

Mechanisms of overharvesting in patch foraging

Gary A. Kane (gkane@princeton.edu), Aaron M. Bornstein (aaronmb@princeton.edu),
Nathaniel D. Daw (ndaw@princeton.edu), Jonathan D. Cohen (jdc@princeton.edu)

Department of Psychology and Neuroscience Institute, Princeton University
Princeton, NJ 08544 USA

Robert C. Wilson (bob@email.arizona.edu)

Department of Psychology, University of Arizona
Tucson, AZ 85721 USA

Amitai Shenhav (amitai_shenhav@brown.edu)

Department of Cognitive, Linguistic, and Psychological Sciences, Brown University
Providence, RI 02912 USA

Abstract

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 decide whether to stick with a depleting reward vs search for a new source. The optimal strategy in patch foraging problems, described by the Marginal Value Theorem (MVT; Charnov, 1976), is to leave the depleting patch when the local reward rate within a patch matches the overall long-run reward rate. Many species of animals, ranging from birds to rodents, monkeys, and humans, adhere to this policy in important respects, but tend to overharvest, or stick with the depleting resource too long. Here we attempt to determine the cognitive biases that underlie overharvesting in one of these species (the rat). We characterized rat behavior in response to two basic manipulations in patch foraging tasks: to travel time between patches and depletion rate, and two novel manipulations to the foraging environment: the size of reward and length of delays, and placement of delays (pre- vs. post-reward). In response to the basic manipulations, rats qualitatively followed predictions of MVT, but stayed in patches for longer than is predicted. In the latter two manipulations, rats deviated from predictions of MVT, exhibiting changes in behavior not predicted by MVT. We formally tested whether four separate cognitive biases – subjective costs, decreasing marginal utility for reward discounting of future reward, and ignoring post-reward delays – could explain overharvesting in the former two manipulations and deviations from MVT in the latter two. All of the biases tested explained overharvesting behavior in the former contexts, but only one bias – in which rats ignore post-reward delays – also explained deviations from MVT due to larger rewards with longer delays and due to introduction of a pre-reward delay. Our results show that multiple biases can explain certain aspects of overharvesting behavior, and, while foraging behavior may be the result of the use of multiple biases, inaccurate estimation of post-reward delays likely contributes to overharvesting.

Keywords: foraging; decision-making, subjective utility; delay discounting

Introduction

Patch foraging refers to situations in which one must decide when to leave a depleting resource patch to search for a

new, likely richer one, that comes at the cost of time and/or effort. The optimal solution in patch foraging is given by the Marginal Value Theorem (MVT; Charnov, 1976): leave when the local reward rate within a patch depletes below the global reward rate across all patches – the average reward rate for the environment. MVT makes two main predictions: i) in patches that contain more reward than average, stay longer to exploit such reward, and ii) when the cost of searching for a new patch is greater (e.g. the time or effort required to travel to a new patch is greater), stay longer in all patches. Many animals, ranging from invertebrates to birds to mammals, qualitatively follow predictions of MVT (Stephens & Krebs, 1986). However, in most tests, animals, including rats, monkeys, and humans, tend to stay in patches longer than is predicted by MVT (Constantino & Daw, 2015; Hayden, Pearson, & Platt, 2011; Nonacs, 1991; Stephens & Krebs, 1986).

Hypotheses to explain overharvesting include common biases in intertemporal choice, such as i) subjective costs, such as an aversion to leaving the patch (Carter & Redish, 2016; Wikenheiser, Stephens, & Redish, 2013); ii) decreasing marginal utility in which large rewards available in a new patch are not viewed as proportionally larger than the smaller, depleted rewards available in the current patch (Constantino & Daw, 2015); iii) discounting future rewards, in which the value of large rewards available in a new patch are discounted by virtue of being available later, above and beyond the time it takes to travel to the new patch (Blanchard, Pearson, & Hayden, 2013; Carter & Redish, 2016; Constantino & Daw, 2015); and iv) ignoring post-reward delays, which causes overestimation of reward rate within the patch due to inaccurate estimation of the time taken to obtain reward (Bateson & Kacelnik, 1996; Blanchard et al., 2013; Carter & Redish, 2016; Gallistel & Gibbon, 2000; Kacelnik, 1997). Although overharvesting is widely observed, there have been few direct investigations into the underlying mechanisms. In this paper, we directly test these hypotheses to rat foraging behavior in an operant chamber based patch foraging task.

