Rats pursue food and leisure following the same rational principles

Raegan S. Logue[1,2], Ana K. Garcia[1,2], Brissa A. Bejarano[1,2], Riley K. Kendall[1,2], Andrew M. Wikenheiser[1,2,3]

1) Department of Psychology, University of California, Los Angeles, CA
2) Brain Research Institute, University of California, Los Angeles, CA
3) Corresponding author (amwikenheiser@psych.ucla.edu)

Abstract

Animals in the wild must balance multiple, potentially mutually-exclusive goals simultaneously in order to survive. Yet laboratory tests of decision making often investigate how animals optimize their behavior to achieve a single, well-defined goal, which is often a nutritive reward. Thus, how animals solve multi-objective optimization problems is not well understood. Here, we devised an ethologically-inspired decision making task to examine how rats balance the pursuit of food and non-food reinforcement. Rats performed a free-choice patch-foraging task, in which they could earn food in one location (food patch) or interact with a rodent play structure in a different location (toy patch). The cost of switching between patches was manipulated by requiring rats to endure a long or short "travel time" penalty during which they were not able to access either patch. Rats devoted a considerable amount of their limited foraging time to patches of both types, showing a small but significant preference for food patches. In accordance with theoretical models of foraging, when the cost of switching patches was high rats chose longer stay durations in both types of patches, suggesting that similar rational principles guided their pursuit of food and non-food rewards. Examining the within-session dynamics of time allocation revealed that rats showed an early preference for spending time in toy patches that reversed over the course of the session. These results validate a naturalistic approach to testing decision making in rats over a range of food and non-food goods.

Key words: Patch foraging, decision making, multi-objective optimization, play, leisure

Introduction

Animals in the wild must manage the pursuit of many goals simultaneously. Goals differ across species, but may include finding food, protecting territory, pursuing mates, and caring for offspring—all while avoiding predators (1, 2). In times of relative abundance, and absent heavy threat of predation, some species also divide their time between work-like and leisure-like activities including rest and play (3, 4). Balancing these competing demands requires organisms to make tradeoffs, prioritizing different goals at different times depending their current biological needs and the state of the environment (5–10). For most organisms, survival therefore depends not on optimizing behavior to achieve any single goal, but rather adaptively compromising to pursue a range of ends whose costs and benefits vary over time (11, 12). Understanding how animals prioritize competing goals is important for understanding learning and decision making at the behavioral, computational, and neural levels (5, 13–20).

In contrast, laboratory tests of decision making typically ask animals to solve a single, well-defined problem in exchange for a single type of programmed reinforcer, often a nutritive food or liquid. This framework has revealed much about the neural and behavioral mechanisms of decision making, but is not well-suited to investigating how animals balance the pursuit of multiple, competing goals. Nevertheless, there is evidence that even in the context of a single focal decision-making problem animals optimize over multiple distinct objectives beyond the reinforcer programmed by the experimenter.

For instance, animals performing behavioral tasks frequently intersperse bouts of steady task performance with breaks, during which they rest, groom, or explore their surroundings. While it is tempting to view such "off-task" behaviors as undesirable interruptions to good performance, perhaps arising due to attentional or cognitive constrains, other interpretations are possible. Just as task performance provides utility in the form of programmed reinforcers, breaks in task performance may offer utility by reducing effort costs and attentional demands, as well as freeing up time for subjects to engage in other potentially rewarding behaviors. Indeed, economic and reinforcement learning frameworks have been used to model breaks in task performance as a simple form of leisure (21–25). However, the typically impoverished set of alternative reinforcers available in a standard operant box make it difficult to ascertain whether breaks in task performance are caused by a positive attraction to some alternative option or processes like boredom or fatigue induced by the focal task.

Here, we tested how rats chose to spend a limited time budget in pursuit of two qualitatively-distinct forms of reinforcement. We offered hungry rats a choice between earning food reward in one location or engaging with a playground-like toy apparatus in a different location. The reward schedule for the food option was fixed, which made the opportunity cost of engaging with the toy structure well defined and easily computed as the amount of food rats would have earned had they eschewed the toy option in favor of pursuing food. By varying the time penalty associated with switching between the food and toy options, we investigated whether rats' consumption of either or both options responded rationally to changes in the opportunity cost of time.

Results

A free-choice foraging task for comparing food and non-food reward.

Long-Evans rats (n = 10, 5 female) were tested on a three-chamber apparatus, consisting of two open-field arenas connected to one another by a corridor with computer-controlled doors (fig. 1a). Rats were placed in the corridor and the session began when both doors opened simultaneously, allowing them to choose an arena to enter. Rats were free to stay in the arena they selected for as long as they wished. Upon entering one arena, the door to the unchosen arena closed, preventing rats from transiting directly between arenas. To switch between arenas the rat entered the travel corridor, which caused the door to the previously-occupied arena to close, confining the rat to the corridor for a fixed duration to simulate travelling to a new foraging location. After each travel duration, the doors to both arenas opened simultaneously as before. Behavioral sessions continued in this way for 30 minutes. This task comprises a free-choice version of the patch-leaving problem, which has been studied extensively in behavioral ecology, psychology, and neuroscience (12, 26–34).

In each session the arenas were designated as food or toy patches. In food patches, sucrose pellets were dropped from above while rats visited the arena. The reward rate decreased over time during each visit (fig. 1b), but reset to the same initial rate when rats re-entered a food patch after a travel delay. In toy patches, a structure consisting of climbing platforms, tunnels, and fixed, manipulable objects (strings, wood blocks, etc.) was positioned in the center of the arena (fig. 1c) and rats were free to interact with it. Food was not available in toy patches. Rats experienced a randomized sequence of eight session types, which fully

