
 485 demonstrating that the appearance of time preferences in intertemporal choice tasks can emerge
 486 rationally from a value construction process by which estimates increase in variability with the
 487 delay until reward receipt — an account that shares features with the short-term rate maximization
 488 hypotheses (*Stephens et al., 2004*). Under this account, “as-if” discounting is hyperbolic when
 489 variability increases linearly with delay (*Gabaix and Laibson, 2017*). Further, a sequential sampling
 490 model of two-alternative forced choice (*Bogacz et al., 2006*), parameterized such that outcome
 491 delay scales variability in this way, has recently been shown to capture key dynamical features
 492 of both patch foraging (*Davidson and El Hady, 2019*) and hyperbolic discounting in intertem-
 493 poral choice (*Hunter et al., 2018*). Future work should build on these findings to explore directly
 494 whether the common biases identified here reflect a core computation underlying sampling and
 495 decision-making under uncertainty and across time.

496 Methods

497 Animals

498 Adult Long-Evans rats were used (Charles River, Kingston, NY). One group of eight rats participated
 499 in the scale, travel time, depletion rate, and handling time experiments (in that order), a different set
 500 of eight rats were tested on the post-reward delay foraging experiment then the delay discounting
 501 task. Rats were housed on a reverse 12 h/12 h light/dark cycle. All behavioral testing was conducted
 502 during the dark period. Rats were food restricted to maintain a weight of 85-90% ad-lib feeding
 503 weight, and were given ad-lib access to water. All procedures were approved by the Princeton
 504 University and Rutgers University Institutional Animal Care and Use Committee.

505 Foraging Task

506 Animals were trained and tested as in *Kane et al. (2017)*. Rats were first trained to lever press for
 507 10% sucrose water on an FR1 reinforcement schedule. Once exhibiting 100+ lever presses in a one
 508 hour session, rats were trained on a sudden patch depletion paradigm — the lever stopped yielding

509 reward after 4-12 lever presses — and rats learned to nose poke to reset the lever. Next rats were
 510 tested on the full foraging task.

511 A diagram of the foraging task is in **Figure 1–Figure Supplement 1**. On a series of trials, rats had
 512 to repeatedly decide to lever press to harvest reward from the patch or to nose poke to travel to a
 513 new, full patch, incurring the cost of a time delay. At the start of each trial, a cue light above the
 514 lever and inside the nose poke turned on, indicating rats could now make a decision. The time
 515 from cues turning on until rats pressed a lever or nose poked was recorded as the decision time
 516 (DT). A decision to harvest from the patch (lever press) yielded reward after a short pre-reward
 517 delay (referred to as the handling time delay, simulating the time to “handle” prey after deciding to
 518 harvest). Reward (sucrose water) was delivered when the rat entered the reward magazine. The
 519 next trial began after an inter-trial interval (ITI). To control the reward rate within the patch, the
 520 length of the ITI was adjusted based on the DT of the current trial, such that the length of all harvest
 521 trials was equivalent. With each consecutive harvest, the rat received a smaller volume of reward
 522 to simulate depletion from the patch. A nose poke to leave the patch caused the lever to retract
 523 for a delay period simulating the time to travel to a new patch. After the delay, the opposite lever
 524 extended, and rats could harvest from a new, replenished patch.

525 Details of the foraging environment for each experiment can be found in **Table 1**. For each
 526 experiment, rats were trained on a specific condition for 5 days, then tested for 5 days. Conditions
 527 within experiments were counterbalanced.

528 **Foraging data analysis**

529 Rat foraging behavior was assessed using linear mixed effects models. Models were fit using
 530 the lme4 package in R (**Bates et al., 2015**). The lme4 package provides only t-statistics for fixed
 531 effects; p-values were calculated using the lmerTest package (**Kuznetsova et al., 2017**), which uses
 532 Scatterwaite’s method to approximate the degrees of freedom for the t-test. In the Travel Time
 533 Experiment, we assessed the effect of starting volume of the patch and the travel time on number
 534 of harvests per patch, with random intercepts and random slopes for both variables across subjects
 535 (lme4 formula: *HarvestsPerPatch ~ PatchStartingVolume * TravelTime + (PatchStartingVolume +*
 536 *TravelTime || Rat)*). In all other foraging experiments, we assessed the effect of experimental
 537 condition on harvests per patch, with random intercepts and random effect of experimental
 538 condition across subjects (lme4 formula: *HarvestsPerPatch ~ Condition + (Condition | Rat)*).

539 We also tested whether rats overharvested relative to MVT predictions in each experiment,
 540 and whether the degree of overharvesting was different between conditions within each exper-
 541 iment. To do so, we subtracted the MVT predicted number of harvests in each patch from the
 542 observed number of harvests (see “Foraging Models” section for details on the calculation of
 543 the optimal number of harvests). Mixed effects models were used to fit an intercept and effect
 544 of experimental condition on the difference from optimal number of harvests (lme4 formula:
 545 *DifferenceFromOptimal ~ Condition + (Condition | Rat)*). In this model, an intercept greater than zeros
 546 would indicate that rats harvested more trials than was predicted by MVT, and a difference in the
 547 effect of task condition would indicate that the degree to which rats differed from optimal was
 548 affected by the task condition.

549 **Intertemporal choice task**

550 Rats were immediately transferred from the foraging task to the intertemporal choice task with no
 551 special training; rats were given three 2-hour sessions to learn the structure of the new task. This
 552 task consisted of a series of episodes that lasted 20 trials. At the beginning of each episode one
 553 lever was randomly selected as the shorter-sooner lever, yielding 40 μ L of reward following a 1 s
 554 delay. The other lever (larger-later lever) was initialized to yield a reward of 40, 80, or 120 μ L after a
 555 1, 2, 4 or 6 s delay. For the first 10 trials of each episode, only one lever extended, and rats were
 556 forced to press that lever to learn its associated reward value and delay. The last four forced trials
 557 (trials 7-10) were counterbalanced to reduce the possibility of rats developing a perseveration bias.

558 For the remaining 10 trials of each episode, both levers extended, and rats were free to choose the
 559 option they prefer. At the beginning of each trial, cue lights turned on above the lever indicating
 560 rats could now make a decision. Once the rat pressed the lever, the cue light turned off, and the
 561 delay period was initiated. A cue light turned on in the reward magazine at the end of the delay
 562 period, and rats received reward as soon as they entered the reward magazine. Reward magnitude
 563 was cued by light and tone. Following reward delivery, there was an ITI before the start of the next
 564 trial. At the completion of the episode, the levers retracted, and rats had to nose poke to begin the
 565 next episode, which reset the larger-later reward and delay.