First, we characterized rat foraging behavior in response to four manipulations to the foraging environment: to travel

time between patches, rate of reward depletion within patches, scale of reward size and length of delay, and placement of delays (pre- vs. post-reward). Next, we fit formal models representing the four hypotheses to rats' behavior to examine how well each hypothesis explained foraging behavior across all manipulations.

Methods

Animals

Adult Long-Evans rats were used (Charles River, Kingston, NY; $n = 8$). Rats were housed on a reverse 12 h/12 h light/dark cycle (lights off at 7 A.M.). All behavioral testing was conducted during the dark period. Throughout behavioral testing, rats were food restricted to maintain a weight of 85-90% ad-lib feeding weight, and were given ad-lib access to water. All procedures were approved by the Princeton University Institutional Animal Care and Use Committee.

Operant Foraging Task

This task simulated foraging in a patchy environment, resembling the task used with monkeys by Hayden et al (2011). On a series of trials performed in a standard operant chamber (Med Associates, St. Albans, VT), rats had to repeatedly decide to stay in a patch to continue harvesting a depleting reward source or leave the patch to travel to a new, full patch, incurring a cost of time to travel to a new patch. Rats' decided to harvest from a patch by pressing an activated lever on one side of the front of the chamber, or to travel to a new, full patch by nosepoking at the back of the chamber and then returning to a newly activated lever on the other side of the front of the chamber. To cue the beginning of a trial, lights above the activated lever and the nosepoke illuminated, indicating that the rat could decide to harvest reward from the activated patch (lever press) or to travel to a new patch (nosepoke). The time from start of trial to the decision was recorded as decision time (DT). If the rat pressed the lever to harvest from the activated patch, a cue light turned on in the reward magazine next to the lever following a short handling time delay (HT), and liquid sucrose was delivered when the rat's head entered the magazine. An intertrial interval (ITI) began as soon as the rat entered the reward magazine. To control the reward rate within the patch, the length of the ITI was adjusted based on the DT of the current trial, such that the length of all lever press trials was equivalent. If the rat nosepoked to leave the patch, the lever retracted for a delay period, simulating the travel time, after which, the lever on the opposite side of the chamber extended, representing a new patch from which the rat could harvest.

Each manipulation (travel time, depletion rate, scale, and delay placement) was conducted in separate experiments, with two conditions in each experiment. Rats were trained on each condition for 5 days, and tested for a subsequent 5 days. Conditions within each experiment were

counterbalanced. Details regarding reward size and timing for each experiment can be found in Figure 1. T-tests or ANOVA with repeated measures were used to compare the number of harvests per patch, a proxy for time in the patch, across conditions.

Models

All models were constructed as Markov Decision Processes. States were represented as trials within patches. A decision to stay in the patch (i.e. harvest from the patch) provided reward for staying in state s , $r_{stay,s}$, and caused transition to state $s + 1$. A decision to leave the patch resulted in travel time delay, τ , followed by the first reward in the next patch, r_{leave} , and associated ITI following the reward, ITI_{leave} . We fit three models based on MVT: a model incorporating a constant subjective cost (subjective cost), a model that accounted for diminishing marginal returns for larger rewards (subjective utility), and a model ignoring post-reward delays, as well as a delay discounting model.

For each of the MVT models, we calculated the value of staying in the patch in state s , $V_{stay,s}$, as the reward rate within the patch,

$$T_{stay,s} = DT_{stay,s} + HT + ITI_{stay,s} ,$$

$$V_{stay,s} = \frac{r_{stay,s}}{T_{stay,s}} ,$$

and the value of leaving the patch in state s , $V_{leave,s}$, as the cumulative reward rate across patches,

$$T_{leave,s} = DT_{leave,s} + \tau + ITI_{leave} ,$$

$$R_s = p_{stay,s} * r_{stay,s} + p_{leave,s} * r_{leave} ,$$

$$T_s = p_{stay,s} * T_{stay,s} + p_{leave,s} * T_{leave,s} ,$$

$$V_{leave} = \frac{\sum_s p_s * R_s}{\sum_s p_s * T_s} ,$$

where R_s and T_s was the average reward and average time for state s , $p_{stay,s}$ was the probability of choosing to stay in state s , and p_s was the probability of being in state s . Optimal behavior was to leave the patch when $V_{leave} \geq V_{stay}$ (i.e. when the long-run average reward rate is greater than the local reward rate in the patch). To model rats' behavior, patch leaving distributions were assumed to be normally distributed with respect to $V_{leave} - V_{stay}$, with mean $\mu = 0$ (i.e. $V_{leave} = V_{stay}$) and variance σ^2 , a free parameter.