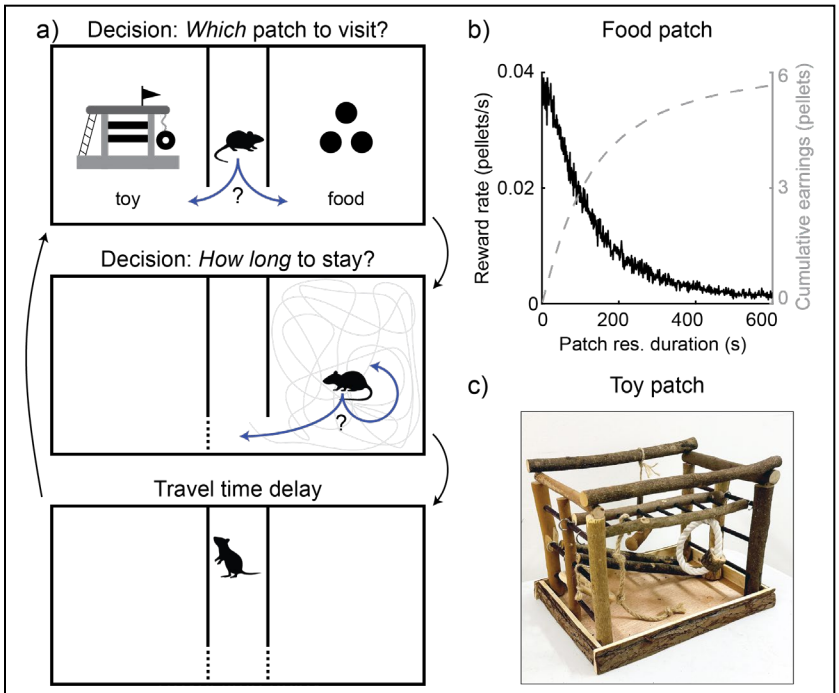

Figure 1. The free-choice foraging task. (a) Rats began each session in the travel corridor, and first selected which of two foraging patches to visit. After choosing one, they decided how long to remain in that patch. Returning to the travel corridor ended the visit to a patch, and triggered the travel time delay. After the travel time elapsed, rats were again free to select a patch to forage in. This sequence continued for the duration of behavioral sessions. Automated doors ensured that rats could not bypass the travel corridor and pass directly between the foraging patches without experiencing a travel time delay. (b) When rats visited food patches, sucrose pellets were dispensed into the arena at a rate that decreased over time. The reward rate was maximal when rats first entered the patched, decreased over the course of their visit, and reset to the maximal rate on each subsequent visit to the patch. (c) In toy patches, rats encountered a rodent play structure in the center of the patch, and were free to interact with it.

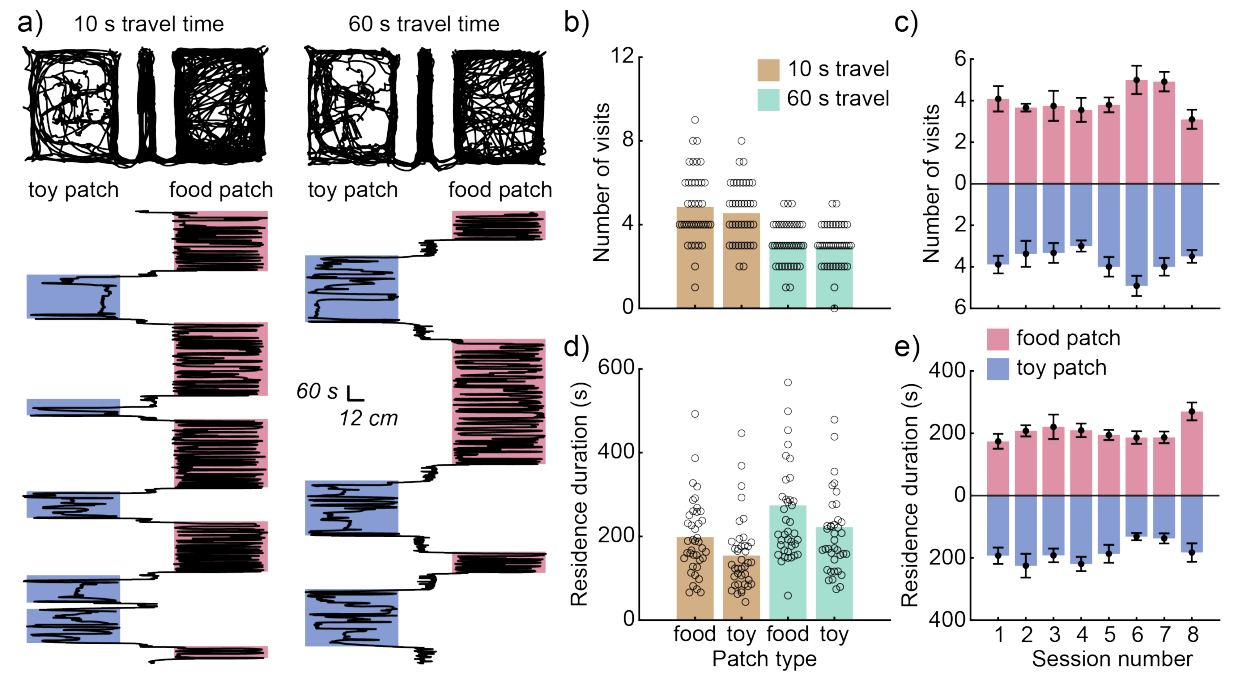

Figure 2. Behavioral results. (a) Position tracking data is shown for two behavioral sessions, one with a short (left) and one with a long (right) travel time. Plots at the top of the panel show top-down tracking data cumulated over the entirety of each session. In the bottom portion of the panel, the rat's x-coordinate is plotted as a function of time within the session, with the start of the session at the bottom of the panel. Foraging bouts to toy and food patches are outlined with colored boxes. (b) Bars show the mean number of visits per session rats made to food and toy patches, separated by the long and short travel time conditions. Each data point indicates the number of visits rats made to each type of patch in an individual session (N = 160 patches in 80 sessions from 10 rats). (c) Bars depict the mean number of visits per session rats made to food and toy patches over the course of the eight-session testing sequence. Error bars depict the standard error of the mean. (d) Bars denote the average residence duration for patches of each type, separated by the long and short travel time conditions. Data points depict the residence time for a single patch within a single session (N = 160 patches in 80 sessions from 10 rats). (e) Bars display the mean residence time in food and toy patches for each of the eight sessions that comprised the testing sequence. Error bars depict the standard error of the mean.

crossed patch type (food or toy) and arena location (left or right) at two levels of travel time (60 s or 10 s). Rats were food restricted for the duration of the experiment.

Rats visited toy and food patches frequently and spent considerable time in each (fig. 2a). We fit mixed-effects models to understand which factors affected the frequency and duration of patch visits (full model results reported in fig. S1). Rats visited toy and food patches with equal frequency (fig. 2b; visit-number model; $\beta_{\text{patch type}} = -0.04$; $p = 0.74$), but stayed ~20% longer on visits to food patches (fig. 2d; visit-duration model; $\beta_{\text{patch type}} = -0.22$; $p = 0.03$). Travel time was a significant predictor of visit number (visit-number model; $\beta_{\text{patch type}} = -0.22$; $p = 5.71\times10^{-7}$) and visit duration (visit-duration model; $\beta_{\text{travel time}} = 0.24$; $p = 3.78\times10^{-7}$), with longer travel times increasing patch residence duration, and thereby decreasing the number of visits to patches. This is consistent with models of optimal foraging, which suggest that visit duration should increase with the cost of seeking out a new patch (12, 35). The long travel time affected patch residence time similarly for toy and food patches, increasing average visit duration in food patches by 36.0% compared to short travel time sessions, and increasing the average visit duration in toy patches by 35.3% relative to

short travel time sessions. This suggests that the cost of switching foraging locations modulated decisions about food and non-food reward in a similar, rational way.