566 Intertemporal choice data analysis

567 Intertemporal choice data was analyzed using a mixed effects logistic regression, examining the the
 568 effect of larger-later reward value, larger-later delay, and task condition on rats choices, with random
 569 intercepts and random effects for all three variables. This model was fit as a generalized linear
 570 mixed effects model using the lme4 package in R (lme4 formula: *ProbLL ~ RewardLL * DelayLL * Condition + (RewardLL + DelayLL + Condition || Rat)*; **Bates et al. 2015**). Post-hoc comparisons of
 571 interest were tested using the phia package in R (**De Rosario-Martinez, 2015**), using Holm's method
 572 to correct for multiple comparisons.

574 Foraging Models

575 All models were constructed as continuous time semi-markov processes. This provided a convenient
 576 way to capture the dynamics of timing in both tasks, such as slow delivery and consumption of
 577 reward (up to 6 s for the largest rewards). To model the foraging task, each event within the task (e.g.
 578 cues turning on/off, lever press, reward delivery, etc.) marked a state transition (abbreviated state
 579 space diagram in **Figure 3-Figure Supplement 1**. All state transitions were deterministic, except
 580 for decisions to stay in vs. leave the patch, which occurred in 'decision' states (the time between
 581 cues turning on at the start of the trial and rats performing a lever press or nosepoke). In decision
 582 states, a decision to stay in the patch transitioned to the handling time state, then reward state, ITI
 583 state, and to the decision state on the next trial. A decision to leave transitioned to the travel time
 584 state, then to the first decision state in the patch. Using the notation of **Bradtko and Duff (1995)**,
 585 the value of staying in state s , $Q(stay, s)$, is the reward provided for staying in state s , $R(stay, s)$, plus
 586 the discounted value of the next state:

$$Q(stay, s) = R(stay, s) + \gamma(stay, s) * V(s_{next})$$

587 where $\gamma(stay, s)$ is the discount applied to the value of the next state for staying in state s , and
 588 $V(s_{next})$ is the value of the next state in the patch. For all non-decision states, rats did not have the
 589 option to leave the patch, so for these states, $V(s) = Q(stay, s)$. For decision states, the value of the
 590 state was the greater of $Q(stay, s)$ and $Q(leave)$.

591 For simplicity, we assume the time spent in a given state is constant, calculated as the average
 592 amount of time a given rat spent in the state. Under this assumption, the reward in a given state,
 593 $R(stay, s)$, is equal to the reward rate provided over the course of the state, $r(s)$, multiplied by the
 594 time spent in that state $T(s)$, discounted according to discount factor β :

$$R(stay, s) = \frac{1 - e^{-\beta * T(s)}}{\beta} * r(s), \text{ and}$$

$$\gamma(stay, s) = e^{-\beta * T(s)}.$$

595 The value of leaving a patch, $Q(leave)$, was equal to the discounted value of the first state in the
 596 next patch, $V(s_{first})$:

$$Q(leave) = \gamma(leave) * V(s_{first})$$

597 where $\gamma(leave)$ is the discount factor applied to the next state in the first patch. Assuming no
 598 variance in the travel time τ , $\gamma(leave) = e^{-\beta * \tau}$. Per MVT, we assumed rats left patches at the first

600 state in the patch in which $Q(stay, s) \leq Q(leave)$. To model variability in the trial at which rats left
 601 patches, we added gaussian noise to $Q(leave)$. As decisions within each patch are not independent,
 602 the patch leaving threshold did not vary trial-by-trial, but rather patch by patch, such that the
 603 cumulative probability that a rat has left the patch by state s , $\pi(leave, s)$, was the probability that
 604 $Q(stay, s) \leq Q(leave) + Q(leave) * \epsilon$, where $\epsilon \sim \mathcal{N}(0, \sigma^2)$, with free parameter σ . ϵ scaled with $Q(leave)$
 605 to enable comparisons across conditions within experiments.

606 The optimal policy for a given set of parameters was found using value iteration (**Sutton and**
Barto, 1998). MVT predictions (maximization of undiscounted long-term reward rate) were deter-
 607 mined by fixing the discount factor $\beta = .001$ and assuming no decision noise ($\epsilon = 0$). MVT predictions
 608 were determined for each rat; the time spent in each state was taken from a given rat's data. For
 609 each model, we fit both group level parameters and individual parameters for each rat using an
 610 expectation-maximization algorithm (**Huys et al., 2011**).

612 To model subjective costs, a free parameter c representing an aversion to leaving the patch was
 613 subtracted from the leaving threshold (**Wikenheiser et al., 2013; Carter and Redish, 2016**):

$$Q_{cost}(leave) = -c + \gamma(leave) * V_{cost}(s_{first}).$$

614 To investigate whether nonlinear reward utility could explain rats' overharvesting behavior,
 615 we tested models in which the utility of a reward received in the task increased in a sublinear
 616 fashion with respect to the magnitude of the reward. Two different utility functions were tested: a
 617 power law function and a steeper constant relative risk aversion (CRRA) utility function that became
 618 increasingly risk averse with larger rewards, both with free parameter η :

$$\begin{aligned} Q_{utility}(stay, s) &= U(stay, s) + \gamma(stay, s) * V_{utility}(s_{next}) \\ 619 \\ U_{power}(stay, s) &= R(stay, s)^\eta, \text{ or} \\ 620 \\ U_{CRRA}(stay, s) &= \frac{R(stay, s)^{1-\eta} - 1}{1 - \eta}. \end{aligned}$$

621 To examine linear and nonlinear underestimation of post-reward delays, respectively, the time
 622 spent in post-reward delay (ITI) states was transformed, with free parameter α :

$$T_{post-linear}(s_{ITI}) = \alpha T(s_{ITI}), \text{ where } 0 < \alpha < 1, \text{ or}$$

$$T_{post-power}(s_{ITI}) = T(s_{ITI})^\alpha$$

624 Similarly, for overestimation of pre-reward delays, the handling time and travel time were trans-
 625 formed:

$$\begin{aligned} 626 \\ T_{pre-delay}(s_{HT}) &= \alpha T(s_{HT}), \text{ and} \\ \tau_{pre-delay} &= \alpha \tau, \text{ where } \alpha > 1. \end{aligned}$$