To account for subjective costs, a constant, c , representing an aversion to leaving the patch, was added to the model, such that the patch leaving distribution was normally distributed with respect to $V_{leave,s} - V_{stay,s} - c$.

For the subjective utility model, the utility for taking action a in state s increased monotonically, but sublinear to the size of the reward, according to a power utility function, dependent on a free parameter, η ,

$$u_{a,s} = r_{a,s}^\eta ,$$

$$V_{stay,s} = \frac{u_{stay,s}}{T_{stay,s}} ,$$

$$U_s = p_{stay,s} * u_{stay,s} + p_{leave,s} * u_{leave} ,$$

$$V_{leave} = \frac{\sum_s p_s * U_s}{\sum_s p_s * T_s}$$

For the ignoring post-reward delays model, delays that occur after receiving reward, but before a decision was made on the next trial (e.g. ITI after reward and DT prior to making next decision), T_{post} , were treated differently than time delays that occurred between the decision and receiving a reward (e.g. handling time delay between lever press and reward, or travel time delay between nosepoke and first reward in the next patch). We tested multiple functions for how post-reward delays might have been treated, all in which the increase in perceived time increased monotonically, but sublinear to actual time, including a linear function with slope < 1, a power function, and an exponential function. The exponential function provided the best fit across all experiments, and was used for further testing:

$$T_{post,stay} = \frac{1 - e^{-\alpha(DT_{stay,s} + ITI_{stay,s})}}{\alpha}$$

$$T_{post,leave} = \frac{1 - e^{-\alpha(DT_{leave} + ITI_{leave})}}{\alpha}$$

$$V_{stay,s} = \frac{r_{stay,s}}{HT + T_{post}}$$

$$T_s = p_{stay,s} * (HT + T_{post,stay}) + p_{leave,s} * (\tau + T_{post,leave})$$

$$V_{leave} = \frac{\sum_s p_s * R_s}{\sum_s p_s * T_s}$$

Whereas MVT optimizes all future reward, the delay discounting model, a hyperbolic discounting model, optimizes discounted future reward (i.e. it similarly optimizes future reward, but with less weight to rewards that occur further in the future):

$$d(t, k) = 1 / (1 + k * t)$$

$$V_{stay,s} = d(HT, k) \left[r_{stay,s} + d(ITI_{stay,s}, k) \sum_{s'} p(s' | stay, s) \sum_a p_{a,s'} V_{a,s'} \right]$$

$$V_{leave,s} = d(\tau, k) \left[r_{leave} + d(ITI_{leave}, k) \sum_{s'} p(s' | stay, s) \sum_a p_{a,s'} V_{a,s'} \right]$$

where $d(t, k)$ was the hyperbolic discount function of time t , with a free parameter, k . $p(s' | a, s)$ was the conditional probability of being in future state s' given action a was taken in state s , $p_{a,s'}$ was the probability of taking action a in future state s' , and $V_{a,s'}$ was the value of for taking action a in future state s' .

As the discount parameter, k , approached zero (no discounting of future reward), this model converged to MVT; that is, it sought to maximize all future reward. As k increases, future rewards are discounted, such that i) the value of large rewards in a new patch are discounted above and beyond the travel time between patches, and ii) the model sought to maximize reward into the future, but over shorter periods of time.

For all models, one set of parameters was fit to each animal per experiment, to maximize the likelihood of the data from both conditions in that experiment. To test whether the model could explain rat overharvesting behavior in each experiment, we generated predicted patch leaving distributions from the best fit parameters for each

model, then perform a repeated measures ANOVA, to test whether there is an interaction between model predictions and observed behavior (i.e. whether the effect of each experimental manipulation was different between model predictions and observed behavior).