We assessed whether preferences for food and toy patches were stable over time by including session number and its interaction with patch type as a predictor in both models. Residence duration was modulated by the interaction of session number and patch type (visit-duration model; $\beta_{\text{patch type*session number}} = -0.22$; $p = 8.92\times10^{-3}$), while the number of visits to patches was not (visit-number model; $\beta_{\text{patch type*session number}} = 0.009$; $p = 0.91$). To further investigate how session number and patch type interacted to affect residence time, we computed the correlation between residence time and session number separately for toy and food patches. Across sessions, rats subtly but significantly decreased the amount of time they spent in toy patches ($R = -0.15$; $p = 0.02$; fig. 2e). This may indicate that rats' initial attraction to toy patches was partially driven by novelty, which decreased over time, or that rats learned to optimize the amount of time they spent in different patch types with experience, eventually settling into a preferred ratio of food and toy reward. The correlation between session number and residence duration in food patches was not significant ($R = 0.07$; $p = 0.23$; fig. 2e), nor was session number significantly correlated with the number of visits to food ($R = 0.06$; $p = 0.60$) or toy ($R = 0.13$; $p = 0.26$) patches (fig. 2c). Thus, despite a subtle change in time allocation to toy patches across sessions, rats' preferences were largely stationary.

Rats could not initially know that food would never be available in toy patches, so it is possible that visits to toy patches reflected exploratory food seeking rather than an attraction to the play structure. Because the play structure was consistently located in the center of patches, we used time spent in the center of the arenas as a proxy for rats' engagement with the toys (fig. 3a,b). The amount of time rats spent in the center of patches was greater in toy patches compared to food patches (fig. 3c; $\beta_{\text{patch type}} = 1.37$; $p = 6.82\times10^{-23}$). In addition, rats traversed the center region of food patches quickly, but moved much more slowly through the center of toy patches (fig. 3d; $\beta_{\text{patch type}} = -1.78$; $p = 2.68\times10^{-59}$), consistent with them pausing to engage with the play structure in toy patches, but moving swiftly through the center of food patches in pursuit of sucrose pellets.

To quantify the cost rats paid by visiting toy patches, we focused our analyses on session types with one food and one toy patch ($n = 40$ sessions). In these toy-patch vs. food-patch sessions, the opportunity cost of spending time in the toy patch was clearly defined because visits to the toy patch could have instead been devoted to the food patch, where pellets would have been delivered at a predictable rate. We computed the number of additional pellets rats would have earned on average had they allocated all their toy-patch visits in a session to the food patch (fig. 3e). The number of additional pellets rats could have earned in each session by only visiting the food patch was not significantly different than their actual earnings ($\beta_{\text{earned/foregone}} = 0.14$; $p = 0.53$). Thus, for sessions in which one patch offered food and the other offered access to the play structure, rats could have earned nearly twice as much food by avoiding the toy patch, indicating that rats' pursuit of non-food reward came at a substantial cost in foregone food.

Finally, we tested whether rats had consistent preferences for the order in which they consumed food and non-food reinforcement. For instance, one might expect a food-restricted rat to visit food

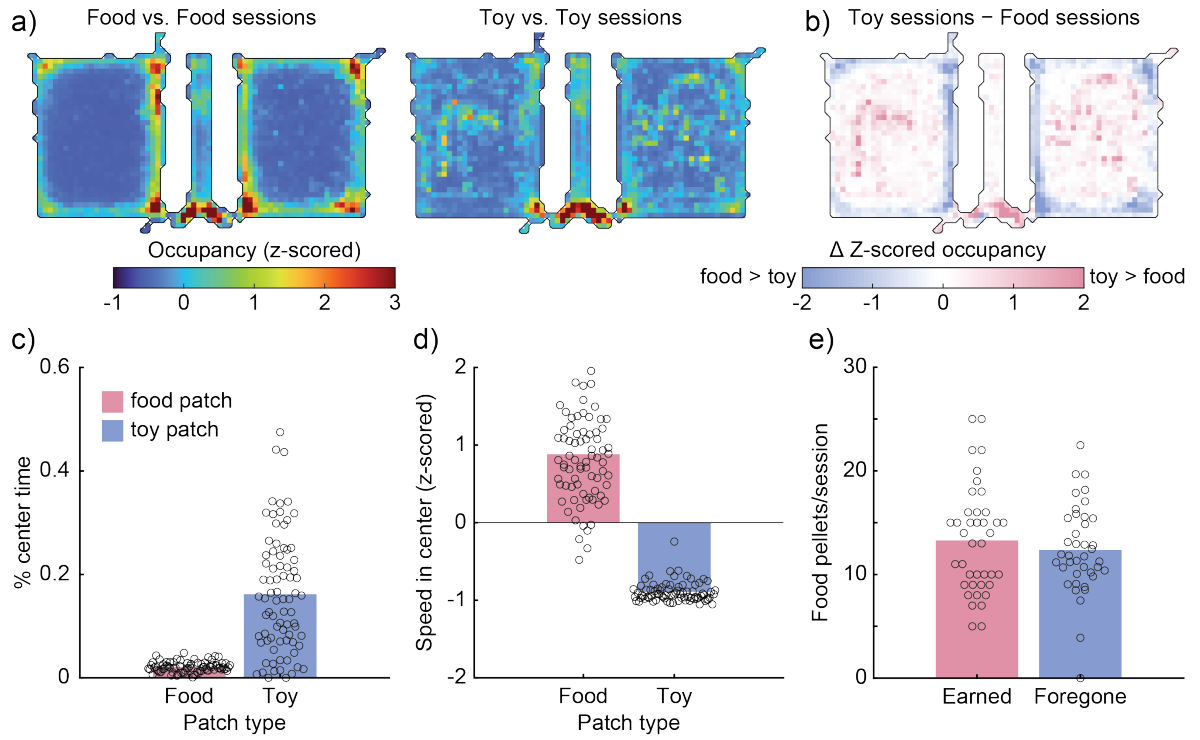

Figure 3. Rats engaged with the play structure, and visits to toy patches were costly. (a) The heat maps depict the average amount of time rats spent at each location in the apparatus in food vs. food session (left; N = 20 sessions from 10 rats) and toy vs. toy session (right; N = 20 session from 10 rats). Note the increase in time spent near the center of patches during toy vs. toy sessions, suggesting that rats actively engaged with the play structure. (b) The heat map plots the average difference in time spent at each location in the apparatus between toy vs. toy and food vs. food sessions. White colors indicate little or no difference in time spent between the two session types, while red (blue) colors denote more time spent in locations during toy vs. toy (food vs. food) sessions. Rats spent more time near the center of patches near the play structures in toy vs. toy sessions compared to food vs. food sessions. (c) We computed the fraction of time that rats spent in the center of patches during their visits to patches of either type (N = 160 patches in 80 sessions from 10 rats). Bars denote mean percentage of time spent in the center of food and toy patches, and data points show results from one patch in one session. Rats spend substantially more time in the center of patches when visiting toy patches. (d) We measured rats' mean running speed in the center of patches and split the results by patch type (N = 160 patches in 80 sessions from 10 rats). Bars depict average center running speed for food and patches, while individual data points show average center running speed in one patch for one session. Rats moved quickly through the center of food patches, but moved much more slowly through the same portion of the apparatus when the toy structure was present, consistent with rats pausing to engage with the toys. (e) For sessions in which rats choose between food and toy patches (N = 80 sessions from 10 rats) we compared the total amount of food they earned and the amount of food they could have earned on average by re-allocating their toy-patch visits to the food patch. The number of earned and foregone pellets was similar, indicating that visits to toy patches were costly in terms of lost opportunity to earn food.