627 For the exponential discounting model, β was fit as a free parameter.

628 As standard hyperbolic discounting cannot conveniently be expressed recursively, this model
 629 was implemented using the μ Agents model described by **Kurth-Nelson and Redish (2009)**. The
 630 value functions of the overall model, $Q^{\mu\text{Agent}}(stay, s)$ and $Q^{\mu\text{Agent}}(leave)$, were the average of μ Agents,
 631 each with their own exponential discount factor β_i , and thus individual reward functions $R_i(stay, s)$,
 632 discount functions $\gamma_i(stay, s)$ and $\gamma_i(leave)$, and value functions $Q_i(stay, s)$, $Q_i(leave)$, and $V_i(s)$:

$$\begin{aligned} 633 \\ Q_i(stay, s) &= R_i(stay, s) + \gamma_i(stay, s) * V_i(s_{next}) \\ Q^{\mu\text{Agent}}(stay, s) &= \frac{1}{10} \sum_i R_i(stay, s) + \gamma_i(stay, s) * V_i(s_{next}) \\ 634 \\ Q_i(leave) &= \gamma_i(leave) * V_i(s_{first}) \\ 635 \\ Q^{\mu\text{Agent}}(leave) &= \frac{1}{10} \sum_i \gamma_i(leave) * V_i(s_{first}) \end{aligned}$$

636 If the μ Agent discount factors, β_i , are drawn from an exponential distribution with rate parameter
 637 $\lambda > 0$, the discounting function of the overall model approximated the standard hyperbolic discount
 638 function, $reward/(1 + k * delay)$, with discount rate $k = 1/\lambda$. This model was implemented using
 639 10 μ Agents with β_i equal to the 5%, 15%, ..., 95% quantile of the exponential distribution. The
 640 relationship of this implementation of the μ Agent model to the standard hyperbolic discount
 641 function is presented in **Figure 6**. k was fit as a free parameter.

642 The constant sensitivity discounting model was based on **Ebert and Prelec (2007)**. In this model,
 643 hyperbolic time preferences are produced via exponential discounting with insensitivity to longer
 644 delays. To implement this model, insensitivity to all time delays — the decision time, pre-reward
 645 delay, reward time, and post-reward delay, and travel time — was achieved using a power function,
 646 just as in the nonlinear post-reward delay model. This model was then equivalent to the exponential
 647 discounting model, replacing the time in each state $T(s)$ with a power function of the time in each
 648 state $T(s)^\alpha$.

649 Quasi-hyperbolic discounting was originally formulated for discrete time applications (**Laibson,**
 650 **1997**). We used the continuous time formulation from **McClure et al. (2007)**, in which the value
 651 functions of the overall model were the weighted sum of two exponential discount systems, a steep
 652 discounting β system that prefers immediate rewards and a slower discounting δ system, each with
 653 their own reward functions, $R_\beta(stay, s)$ and $R_\delta(stay, s)$, and discount functions $\gamma_\beta(stay, s)$, $\gamma_\beta(leave)$,
 654 $\gamma_\delta(stay, s)$, and $\gamma_\delta(leave)$:

$$Q_\beta(stay, s) = R_\beta(stay, s) + \gamma_\beta(stay, s) * V_\beta(s_{next})$$

$$Q_\beta(leave) = \gamma_\beta(leave) * V_\beta(s_{first})$$

$$Q_\delta(stay, s) = R_\delta(stay, s) + \gamma_\delta(stay, s) * V_\delta(s_{next})$$

$$Q_\delta(leave) = \gamma_\delta(leave) * V_\delta(s_{first})$$

658 The value function of the overall quasi-hyperbolic discounting model were:

$$Q^{quasi}(stay, s) = \omega * Q_\beta(stay, s) + (1 - \omega) * Q_\delta(stay, s)$$

$$Q^{quasi}(leave) = \omega * Q_\beta(leave) + (1 - \omega) * Q_\delta(leave)$$

660 where $0 < \omega < 1$ was the weight of the β system relative to the δ system. β , δ , and ω were all free
 661 parameters.

662 Intertemporal Choice Task Models

663 Similar to the foraging task, events within the intertemporal choice task marked state transitions,
 664 and all state transitions were deterministic except for decisions to choose the smaller-sooner option
 665 (SS) or larger-later option (LL), which occurred only in decision states (abbreviated state space
 666 diagram in **Figure 4–Figure Supplement 1**). From decision states, animals transitioned to delay,
 667 reward, and post-reward delay (ITI) states for the chosen option — the delay, reward and ITI for the
 668 SS and LL options were represented by separate states. The value of choosing SS or LL in decision
 669 state s is the discounted value of the next state, the following delay state:

$$Q(SS, s) = \gamma(s) * Q(SS\ Delay)$$

$$Q(LL, s) = \gamma(s) * Q(LL\ Delay)$$

671 The value of delay states were the discounted value of the reward state for that action, the value
 672 of reward states were the reward for that action plus the discounted value of the ITI state for that
 673 action, and the value of ITI states were the discounted value of the next decision state:

$$Q(SS\ Delay) = \gamma(SS\ Delay) * Q(SS\ Reward)$$

$$Q(SS\ Reward) = R(SS\ Reward) * \gamma(SS\ Reward) * Q(SS\ ITI)$$

$$Q(SS\ ITI) = \gamma(SS\ ITI) * V(s_{next\ dec})$$

676 where the value of the next decision state, $V(s_{next\ dec})$ is the greater of $Q(SS, s_{next\ dec})$ and $Q(LL, s_{next\ dec})$.
 677 Decisions were made assuming the value of $Q(SS, s)$ and $Q(LL, s)$ were represented as Gaussian
 678 distributions with noise that scaled with their magnitude. The probability of choosing the LL option
 679 was the probability that a random sample from the LL distribution was greater than a random
 680 sample from the SS distribution for that state:

$$p(\text{choose LL}, s) = 1 - \phi\left(\frac{Q(SS, s) - Q(LL, s))}{\sqrt{\sigma^2 * [Q(SS, s)^2 + Q(LL, s)^2]}}\right)$$

681 The nonlinear reward utility, biased time perception, and temporal discounting models were
 682 implemented as they were in the foraging task.

683 For the one-trial horizon discounting models, the value of choosing a given option was the
 684 discounted value of the reward on the current trial only, with delay d and reward r :

$$\begin{aligned} Q_{exp} &= e^{-\beta*d} * r \\ Q_{hyp} &= \frac{r}{1 + k * d} \\ Q_{cs} &= e^{-\beta*d^\alpha} * r \\ Q_{quasi} &= [e^{-\beta*d} * (1 - \omega) + e^{-\delta*d} * \omega] * r \end{aligned}$$

685 To calculate the probability of choosing the LL option, the same decision rule was used as in the
 686 all-future horizon model.