Results

Foraging Behavior

Rats were first tested on a manipulation of travel time. With longer travel time, the long-run average reward rate is lower, thus MVT predicts rats should stay in patches longer. Within behavioral sessions, rats encountered three different patch types, which started with varying amount of reward (60, 90, or 120 μ L) and depleted by the same rate (8 μ L). Between sessions, rats were tested on either a 10 s or 30 s travel time delay following their decision to leave the patch. As predicted by MVT, rats stayed longer in patch types that started with larger reward volume, indicated by more harvests per patch, $F(2, 14) = 3145$, $p < .001$, and rats stayed longer in all patches with longer travel time, $F(1, 7) = 71.4$, $p < .001$. However, rats overharvested, staying longer in all patches than is predicted by MVT (Figure 1A).

Next, rats were tested on a manipulation of depletion rate. Quicker reward depletion causes the local reward rate to deplete to the long-run average reward rate quicker, such that MVT predicts earlier patch leaving. Within sessions, rats encountered a single patch type (starting volume of 90 μ L), which depleted at a rate of either 8 or 16 μ L/trial, tested between sessions. As predicted by MVT, rats left patches earlier when they depleted more quickly, $t(7) = 15.835$, $p < .001$. But, again, rats stayed in patches longer than is predicted by MVT (Figure 1B).

Rats were then tested on a manipulation of the scale of rewards and time. In one condition, the size of rewards and length of delays was twice that of the other: patches started with 90 or 180 μ L of reward, depleted at a rate of 8 or 16 μ L/trial, and travel time between patches was 10 or 20 s. Both reward rate within the patch and reward rate across patches were equivalent in the two conditions; thus, MVT predicts no change in behavior. Contrary to predictions of MVT, rats stayed in patches significantly longer when given larger rewards with longer delays, $t(7) = 10.039$, $p < .001$. And, again, rats overharvested in both conditions (Figure 1C).

Lastly, rats were tested on a manipulation of the placement of delays. In one condition, rats experienced no pre-reward delay, and a long post-reward delay (ITI \sim 10 s, adjusted based on DT). In the other condition, rats experienced a 3 s pre-reward delay, and shorter post-reward delay (ITI \sim 7 s). The duration of each trial did not change, so both the local reward rate within the patch and long-run average reward rate across patches were equivalent between the conditions, and MVT predicts no change in behavior. Rats overharvested in both conditions, but they left patches earlier when part of the delay occurred prior to the reward, $t(7) = 7.453$, $p < .001$ (Figure 1D).