patches in the early part of sessions and allocate time to toy patches later, after its hunger was at least partially sated. In fact, we observed the opposite pattern (fig. 4). Examining how rats allocated time to the food and toy patches over the course of the session revealed that rats initially devoted more time to the toy patch, and later shifted their time allocation to favor the food patch (fig. 4a). We computed the latency for rats to accumulate 100 s, 300 s, and 500 s of time spent in each patch (fig. 4b) in toy-patch vs. food-patch sessions. Rats reached 100 s ($\beta_{patch\ type} = -0.82$; $p = 1.27\times10^{-4}$) and 300 s ($\beta_{patch\ type} = -0.72$; $p = 0.001$) of residence time more quickly for toy patches than food patches. Patch type did not affect latency to reach 500 s of residence time ($\beta_{patch\ type} = -0.07$; $p = 0.76$). This is

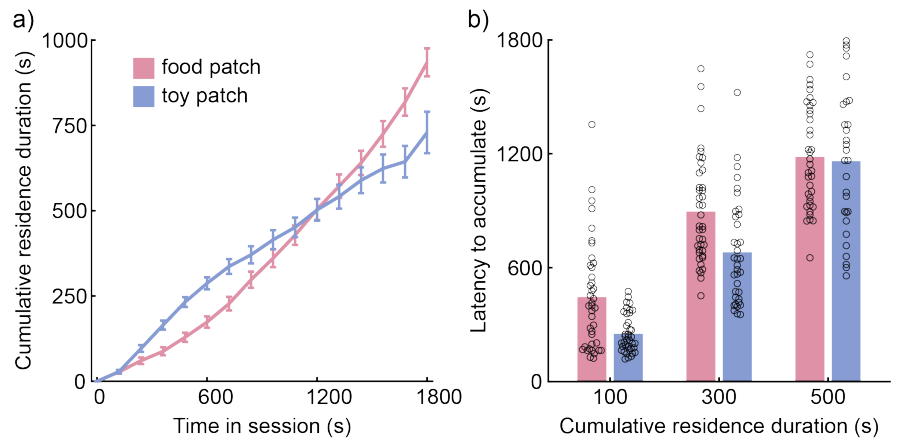

Figure 4. Temporal dynamics of food and toy patch visits. (a) For toy patch vs. food patch sessions, (N = 80 sessions from 10 rats) we tracked cumulative time spent in patches of each type across the 30-minute behavioral session. Early in sessions, rats devoted more time to toy patches, with the balance eventually shifting to favor food patches after around 2/3 of the session elapsed. (b) To quantify rats' early preference for toy patches, we computed the amount of session time required for rats to accumulate 100, 300, and 500 s in patches of each type. The latency to reach 100 and 300 s was significantly lower for toy patches, but there was no difference in latency to reach 500 s in patches of either type, suggesting that rats initially preferred to devote time to the toy patch, but shifted to allocating longer-duration visits to the food patch in the final 1/3 of the session.

consistent with time allocation to food patches eventually catching up with and overtaking time allocation to toy patches (fig. 4a).

Discussion

Using a free-choice foraging task, we found that hungry rats devoted a surprisingly large fraction of their limited time budget to seeking non-food reinforcement. Importantly, rats were sensitive to the opportunity cost of time in their pursuit of both food and non-food patches (27, 35–38); this suggests that similar principles of rational decision making hold over a range goods, and may point to a common, integrated neural substrate that supports utility-maximizing decisions across qualitatively-distinct sources of reward. It has long been known that laboratory rats are motivated to explore unvisited spaces and investigate novel objects (39, 40); our data situate these findings in a quantitative, economic framework, enabling us to calculate the cost (in terms of lost food earnings) rats were willing to pay to engage with the play structure in toy patches. Rats in our experiment could have nearly doubled their food earnings by not visiting toy patches, which suggests that the value they derived from interacting with toy objects was considerable.

There are at least two broad reasons that rats might have devoted so much time to toy patches in our experiment. First, exploration of novel stimuli, spaces, or actions may have instrumental value by leading to the discovery of new sources of reward. In reinforcement learning (41) terminology, exploratory choices ensure that agents learn an accurate value function over their surroundings (42–46). Second, interacting with toy patches may have been an end in itself, just as many behaviors that humans engage in offer subjective satisfaction without easily measured instrumental benefit (47). These two reasons are not mutually exclusive; for instance, reinforcement learning agents can be incentivized to explore by providing novelty bonuses, internally-generated reward signals that increase the value of strategically exploring undersampled features of the environment, making exploration both an end in itself and a component of an adaptive learning system. More complex directed exploration strategies similarly rely on internally-generated reward signals to direct an agent's efforts towards productive learning opportunities (48,

49). Our data are consistent with both of these explanations. Rats decreased visit durations in toy patches across successive exposures but never stopped engaging with toy patches entirely. This suggests that rats continued to derive value from engaging with toy patches even after learning they were not predictive of food.

Foraging models have increasingly been applied to decision neuroscience experiments (37, 50–52). Classic models of patch foraging maximize a single, intuitive currency: net rate of food intake (12, 27, 35, 53, 54). Such models would predict very short residence durations in toy patches since visits there never produced food. Although more complex foraging models have been developed to optimize over multiple currencies simultaneously (12), single-currency models remain in wide use and often make accurate behavioral predictions.

One well known systematic mismatch between model and data is the tendency of foragers to "overharvest" patches by staying longer than the optimal residence time (26, 28, 32, 55–58). Many explanations for overstaying have been proposed, and it seems likely that multiple factors contribute to this tendency (59–62). Our data raise the possibility that multi-objective optimization could play a role in overharvesting. Although rats in our experiment were hungry, they continued to value consumption of non-food reinforcement, and this is likely the case in experiments in which food reward is the sole programmed reinforcer. Overstaying in such settings might be driven by rats deriving satisfaction from whatever alternative behaviors are available to them within a patch, even if such opportunities are minimized by experimental design. In this vein, one recently-developed model of patch-leaving decisions considered both the amount of reward human foragers could gain by remaining in a patch and how much they would learn about the reward statistics of the environment by staying longer (30). Prioritizing both reward intake and uncertainty reduction provided a better description of decisions than models that maximized reward rate only. A challenge for future research is to better characterize the currency that foragers aim to maximize, which may be complex even in relatively constrained behavioral contexts.