687 Model Comparison

688 All models had two parameters except for the constant sensitivity discounting model with three
 689 and the quasi-hyperbolic discounting model with four. To determine the model that provided the
 690 best fit to the data, while accounting for the increased flexibility of these models, we calculated
 691 the Bayesian Information Criterion over the group level parameters (iBIC) (*MacKay, 2003; Huys*
 692 *et al., 2011*). iBIC penalizes the log marginal likelihood, $\log p(D | \theta)$, which is the integral of the log
 693 likelihood of the data D over the distribution of group level parameters θ , for model complexity.
 694 Complexity is determined by the number of parameters k , and the size of the penalty depends on
 695 the total number of observations, n :

$$iBIC = \log p(D | \theta) + \frac{k}{2} \log(n).$$

696 As in *Huys et al. (2011)*, we use a Laplace approximation to the log marginal likelihood:

$$\log p(D | \theta) = -\frac{n}{2} \log(2\pi) * s + \sum_{i=1}^s p(D_i | \theta_i) p(\theta_i | \theta) - \frac{\sum_{i=1}^s \log \det(Hf(\theta_i))}{2}$$

697 where s is the number of subjects, and $Hf(\theta_i)$ is the hessian matrix of the likelihood for subject i at
 698 the individual parameters θ_i .

699 To compare the fit of the quasi-hyperbolic discounting model across the foraging and intertemporal
 700 choice tasks, a cross-validation method was used. Data from each task was separated into
 701 thirds. The quasi-hyperbolic discounting model was fit to 2 of the samples from each task using
 702 maximum likelihood estimation (fitting only individual parameters for each rat). The log likelihood
 703 of the data from the left out sample was evaluated. This process was repeated three times, leaving
 704 out each of the samples once, and we took the sum of the likelihood of the three left out samples.
 705 As the structure of variability was different between the foraging model (variability in the patch
 706 leaving threshold) and intertemporal choice models (noise in the estimates of SS and LL values), to
 707 compare the discount function fit to the foraging task on intertemporal choice data, a new noise
 708 parameter was fit to the intertemporal choice data (and vice-versa). We report the difference in
 709 the log likelihood of the data using parameters fit to the intertemporal choice task and of the log
 710 likelihood using parameters fit to the foraging task (*Figure 5*).

714 Acknowledgements

715 We thank Dr. Gary Aston-Jones for helpful discussions. This work was supported by NIH grant
 716 F31MH109286 (GAK) and the Princeton Program in Cognitive Science.

717 References

- 718 **Ainslie G.** Picoeconomics: The Strategic Interaction of Successive Motivational States Within the Person.
 719 Cambridge University Press; 1992.
- 720 **Bates D**, Mächler M, Bolker BM, Walker SC. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. 2015; 67(1):1–48. <http://cran.r-project.org/package=lme4%5Cnhttp://www.jstatsoft.org/index.php/jss/article/view/v067i01/v67i01.pdf>, doi: 10.18637/jss.v067.i01, citation Key Alias: BatesFittingLinearMixedEffects2015.
- 721 **Bateson M**, Kacelnik a. Rate currencies and the foraging starling: The fallacy of the averages revisited. *Behavioral Ecology*. 1996; 7(3):341–352. doi: 10.1093/beheco/7.3.341.
- 722 **Blanchard TC**, Hayden BY. Monkeys Are More Patient in a Foraging Task than in a Standard Intertemporal
 723 Choice Task. *PLOS ONE*. 2015 Feb; 10(2):e0117057. <http://dx.plos.org/10.1371/journal.pone.0117057>, doi:
 724 10.1371/journal.pone.0117057.
- 725 **Blanchard TC**, Pearson JM, Hayden BY. Postreward delays and systematic biases in measures of animal
 726 temporal discounting. *Proceedings of the National Academy of Sciences*. 2013; 110(38):15491–6. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC380845/>&tool=pmcentrez&rendertype=abstract, doi:
 727 10.1073/pnas.1310446110.
- 728 **Bogacz R**, Brown E, Moehlis J, Holmes P, Cohen JD. The physics of optimal decision making: a formal analysis of
 729 models of performance in two-alternative forced-choice tasks. *Psychological review*. 2006; 113(4):700–765.
 730 doi: 10.1037/0033-295X.113.4.700.
- 731 **Bradtko SJ**, Duff MO. Reinforcement learning methods for continuous-time Markov decision problems. In:
 732 *Advances in neural information processing systems*; 1995. p. 393–400.
- 733 **Carter EC**, Pedersen EJ, McCullough ME. Reassessing intertemporal choice: human decision-making is more
 734 optimal in a foraging task than in a self-control task. *Frontiers in Psychology*. 2015 Feb; 6. <http://journal.frontiersin.org/Article/10.3389/fpsyg.2015.00095/abstract>, doi: 10.3389/fpsyg.2015.00095.
- 735 **Carter EC**, Redish AD. Rats value time differently on equivalent foraging and delay-discounting tasks. *Journal of Experimental Psychology: General*. 2016; 145(9):1093–1101. <http://doi.apa.org/getdoi.cfm?doi=10.1037/xge0000196>, doi: 10.1037/xge0000196.
- 736 **Charnov EL**. Optimal foraging, the marginal value theorem. *Theoretical population biology*. 1976; 9:129–136.
 737 doi: 10.1016/0040-5809(76)90040-X.
- 738 **Constantino SM**, Daw ND. Learning the opportunity cost of time in a patch-foraging task. *Cognitive, affective & behavioral neuroscience*. 2015; 15(4):837–53. <http://www.ncbi.nlm.nih.gov/pubmed/25917000>, doi:
 739 10.3758/s13415-015-0350-y.
- 740 **Davidson JD**, El Hady A. Foraging as an evidence accumulation process. *PLoS Computational Biology*. 2019 Jul;
 741 15(7). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6682163/>, doi: 10.1371/journal.pcbi.1007060.
- 742 **De Rosario-Martinez H.** phia: Post-Hoc Interaction Analysis. . 2015; <https://CRAN.R-project.org/package=phia>.
- 743 **Ebert JEJ**, Prelec D. The Fragility of Time: Time-Insensitivity and Valuation of the Near and Far Future. *Management Science*. 2007 Jul; 53(9):1423–1438. <https://pubsonline.informs.org/doi/abs/10.1287/mnsc.1060.0671>, doi: 10.1287/mnsc.1060.0671.
- 744 **Gabaix X**, Laibson D. Myopia and Discounting. Cambridge, MA: National Bureau of Economic Research; 2017.
- 745 **Gallistel CR**, Gibbon J. Time, rate, and conditioning. *Psychological review*. 2000; 107(2):289–344. doi:
 746 10.1037/0033-295X.107.2.289.
- 747 **Hayden BY**, Pearson JM, Platt ML. Neuronal basis of sequential foraging decisions in a patchy environment.
 748 *Nature Neuroscience*. 2011; 14(7):933–939. <http://dx.doi.org/10.1038/nn.2856>, doi: 10.1038/nn.2856.