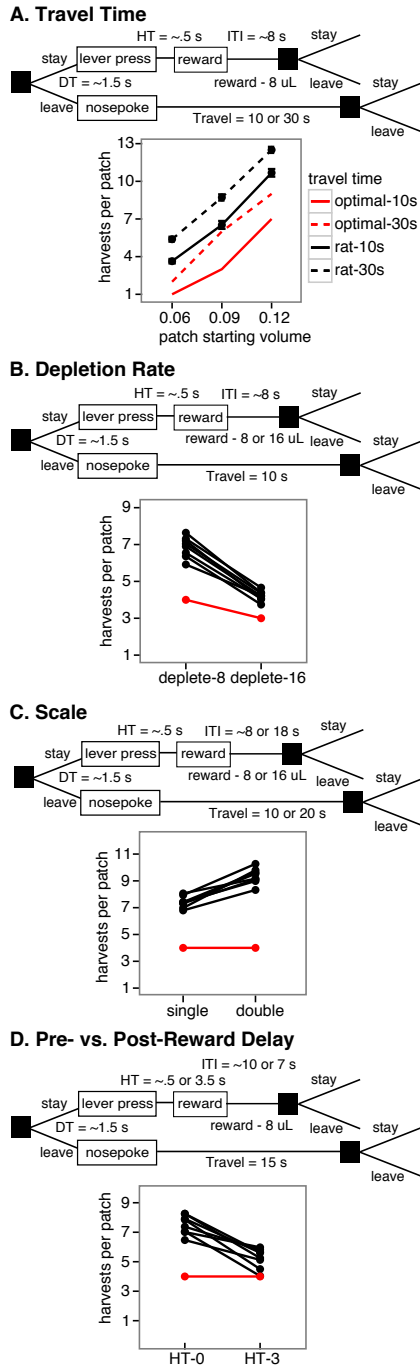


Figure 1: Diagram of each foraging experiment and behavioral data. In diagrams, black boxes represent the start of a trial, at which a decision to lever press or nosepoke must be made. DT = decision time, HT = handling time, ITI = intertrial interval. In graphs, black points and lines represent rat data, and red points and lines the optimal behavior predicted by MVT. A) Points represent the mean number of lever presses in each patch from each animal, error bars representing standard error are obstructed by the points. B-D) Each point is the mean number of lever presses in each patch for a single rat, with lines connecting each rat's behavior in the two conditions.

Models of overharvesting

We first tested a model that includes a subjective cost to foraging – a constant that represents a bias towards staying in the patch. Predictions from the model, fit to each rat, are presented in Figure 2. Qualitatively, this model explained rat behavior on the travel time and depletion rate experiments well, producing a predicted number of harvests per patch similar to that exhibited by the rats. However, there was a significant interaction between travel time and predicted vs. observed behavior, $F(1, 7) = 7.391, p = .030$, indicating a difference between how the model vs. the rats responded to the change in travel time. This is likely driven by the model predicting slightly earlier patch leaving in the 30 s travel time relative to rats' behavior. The interaction between depletion rate and predicted vs. observed behavior was not significant, $F(1, 7) = .124, p = .735$.

As this model only allows for a constant change in the reward rate threshold to leave patches, it is unlikely to account for behavior in which rats select a different threshold between conditions. When rats were given longer rewards with longer delays, they stayed in patches longer, allowing patches to deplete to a lower reward rate before leaving. Similarly, when a pre-reward delay was introduced, rats left patches earlier, at a higher reward rate. The model failed to account for both of these effects (interaction between scale x predicted vs. observed behavior, $F(1, 7) = 58.43, p < .001$; delay x predicted vs. observed behavior, $F(1, 7) = 48.79, p < .001$).

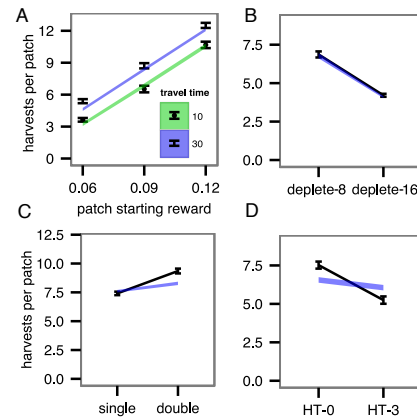


Figure 2: Predictions of the subjective cost model for the A) travel time, B) depletion rate, C) scale, and D) pre- vs. post-reward delay experiments. Black points and errorbars represent the mean number of harvests per patch \pm standard error. Colored lines represent the average model predicted number of harvests. The width of the colored line represents the standard error of the predicted number of harvests. There were significant interactions between model predictions and observed behavior in the travel time (A), scale (C), and pre vs. post-reward delay (D) experiments.

We next tested whether diminishing marginal returns could explain overharvesting (Figure 3). Under this

hypothesis, large rewards in a new patch were not valued as proportionally larger to smaller rewards in the current, depleting patch. Predictions from the subjective utility model are presented in Figure 3. As did the subjective cost model, the subjective utility model qualitatively explained overharvesting behavior in the travel time and depletion rate experiments. This was supported by the lack of a significant interaction between travel time and predicted vs. observed behavior, $F(1, 7) = 4.501, p = .072$, although there was a significant interaction between depletion rate and predicted vs. observed, $F(1, 7) = 14.12, p = .007$.

In the scale experiment, the subjective utility model should estimate a lower reward rate in the environment with larger rewards, and thus predict later patch leaving. However, this model could not explain both general overharvesting, as well as the change in behavior due to scale, $F(1, 7) = 112, p < .001$. Additionally, this model is insensitive to the placement of delays, and failed to predict that rats would leave patches earlier when a pre-reward delay was introduced, $F(1, 7) = 77.22, p < .001$.