Future work could further test whether rats' foraging decisions comport with additional principles of rationality. For instance, manipulating rats' access to food and toys outside of the decision-making experiment should cause predictable shifts in how rats prioritize pursuit of goods during their limited time in the foraging task. Similarly, increasing or decreasing the rate of reward in food patches offers another means of varying the opportunity cost of spending time in the toy patch; a higher rate of reward might shift time allocations to further favor the food patch, while a decrease in food delivery rate would be expected to drive additional time allocation to toy patches. Parametric variation of non-food reinforcer options could be used to further fractionate rats' preferences and provide more detailed insights into how rats derive utility from engaging with different types of objects. More generally, the free-choice foraging behavioral framework used here could be extended to investigate a wider range of both nutritive (63, 64) and non-nutritive (65–68) rewards.

Methods

Subjects. Long-Evans rats (n = 10; 5 female) were purchased from Charles River at approximately 300 g, and were maintained on a standard 12-hour light cycle. Rats were housed singly and food restricted to maintain them at 85% of their free-feeding for the duration of the experiment. Rats earned food reward in the behavioral task, but our experiments did not follow a "closed economy" design, so supplemental rodent chow was provided daily to ensure they maintained a stable weight. Rats were handled 10–15 minutes per day for a week before behavioral testing commenced. During the week of handling, rats were habituated to the behavioral apparatus by allowing them to explore it for 10 minutes per day with all doors open. Before being tested in the experiments described here, rats completed 24 sessions of a similar foraging task in the same apparatus; these sessions involved food-patch vs food-patch comparisons exclusively. Rats never experienced the toy option prior to the experimental sessions reported here. All procedures were approved by the Chancelor's Animal Research Committee at the University of California, Los Angeles.

Behavioral task. Rats were tested once per day in sessions lasting 30 minutes. In each session, rats foraged in an automated 3-chamber apparatus with 80 cm walls. Rats were placed in the travel corridor (18×71 cm) at the beginning of each session. After the travel delay elapsed, the doors blocking access to patches (61×71 cm) opened simultaneously, allowing rats to choose which location to visit. Upon entering one patch, the door to the unchosen patch closed. Sucrose pellets (45 mg; Bioserv) were delivered from dispensers (Med Associates) mounted above each patch. The travel time delay was either 10 s or 60 s, and remained fixed for the duration of the session.

The initial reward rate in food patches was 1 pellet every 26 seconds. The reward rate decreased exponentially (τ = 90 s) while the rat remained in the patch, and reset to its maximal value each time the rat re-entered a food patch following a travel time delay. The timing of pellet deliveries was stochastic, approximating an inhomogeneous Poisson process, and a unique set of pellet delivery times was programmed for each food patch visit using a sequential procedure. The latency to the first pellet delivery was drawn from an exponential distribution with λ set to 26 s (i.e., an inter-pellet interval consistent with the maximum rate of reward). Subsequent inter-pellet intervals were similarly drawn from exponential distributions with λ set to the inter-pellet interval corresponding to the patch's current reward rate. The mean inter-pellet interval after *t* seconds in the patch was computed as:

$$ipi(t) = 26 * e^{t/\tau}$$

In toy patches, rats encountered a commercial "hamster playground" (Wontee) containing beams for climbing and fixed strings suspending metal rings and small pieces of wood. The footprint of the playground was 30 cm × 20 cm, and it was 18 cm tall at its highest point. The playground was placed in the center of the patch and was too heavy for rats to move. Rats were free to interact with these objects as they chose to. Rats' location was tracked during behavioral sessions using Bonsai (69). Reward deliveries and door positions were controlled by custom Matlab (version 2024a, Mathworks) code. The play structures and the entire apparatus were cleaned thoroughly with 70% ethanol after each behavioral session.

We tested rats on an eight-session sequence that exhaustively assigned patch type (toy or food) to both locations within the apparatus (left or right) at two levels of travel time (10 s or 60 s). The patch-type assignments were therefore: toy-left/toy-right, food-left/food-right, food-left/toy-right, and food-right/toy-left. These combinations were repeated once with the long travel time and once with the short travel time. Rats experienced session types in random order.

Data analysis and statistics. All data analyses were completed using Matlab 2024a. We used linear or generalized linear mixed-effects models to quantify most aspects of behavior on the task (70). Two main models tested which factors affected the number of visits rats made to each patch within a session (generalized linear mixed effects model with Poisson link function to accommodate count data; Matlab function *fitglme*), and the amount of time rats spent in a patch on each visit (linear mixed effects model; Matlab function *fitlme*). We included the same set of predictors for both models (fig. S1). Continuous predictors (travel time and session number) were z-scored to standardize coefficient estimates and facilitate comparison of effect sizes. Categorical predictors were sex (female coded as one), patch type (toy patch coded as one), and the location of the patch within the apparatus (side; left coded as one). Rat identity was included as a random effect. The residence duration of each visit was defined as the time between when rats first crossed from the travel corridor into one of the two patches, to when rats re-entered the travel corridor far enough to trigger the doors to close and initiate the travel time delay. If the session ended while a rat was still foraging in a patch the final visit was excluded from analysis of residence durations, due to uncertainty over how much longer the rat would have remained in the patch. For mixed effects models quantifying other aspects of behavior, the predictors included are described in the main text; in all of these models, rat identity was included as a random factor.

To visualize the distribution of behavior across the patches, we divided the apparatus into a 100×50 grid of bins, and computed the amount of time rats spent in each bin during each session. Occupancy at each location was z-scored within session and then averaged across sessions, splitting by patch type to produce the heat maps in fig. 3a. The occupancy difference heat map in fig. 3b was produced by subtracting pixel-by-pixel the average heatmap of toy vs. toy sessions from the average heat map of food vs. food sessions. For analyses of time spent and running speed in the center of patches, the rat was counted as occupying the center region of a patch when their distance to the center of the patch was 17 cm or less. Choosing different values for this threshold did not affect the qualitative or statistical pattern of results.

Acknowledgements

This work was supported by Whitehall foundation research grant #2021-12-045 (AMW). RL was supported by the UCLA-HBCU Neuroscience Pathways program. The authors thank members of the Wikenheiser lab for helpful discussions about this work.

References

1. N. B. Davies, J. R. Krebs, S. A. West, *An Introduction to Behavioural Ecology* (John Wiley & Sons, 2012).

2. P. H. Rudebeck, A. Izquierdo, Foraging with the frontal cortex: A cross-species evaluation of reward-guided behavior. *Neuropsychopharmacol.* **47**, 134–146 (2022).

3. J. Berger, K. A. Cassidy, Play is a privilege in both humans and animals: how our recreation influences wildlife. *The Journal of Wildlife Management* **88**, e22664 (2024).

4. P. Bateson, Playfulness and creativity. *Current Biology* **25**, R12–R16 (2015).

5. E. Wajnberg, Multi-objective behavioural mechanisms are adopted by foraging animals to achieve several optimality goals simultaneously. *Journal of Animal Ecology* **81**, 503–511 (2012).