- 760 Hunter LE, Bornstein AM, Hartley CA. A common deliberative process underlies model-based planning and
761 patient intertemporal choice. *bioRxiv*. 2018 Dec; p. 499707. <https://www.biorxiv.org/content/10.1101/499707v2>, doi: 10.1101/499707.
- 763 **Huys QJM**, Cools R, Gölzer M, Friedel E, Heinz A, Dolan RJ, Dayan P. Disentangling the Roles of Approach,
764 Activation and Valence in Instrumental and Pavlovian Responding. *PLoS Computational Biology*. 2011 Apr;
765 7(4):e1002028. <http://dx.plos.org/10.1371/journal.pcbi.1002028>, doi: 10.1371/journal.pcbi.1002028.
- 766 **Huys QJM**, Eshel N, O'Nions E, Sheridan L, Dayan P, Roiser JP. Bonsai Trees in Your Head: How the Pavlovian
767 System Sculpts Goal-Directed Choices by Pruning Decision Trees. *PLOS Computational Biology*. 2012
768 Mar; 8(3):e1002410. <http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1002410>, doi:
769 10.1371/journal.pcbi.1002410.
- 770 **Kacelnik A**, Todd IA. Psychological mechanisms and the Marginal Value Theorem: effect of variability in travel
771 time on patch exploitation. *Animal Behaviour*. 1992; 43(2):313–322. doi: 10.1016/S0003-3472(05)80226-X.
- 772 **Kane GA**, Vazey EM, Wilson RC, Shenhav A, Daw ND, Aston-Jones G, Cohen JD. Increased locus coeruleus tonic
773 activity causes disengagement from a patch-foraging task. *Cognitive, Affective, & Behavioral Neuroscience*.
774 2017; <http://link.springer.com/10.3758/s13415-017-0531-y>, doi: 10.3758/s13415-017-0531-y.
- 775 **Killeen PR**. An Additive-Utility Model of Delay Discounting. *Psychological Review*. 2009 Jul; 116(3):602–
776 619. <http://ezp-prod1.hul.harvard.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=43581320&site=ehost-live&scope=site>, doi: 10.1037/a0016414.
- 778 **Kirby KN**. Bidding on the future: Evidence against normative discounting of delayed rewards. *Journal of
779 Experimental Psychology: General*. 1997; 126(1):54.
- 780 **Kolling N**, Behrens TEJ, Mars RB, Rushworth MFS. Neural Mechanisms of Foraging. *Science*. 2012; 336:95–98.
781 doi: 10.1126/science.1216930.
- 782 **Kurth-Nelson Z**, Redish AD. Temporal-difference reinforcement learning with distributed representations. *PLoS
783 One*. 2009; 4(10):e7362.
- 784 **Kuznetsova A**, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal
785 of Statistical Software*. 2017 Dec; 82(1):1–26. <https://www.jstatsoft.org/index.php/jss/article/view/v082i13>, doi:
786 10.18637/jss.v082.i13.
- 787 **Laibson D**. Golden Eggs and Hyperbolic Discounting. *The Quarterly Journal of Economics*. 1997
788 May; 112(2):443–478. <https://academic.oup.com/qje/article-lookup/doi/10.1162/00335539755253>, doi:
789 10.1162/00335539755253.
- 790 **MacKay DJ**. Information theory, inference and learning algorithms. Cambridge university press; 2003.
- 791 **Mazur JE**. Choice with Probabilistic Reinforcement: Effects of Delay and Conditioned Reinforcers. *Journal of the
792 Experimental Analysis of Behavior*. 1991 Jan; 55(1):63–77. <https://onlinelibrary.wiley.com/doi/abs/10.1901/jeab.1991.55-63>, doi: 10.1901/jeab.1991.55-63.
- 794 **McClure SM**, Ericson KM, Laibson DI, Loewenstein G, Cohen JD. Time Discounting for Primary Rewards. *Journal
795 of Neuroscience*. 2007 May; 27(21):5796–5804. <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.4246-06.2007>, doi: 10.1523/JNEUROSCI.4246-06.2007.
- 797 **McClure SM**, Laibson DI, Loewenstein G, Cohen JD. Separate Neural Systems Value Immediate and Delayed
798 Monetary Rewards. *Science*. 2004 Oct; 306(5695):503–507. <http://science.scienmag.org/content/306/5695/503>, doi: 10.1126/science.1100907.
- 800 **Namboodiri VMK**, Mihalas S, Marton TM, Hussain Shuler MG. A general theory of intertemporal decision-making
801 and the perception of time. *Frontiers in Behavioral Neuroscience*. 2014; 8:61. doi: 10.3389/fnbeh.2014.00061.
- 802 **Nonacs P**. State dependent behavior and the Marginal Value Theorem. *Behavioral Ecology*. 2001 Jan; 12(1):71–83.
803 [https://academic.oup.com/beheco/article/12/1/71/392385](http://academic.oup.com/beheco/article/12/1/71/392385), doi: 10.1093/oxfordjournals.beheco.a000381.
- 804 **Pearson JM**, Hayden BY, Platt ML. Explicit Information Reduces Discounting Behavior in Monkeys. *Frontiers
805 in Psychology*. 2010; 1. <http://journal.frontiersin.org/article/10.3389/fpsyg.2010.00237/abstract>, doi:
806 10.3389/fpsyg.2010.00237.