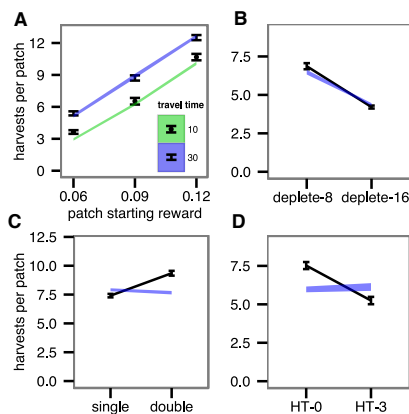


Figure 3: Predictions of the subjective utility model for the A) travel time, B) depletion rate, C) scale, and D) pre- vs. post-reward delay experiments. There were significant interactions between model predictions and observed behavior in the depletion rate (B), scale (C), and pre- vs. post-reward delay (D) experiments.

Next, we tested whether a delay discounting model that considers future rewards could account for rat overharvesting behavior (Figure 4). As rewards are discounted to the future, the value of the first reward in a new patch was discounted due to the travel time between patches, and the model sought to maximize future rewards over a shorter period of time. The discounting model accurately predicted overharvesting behavior in both travel times; interaction between travel time and predicted vs. observed behavior was not significant, $F(1, 7) = .050, p = .830$. This model also predicted earlier patch leaving when reward in the patch depleted quicker, but there was significant interaction between depletion rate and predicted vs. observed behavior, $F(1, 7) = 16.780, p = .005$, indicating that the model-predicted change in behavior is different

from the change in behavior exhibited by rats.

In the scale experiment, when comparing larger rewards with longer delays to smaller rewards with shorter delays, the larger rewards would be discounted to a greater extent. Thus, in this model, the estimate of the overall reward rate would be lower in the environment with larger rewards, predicting that rats would stay longer in this environment. Indeed, this model did predict that rats would stay in patches longer when given larger rewards with longer delays, and the interaction between scale and predicted vs. observed behavior was not significant, $F(1, 7) = .482, p = .510$. This model also should place lower value on rewards in the patch when there is a longer delay between decision to harvest and obtaining reward. However, there was a significant interaction between pre- vs. post-reward delay conditions and predicted vs. observed behavior, $F(1, 7) = 34.650, p < .001$.

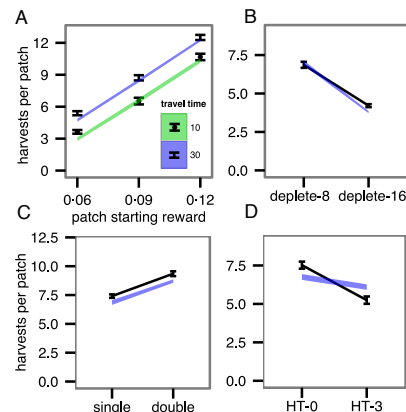


Figure 4: Predictions of the delay discounting model for the A) travel time, B) depletion rate, C) scale, and D) pre- vs. post-reward delay experiments. There were significant interactions between model predictions and observed behavior in the depletion rate (B) and pre- vs. post-reward delay (D) experiments.

Lastly, we tested whether ignoring post-reward delays could explain rats' overharvesting behavior. In this model, time delays that occur after receiving reward, before a decision is made on the next trial (e.g. ITI after reward and DT prior to making next decision), were treated differently than time delays that occur between making a decision and receiving a reward (e.g. handling time delay between lever press and reward, or travel time delay between nosepoke and first reward in the next patch). Time delays that occur after the reward, and before the next decision are assumed to increase monotonically, but sublinear relative to actual time, according to an exponential function. In this model, underestimation of the ITI would cause overestimation of reward rate, and overharvesting. Additionally, in the scale experiment, longer delays would cause greater overestimation of reward rate, and would predict that rats should stay in patches longer when given larger rewards with longer delays. In the pre- vs. post-reward delay

experiment, when the pre-reward delay was introduced, post-reward delays were shorter. In this model, shorter post-reward delays would lead to less overestimation of reward rate, and earlier patch leaving.