6. J. Felts, K. A. Schmidt, Multitasking and eavesdropping in cotton rats foraging under predation risk. *Behavioral Ecology* **21**, 1080–1086 (2010).

7. G. Beauchamp, *Animal vigilance: monitoring predators and competitors* (Academic Press, 2015).

8. J. A. van Gils, I. W. Schenk, O. Bos, T. Piersma, Incompletely Informed Shorebirds That Face a Digestive Constraint Maximize Net Energy Gain When Exploiting Patches. *The American Naturalist* **161**, 777–793 (2003).

9. K. R. L. Janmaat, R. W. Byrne, K. Zuberbühler, Primates Take Weather into Account when Searching for Fruits. *Current Biology* **16**, 1232–1237 (2006).

10. K. R. L. Janmaat, L. Polansky, S. D. Ban, C. Boesch, Wild chimpanzees plan their breakfast time, type, and location. *Proceedings of the National Academy of Sciences* **111**, 16343–16348 (2014).

11. M. Z. Wang, B. Y. Hayden, Latent learning, cognitive maps, and curiosity. *Current Opinion in Behavioral Sciences* **38**, 1–7 (2021).

12. D. Stephens, J. Krebs, *Foraging theory* (Princeton University Press, 1986).

13. C. Liu, X. Xu, D. Hu, Multiobjective Reinforcement Learning: A Comprehensive Overview. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **45**, 385–398 (2015).

14. F. S. Arana, *et al.*, Dissociable Contributions of the Human Amygdala and Orbitofrontal Cortex to Incentive Motivation and Goal Selection. *J. Neurosci.* **23**, 9632–9638 (2003).

15. E. A. Murray, E. J. Moylan, K. S. Saleem, B. M. Basile, J. Turchi, Specialized areas for value updating and goal selection in the primate orbitofrontal cortex. *eLife* **4**, e11695 (2015).

16. F. A. Mansouri, E. Koechlin, M. G. P. Rosa, M. J. Buckley, Managing competing goals — a key role for the frontopolar cortex. *Nat Rev Neurosci* **18**, 645–657 (2017).

17. B. G. C. Dellaert, *et al.*, Individuals' Decisions in the Presence of Multiple Goals. *Cust. Need. and Solut.* **5**, 51–64 (2018).

18. J. Swait, A. A. J. Marley, Probabilistic choice (models) as a result of balancing multiple goals. *Journal of Mathematical Psychology* **57**, 1–14 (2013).

19. D. J. Rapport, Optimal Foraging for Complementary Resources. *The American Naturalist* **116**, 324–346 (1980).

20. B. M. Sweis, M. J. Thomas, A. D. Redish, Beyond simple tests of value: measuring addiction as a heterogeneous disease of computation-specific valuation processes. *Learn. Mem.* **25**, 501–512 (2018).

21. R. C. Battalio, L. Green, J. H. Kagel, Income-Leisure Tradeoffs of Animal Workers. *The American Economic Review* **71**, 621–632 (1981).

22. L. Green, J. H. Kagel, R. C. Battalio, Consumption-Leisure Tradeoffs in Pigeons: Effects of Changing Marginal Wage Rates by Varying Amount of Reinforcement. *Journal of the Experimental Analysis of Behavior* **47**, 17–28 (1987).

23. R. K. Niyogi, *et al.*, Optimal indolence: a normative microscopic approach to work and leisure. *Journal of The Royal Society Interface* **11**, 20130969 (2014).

24. R. K. Niyogi, P. Shizgal, P. Dayan, Some Work and Some Play: Microscopic and Macroscopic Approaches to Labor and Leisure. *PLOS Computational Biology* **10**, e1003894 (2014).

25. J. Duncan, The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends in Cognitive Sciences* **14**, 172–179 (2010).

26. M. Garcia, S. Gupta, A. M. Wikenheiser, Sex differences in patch-leaving foraging decisions in rats. *Oxford Open Neuroscience* **2**, kvad011 (2023).

27. D. W. Stephens, Decision ecology: foraging and the ecology of animal decision making. *Cogn Affect Behav Neurosci* **8**, 475–484 (2008).

28. B. Y. Hayden, J. M. Pearson, M. L. Platt, Neuronal basis of sequential foraging decisions in a patchy environment. *Nat Neurosci* **14**, 933–939 (2011).

29. G. A. Kane, *et al.*, Increased locus coeruleus tonic activity causes disengagement from a patch-foraging task. *Cogn Affect Behav Neurosci* **17**, 1073–1083 (2017).

30. N. C. Harhen, A. M. Bornstein, Overharvesting in human patch foraging reflects rational structure learning and adaptive planning. *Proceedings of the National Academy of Sciences* **120**, e2216524120 (2023).

31. S. J. Shettleworth, "Foraging as Operant Behavior and Operant Behavior as Foraging: What Have We Learned?" in *Psychology of Learning and Motivation*, G. H. Bower, Ed. (Academic Press, 1988), pp. 1–49.

32. S. M. Constantino, N. D. Daw, Learning the opportunity cost of time in a patch-foraging task. *Cogn Affect Behav Neurosci* **15**, 837–853 (2015).

33. A. C. Kamil, S. I. Yoerg, K. C. Clements, Rules to leave by: patch departure in foraging blue jays. *Animal Behaviour* **36**, 843–853 (1988).

34. S. Hall-McMaster, P. Dayan, N. W. Schuck, Control over patch encounters changes foraging behavior. *iScience* **24**, 103005 (2021).

35. E. L. Charnov, Optimal foraging, the marginal value theorum. *Theoretical Population Biology* **9**, 129–136 (1976).

36. Y. Niv, N. D. Daw, D. Joel, P. Dayan, Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology* **191**, 507–520 (2006).

37. D. Mobbs, P. C. Trimmer, D. T. Blumstein, P. Dayan, Foraging for foundations in decision neuroscience: insights from ethology. *Nature Reviews Neuroscience* **19**, 419 (2018).

38. J. A. Beeler, D. Mourra, To Do or Not to Do: Dopamine, Affordability and the Economics of Opportunity. *Front. Integr. Neurosci.* **12** (2018).

39. D. E. Berlyne, Curiousity and Exploration. *Science* **153**, 25–35 (1966).

40. S. A. Barnett, Exploratory Behaviour. *British Journal of Psychology* **49**, 289–310 (1958).

41. R. S. Sutton, A. G. Barto, *Reinforcement Learning: An introduction* (MIT Press, 1998).

42. T. T. Hills, P. M. Todd, D. Lazer, A. D. Redish, I. D. Couzin, Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences* **19**, 46–54 (2015).

43. R. C. Wilson, E. Bonawitz, V. D. Costa, R. B. Ebitz, Balancing exploration and exploitation with information and randomization. *Current Opinion in Behavioral Sciences* **38**, 49–56 (2021).