- 807 **Seinstra MS**, Sellitto M, Kalenscher T. Rate maximization and hyperbolic discounting in human experiential
808 intertemporal decision making. *Behavioral Ecology*. 2018 Jan; 29(1):193–203. <https://academic.oup.com/beheco/article/29/1/193/4633897>, doi: 10.1093/beheco/arx145.
- 810 **Shenhav A**, Straccia Ma, Cohen JD, Botvinick MM. Anterior cingulate engagement in a foraging context reflects
811 choice difficulty, not foraging value. *Nature neuroscience*. 2014; 17(July):1249–1254. <http://www.nature.com/doifinder/10.1038/nn.3771%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/25064851>, doi: 10.1038/nn.3771.
- 813 **Stephens DW**. Discrimination, discounting and impulsivity: a role for an informational constraint. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2002 Nov; 357(1427):1527–1537. <http://rstb.royalsocietypublishing.org/cgi/doi/10.1098/rstb.2002.1062>, doi: 10.1098/rstb.2002.1062.
- 816 **Stephens DW**, Kerr B, Fernandez-Juricic E. Impulsiveness without discounting: the ecological rationality
817 hypothesis. *Proceedings of the Royal Society B: Biological Sciences*. 2004 Dec; 271(1556):2459–2465.
818 <http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2004.2871>, doi: 10.1098/rspb.2004.2871.
- 819 **Stephens DW**. Decision ecology: foraging and the ecology of animal decision making. *Cognitive, Affective &*
820 *Behavioral Neuroscience*. 2008; 8(4):475–484. doi: 10.3758/CABN.8.4.475.
- 821 **Stephens DW**, Anderson D. The adaptive value of preference for immediacy: when shortsighted rules have
822 farsighted consequences. *Behavioral Ecology*. 2001; 12(3):330–339. <http://beheco.oxfordjournals.org/content/12/3/330.abstract>, doi: 10.1093/beheco/12.3.330.
- 824 **Stephens DW**, Krebs JR. Foraging theory. *Monographs in behavior and ecology*, Princeton, NJ: Princeton
825 University Press; 1986.
- 826 **Sutton RS**, Barto AG. Reinforcement Learning : An Introduction. MIT Press; 1998.
- 827 **Thaler R**. Some empirical evidence on dynamic inconsistency. *Economics Letters*. 1981 Jan; 8(3):201–207.
828 <http://www.sciencedirect.com/science/article/pii/0165176581900677>, doi: 10.1016/0165-1765(81)90067-7.
- 829 **Wikenheiser AM**, Stephens DW, Redish aD. Subjective costs drive overly patient foraging strategies in rats on an
830 intertemporal foraging task. *Proceedings of the National Academy of Sciences*. 2013; 110(20):8308–13. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3780236/>, doi: 10.1073/pnas.1220738110.
- 833 **Zauberman G**, Kim BK, Malkoc SA, Bettman JR. Discounting Time and Time Discounting: Subjective Time
834 Perception and Intertemporal Preferences. *Journal of Marketing Research*. 2009 Aug; 46(4):543–556. <https://doi.org/10.1509/jmkr.46.4.543>, doi: 10.1509/jmkr.46.4.543.

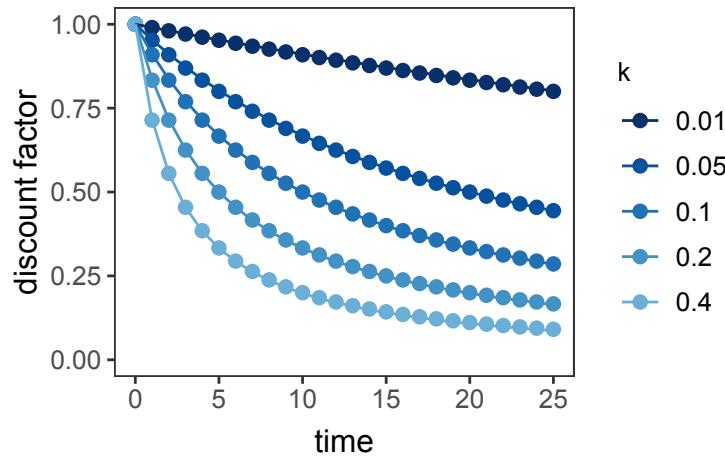


Figure 6. Discount function of the μ Agent hyperbolic discounting model vs. standard hyperbolic discounting. Points represent standard hyperbolic discounting function, $1/(1 + k * \text{time})$. Lines represent the μ Agent discount function in which the discount factor for each of the 10 μ Agents was equal to the 5%-95% quantile of an exponential distribution with rate parameter $\lambda = 1/k$.

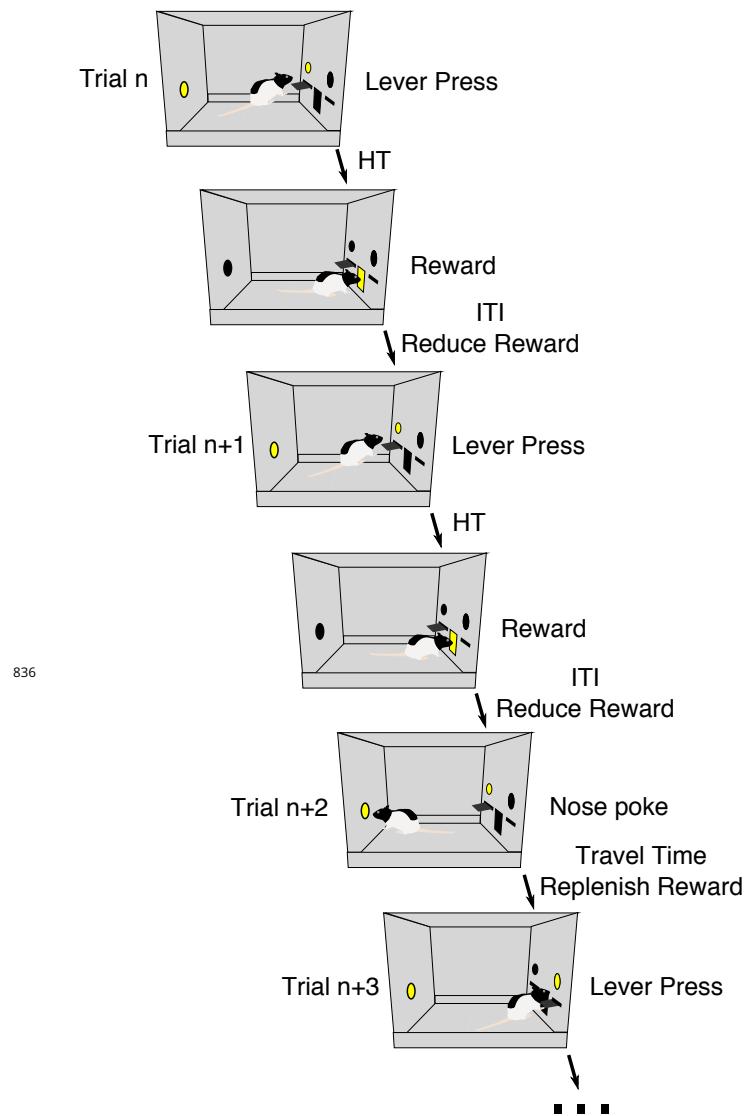


Figure 1–Figure supplement 1. Diagram of the foraging task. Rats press a lever to harvest reward from the patch then receive reward in an adjacent port following a handling time delay. After receiving reward, there is an inter-trial interval (post-reward delay) before rats can make their next decision. Rats can leave the patch by nose poking in the back of the chamber (trial n+2), which initiates a delay simulating time to travel to the next patch, after which, rats can harvest from a new replenished patch.