This model qualitatively explained overharvesting in all four experiments. Additionally, there were no significant interactions between task manipulations and predicted vs. observed behavior (travel time, $F(1, 7) = .416$, $p = .539$; depletion rate, $F(1, 7) = 4.691$, $p = .067$; scale of reward and time, $F(1, 7) = .047$, $p = .835$; pre- vs. post-reward delay, $F(1, 7) = 1.985$, $p = .202$), indicating that there were no differences between rats change in behavior due to experimental manipulation and model predicted change in behavior in all four experiments.

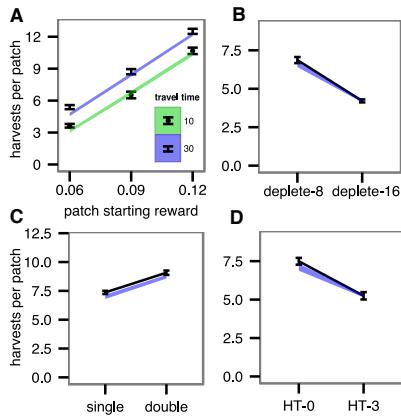


Figure 5: Predictions of the ignore post-reward delays model for the A) travel time, B) depletion rate, C) scale, and D) pre- vs. post-reward delay experiments. Interactions between model predictions and observed behavior were not significant in any of the four experiments.

Discussion

We characterized patch foraging behavior of one of these species, rats, in a variety of foraging environments, and examined the computational mechanisms of overharvesting. We found that rats, like humans (Constantino & Daw, 2015), followed the primary qualitative predictions of MVT, leaving patches earlier when the rate of depletion was quicker, and staying longer in patches when travel time was longer. However, as has consistently been observed in other species, they overharvested (or stayed longer in patches than is predicted by MVT). Furthermore, rats deviated from predictions of MVT in other ways, staying longer in patches that provided larger rewards with longer delays, and leaving patches earlier when delays occurred between the decision to harvest from the patch and receiving reward. To examine the cognitive biases that underlie overharvesting, we fit four models to rats foraging behavior in each context: a model including subjective costs, diminishing marginal returns for larger rewards, discounting of future reward, and ignoring post-reward delays, and tested whether predictions of these models were different from rats' behavior. All four models could qualitatively explain rat foraging behavior in response

to a change in travel time and patch depletion rate, but only the 'ignore post-reward delays' model, in which post reward delays are perceived to be shorter than they actually are, could predict both later patch leaving when given larger rewards with longer delays, and earlier patch leaving when a pre-reward delay was introduced. These results suggest that there are multiple cognitive biases that can explain overharvesting in certain contexts, and that foraging behavior may be the result of the use of multiple biases. However, inaccurate estimation of post-reward delays likely contributes to overharvesting.

References

- Bateson, M., & Kacelnik, a. (1996). Rate currencies and the foraging starling: The fallacy of the averages revisited. *Behavioral Ecology*, 7(3), 341–352.
- Blanchard, T. C., Pearson, J. M., & Hayden, B. Y. (2013). Postreward delays and systematic biases in measures of animal temporal discounting. *Proceedings of the National Academy of Sciences of the United States of America*, 110(38), 15491–6.
- Carter, E. C., & Redish, A. D. (2016). Rats value time differently on equivalent foraging and delay-discounting tasks. *Journal of Experimental Psychology: General*, 145(9), 1093–1101.
- Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*.
- Constantino, S. M., & Daw, N. D. (2015). Learning the opportunity cost of time in a patch-foraging task. *Cognitive, Affective, & Behavioral Neuroscience*.
- Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review*, 107(2), 289–344.
- Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2011). Neuronal basis of sequential foraging decisions in a patchy environment. *Nature Neuroscience*, 14(7), 933–939.
- Kacelnik, A. (1997). Normative and descriptive models of decision making: time discounting and risk sensitivity. In *Characterizing human psychological adaptations* (Vol. 208, pp. 51–66).
- Nonacs, P. (1991). State dependent behavior and the Marginal Value Theorem. *Behavioral Ecology*, 12(1), 71–83.
- Stephens, D. W., & Krebs, J. R. (1986). *Foraging Theory. Evolutionary Behavioral Ecology* (Vol. 121).
- Wikenheiser, A. M., Stephens, D. W., & Redish, a D. (2013). Subjective costs drive overly patient foraging strategies in rats on an intertemporal foraging task. *Proceedings of the National Academy of Sciences of the United States of America*, 110(20), 8308–13.