44. S. Wang, B. Gerken, J. R. Wieland, R. C. Wilson, J.-M. Fellous, The effects of time horizon and guided choices on explore–exploit decisions in rodents. *Behavioral Neuroscience* **137**, 127–142 (2023).

45. R. C. Wilson, A. Geana, J. M. White, E. A. Ludvig, J. D. Cohen, Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General* **143**, 2074–2081 (2014).

46. B. B. Averbeck, Theory of Choice in Bandit, Information Sampling and Foraging Tasks. *PLOS Computational Biology* **11**, e1004164 (2015).

47. J. Chu, J. B. Tenenbaum, L. E. Schulz, In praise of folly: flexible goals and human cognition. *Trends in Cognitive Sciences* **28**, 628–642 (2024).

48. P.-Y. Oudeyer, J. Gottlieb, M. Lopes, “Chapter 11 - Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies” in *Progress in Brain Research*, Motivation., B. Studer, S. Knecht, Eds. (Elsevier, 2016), pp. 257–284.

49. J. Gottlieb, P.-Y. Oudeyer, M. Lopes, A. Baranes, Information-seeking, curiosity, and attention: computational and neural mechanisms. *Trends in cognitive sciences* **17**, 585–593 (2013).

50. A. J. Calhoun, B. Y. Hayden, The foraging brain. *Current Opinion in Behavioral Sciences* **5**, 24–31 (2015).

51. D. Mobbs, J. J. Kim, Neuroethological studies of fear, anxiety, and risky decision-making in rodents and humans. *Current Opinion in Behavioral Sciences* **5**, 8–15 (2015).

52. J. M. Pearson, K. K. Watson, M. L. Platt, Decision Making: The Neuroethological Turn. *Neuron* **82**, 950–965 (2014).

53. R. H. MacArthur, E. R. Pianka, On optimal use of a patchy environment. *The American Naturalist* **100**, 603–609 (1966).

54. D. W. Stephens, J. S. Brown, R. C. Ydenberg, Eds., *Foraging: Behavior and Ecology* (University of Chicago Press, 2007).

55. G. A. Kane, *et al.*, Rats exhibit similar biases in foraging and intertemporal choice tasks. *eLife* **8**, e48429 (2019).

56. P. Nonacs, State dependent behavior and the Marginal Value Theorem. *Behavioral Ecology* **12**, 71–83 (2001).

57. N. Kolling, T. E. J. Behrens, R. B. Mars, M. F. S. Rushworth, Neural Mechanisms of Foraging. *Science* **336**, 95–98 (2012).

58. T. C. Blanchard, B. Y. Hayden, Monkeys Are More Patient in a Foraging Task than in a Standard Intertemporal Choice Task. *PLOS ONE* **10**, e0117057 (2015).

59. T. Cash-Padgett, B. Hayden, Behavioural variability contributes to over-staying in patchy foraging. *Biology Letters* **16**, 20190915 (2020).

60. R. K. Kendall, A. M. Wikenheiser, Quitting while you’re ahead: Patch foraging and temporal cognition. *Behavioral Neuroscience* **136**, 467–478 (2022).

61. D. Brunner, A. Kacelnik, J. Gibbon, Optimal foraging and timing processes in the starling, Sturnus vulgaris: effect of inter-capture interval. *Animal Behaviour* **44**, 597–613 (1992).

62. M. Bateson, “Interval timing and optimal foraging” in *Functional and Neural Mechanisms of Interval Timing*, (CRC Press/Routledge/Taylor & Francis Group, 2003), pp. 113–141.

63. F.-Y. Huang, F. Grabenhorst, Nutrient-Sensitive Reinforcement Learning in Monkeys. *J. Neurosci.* **43**, 1714–1730 (2023).

64. P. Reinagel, Rational regulation of water-seeking effort in rodents. *Proceedings of the National Academy of Sciences* **118**, e2111742118 (2021).

65. J. B. Panksepp, G. P. Lahvis, Social reward among juvenile mice. *Genes, Brain and Behavior* **6**, 661–671 (2007).

66. A. Stijovic, *et al.*, Defining social reward: A systematic review of human and animal studies. *Psychological Bulletin* No Pagination Specified-No Pagination Specified (2024). https://doi.org/10.1037/bul0000455.

67. V. Trezza, P. J. J. Baarendse, L. J. M. J. Vanderschuren, The pleasures of play: pharmacological insights into social reward mechanisms. *Trends in Pharmacological Sciences* **31**, 463–469 (2010).

68. M. Venniro, T. I. Russell, M. Zhang, Y. Shaham, Operant Social Reward Decreases Incubation of Heroin Craving in Male and Female Rats. *Biological Psychiatry* **86**, 848–856 (2019).

69. G. Lopes, *et al.*, Bonsai: an event-based framework for processing and controlling data streams. *Front. Neuroinform.* **9** (2015).

70. Z. Yu, *et al.*, Beyond t test and ANOVA: applications of mixed-effects models for more rigorous statistical analysis in neuroscience research. *Neuron* **110**, 21–35 (2022).

Figure captions

Figure 1. The free-choice foraging task. (a) Rats began each session in the travel corridor, and first selected which of two foraging patches to visit. After choosing one, they decided how long to remain in that patch. Returning to the travel corridor ended the visit to a patch, and triggered the travel time delay. After the travel time elapsed, rats were again free to select a patch to forage in. This sequence continued for the duration of behavioral sessions. Automated doors ensured that rats could not bypass the travel corridor and pass directly between the foraging patches without experiencing a travel time delay. (b) When rats visited food patches, sucrose pellets were dispensed into the arena at a rate that decreased over time. The reward rate was maximal when rats first entered the patched, decreased over the course of their visit, and reset to the maximal rate on each subsequent visit to the patch. (c) In toy patches, rats encountered a rodent play structure in the center of the patch, and were free to interact with it.

Figure 2. Behavioral results. (a) Position tracking data is shown for two behavioral sessions, one with a short (left) and one with a long (right) travel time. Plots at the top of the panel show top-down tracking data cumulated over the entirety of each session. In the bottom portion of the panel, the rat's x-coordinate is plotted as a function of time within the session, with the start of the session at the bottom of the panel. Foraging bouts to toy and food patches are outlined with colored boxes. (b) Bars show the mean number of visits per session rats made to food and toy patches, separated by the long and short travel time conditions. Each data point indicates the number of visits rats made to each type of patch in an individual session (N = 160 patches in 80 sessions from 10 rats). (c) Bars depict the mean number of visits per session rats made to food and toy patches over the course of the eight-session testing sequence. Error bars depict the standard error of the mean. (d) Bars denote the average residence duration for patches of each type, separated by the long and short travel time conditions. Data points depict the residence time for a single patch within a single session (N = 160 patches in 80 sessions from 10 rats). (e) Bars display the mean residence time in food and toy patches for each of the eight sessions that comprised the testing sequence. Error bars depict the standard error of the mean.