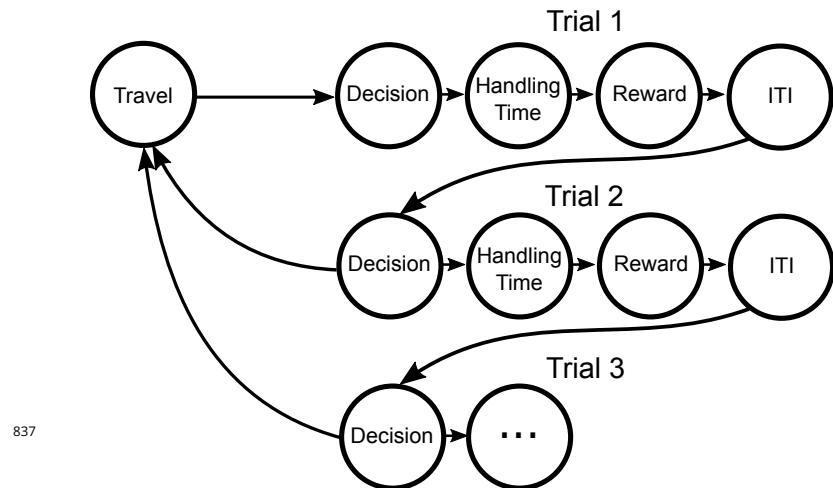


Figure 3–Figure supplement 1. State space diagram for the semi-markov model of the foraging (top) and intertemporal choice (bottom) tasks. In the foraging task, decisions to stay vs. leave are made in Decision states. A Decision to stay causes a transition to the handling time, then reward, ITI, and to the Decision state on the next trial. Reward is delivered uniformly throughout time spent in the each reward state. Reward depletion is achieved via shorter time spent in the reward state (resulting in longer stay in the ITI state). A Decision to leave causes a transition to the travel state, then to the first trial of the patch. In the intertemporal choice task, decisions made in Decision states cause transition to the Delay, Reward, and ITI states for the option chosen (either SS or LL), then back to the next Decision state. The model consisted of 10 consecutive trials—the number of free choice trials.

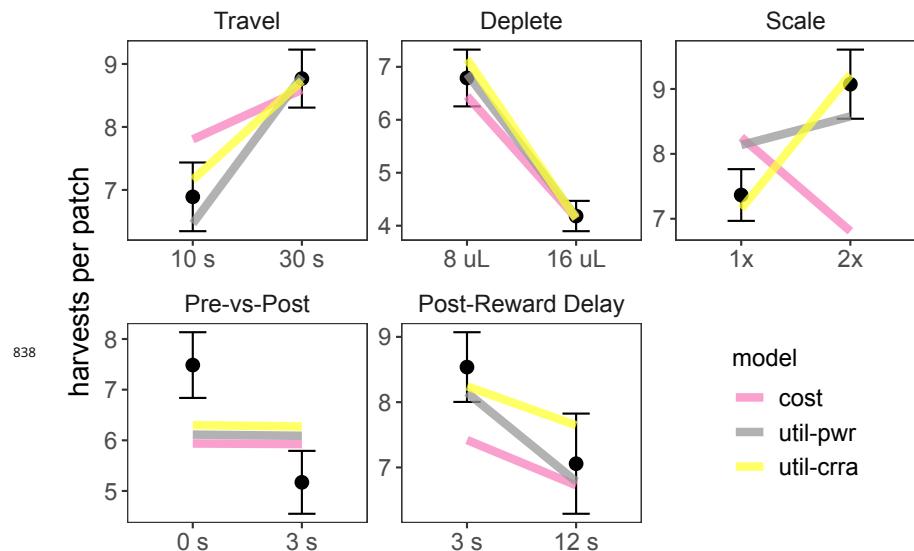


Figure 3–Figure supplement 2. Predictions of the best fit subjective cost and nonlinear reward utility models (power law = util-pwr; constant relative risk aversion = util-CRRA). Black points and error bars represent mean \pm standard error of observed behavior. Colored lines represent the mean model predicted behavior across rats.

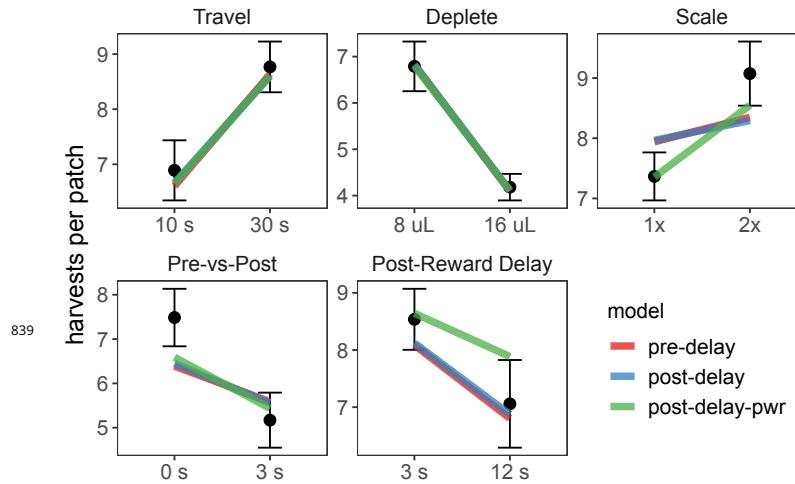


Figure 3–Figure supplement 3. Predictions of the best fit models of overestimation of pre-reward delays (pre-delay), linear underestimation of post-reward delays (post-delay), and nonlinear underestimation of post-reward delays (post-delay-nonlinear). Points and errorbars are the mean \pm standard deviation of rat behavior, colored lines represent the mean model predicted number of harvests across all rats.

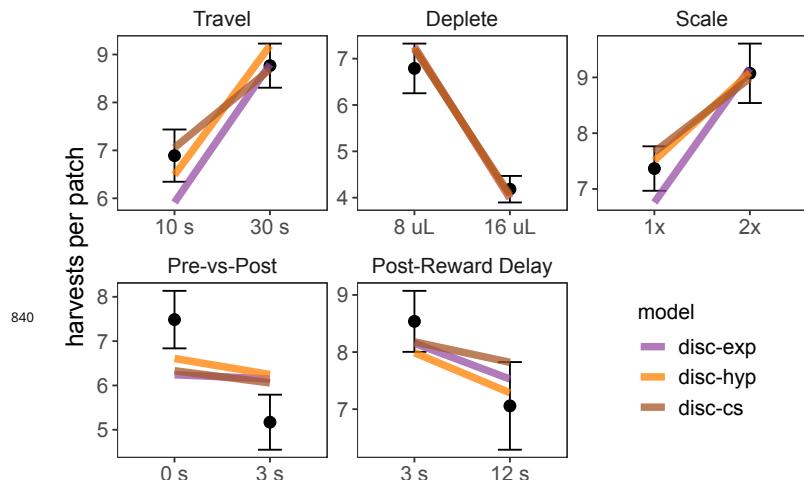


Figure 3–Figure supplement 4. Predictions of the best fit exponential discounting model (disc-exp), hyperbolic discounting model (disc-hyp), and constant sensitivity discounting model (disc-cs). Points and error bars are the mean \pm standard deviation of rat behavior; colored lines represent the mean predicted number of harvests across all rats.