Figure 3. Rats engaged with the play structure, and visits to toy patches were costly. (a) The heat maps depict the average amount of time rats spent at each location in the apparatus in food vs. food session (left; N = 20 sessions from 10 rats) and toy vs. toy session (right; N = 20 session from 10 rats). Note the increase in time spent near the center of patches during toy vs. toy sessions, suggesting that rats actively engaged with the play structure. (b) The heat map plots the average difference in time spent at each location in the apparatus between toy vs. toy and food vs. food sessions. White colors indicate little or no difference in time spent between the two session types, while red (blue) colors denote more time spent in locations during toy vs. toy (food vs. food) sessions. Rats spent more time near the center of patches near the play structures in toy vs. toy sessions compared to food vs. food sessions. (c) We computed the fraction of time that rats spent in the center of patches during their visits to patches of either type (N = 160 patches in 80 sessions from 10 rats). Bars denote mean percentage of time spent in the center of food and toy patches, and data points show results from one patch in one session. Rats spend substantially more time in

the center of patches when visiting toy patches. (d) We measured rats' mean running speed in the center of patches and split the results by patch type (N = 160 patches in 80 sessions from 10 rats). Bars depict average center running speed for food and patches, while individual data points show average center running speed in one patch for one session. Rats moved quickly through the center of food patches, but moved much more slowly through the same portion of the apparatus when the toy structure was present, consistent with rats pausing to engage with the toys. (e) For sessions in which rats choose between food and toy patches (N = 80 sessions from 10 rats) we compared the total amount of food they earned and the amount of food they could have earned on average by re-allocating their toy-patch visits to the food patch. The number of earned and foregone pellets was similar, indicating that visits to toy patches were costly in terms of lost opportunity to earn food.

Figure 4. Temporal dynamics of food and toy patch visits. (a) For toy patch vs. food patch sessions, (N = 80 sessions from 10 rats) we tracked cumulative time spent in patches of each type across the 30-minute behavioral session. Early in sessions, rats devoted more time to toy patches, with the balance eventually shifting to favor food patches after around 2/3 of the session elapsed. (b) To quantify rats' early preference for toy patches, we computed the amount of session time required for rats to accumulate 100, 300, and 500 s in patches of each type. The latency to reach 100 and 300 s was significantly lower for toy patches, but there was no difference in latency to reach 500 s in patches of either type, suggesting that rats initially preferred to devote time to the toy patch, but shifted to allocating longer-duration visits to the food patch in the final 1/3 of the session.

Figure S1. Regression model results. We used mixed-effects models to quantify which factors influenced the number and duration of visits rats made to patches. Coefficient estimates for each model are sorted by their magnitude, with significant predictors plotted in blue. Error bars indicated the 95% confidence interval around each coefficient estimate.

Figure S1

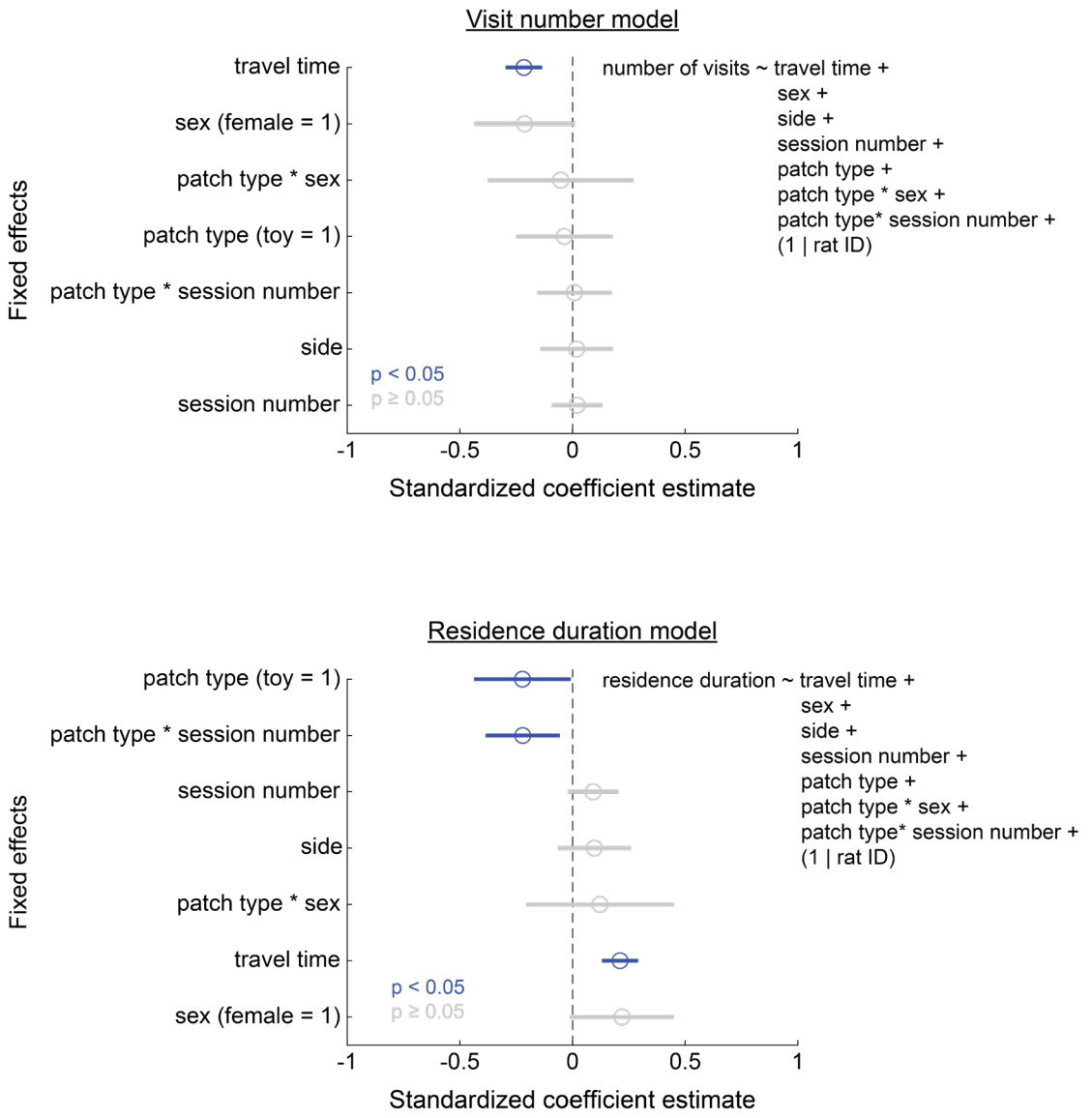

Figure S1. Regression model results. We used mixed-effects models to quantify which factors influenced the number and duration of visits rats made to patches. Coefficient estimates for each model are sorted by their magnitude, with significant predictors plotted in blue. Error bars indicated the 95% confidence interval around each coefficient estimate.