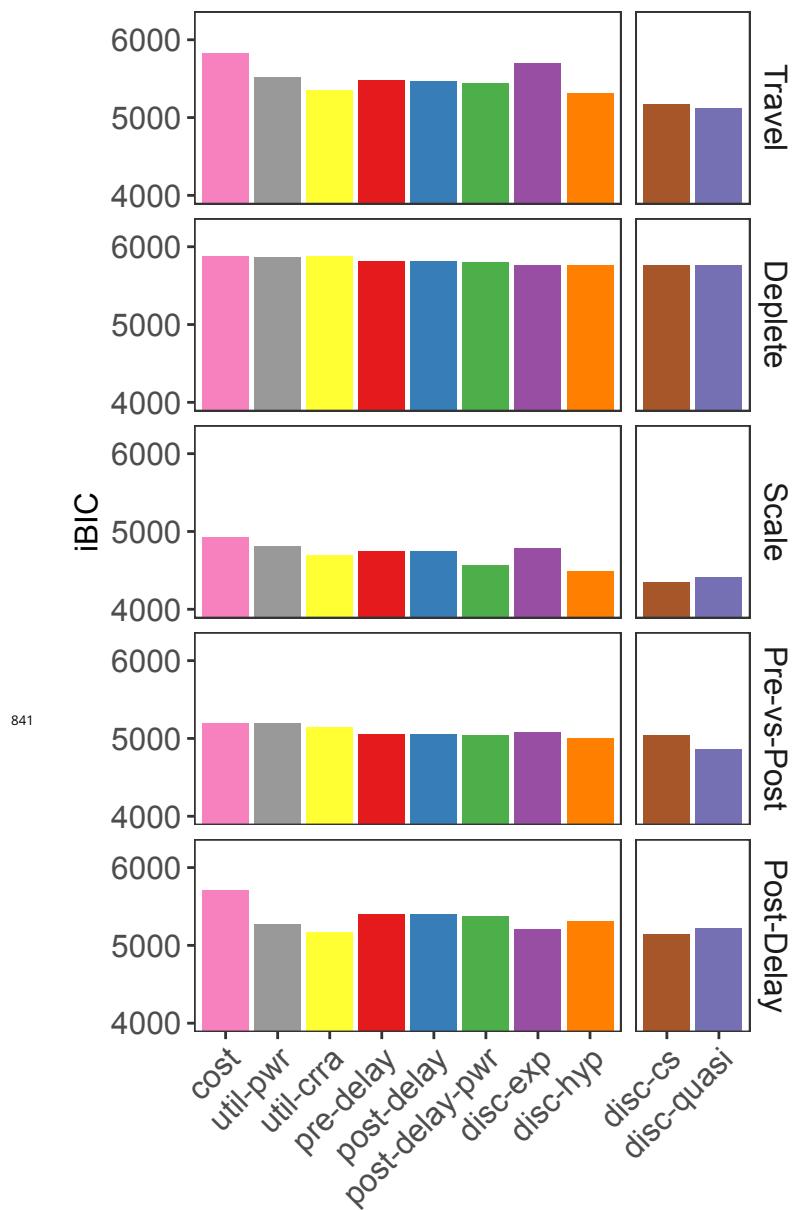


Figure 3-Figure supplement 5. iBIC for each model for each foraging experiment. Cost = subjective cost model, util-pwr and util-crra = nonlinear reward utility with power and CRRA function respectively, pre-del = linear overestimation of pre-reward delays, post-del = linear underestimation of post-reward delays, post-del-pwr = underestimation of post-reward delays according to a power function, disc-exp = exponential discounting, disc-hyp = hyperbolic discounting, disc-CS = constant sensitivity discounting, disc-quasi = quasi-hyperbolic discounting.

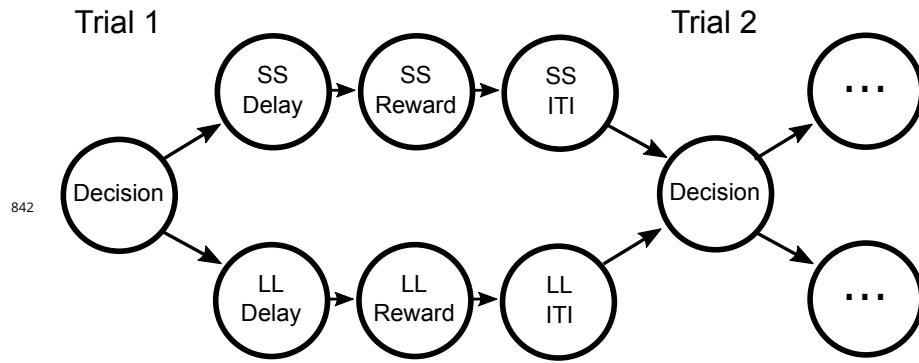


Figure 4-Figure supplement 1. State space diagram of the intertemporal choice task.

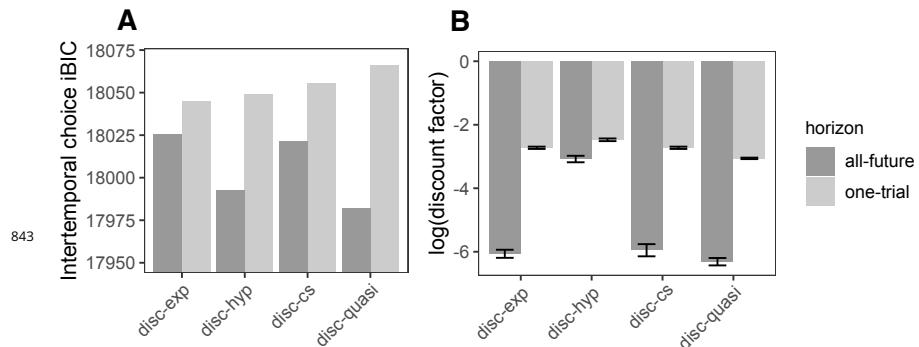


Figure 4-Figure supplement 2. A) iBIC for the full horizon and one-trial horizon discounting models B) Measured log-transformed discount factors for the full horizon and one-trial horizon discounting models. Bars and errorbars represent mean \pm standard error.