American Psychological Association

Behavioral Neuroscience

ISSN: 0735-7044

2021, Vol. 135, No. 4, 469–486
https://doi.org/10.1037/bne0000460

Certainty and Uncertainty of the Future Changes Planning and Sunk Costs

Anneke A. Duin, London Aman, Brandy Schmidt, and A. David Redish
Department of Neuroscience, University of Minnesota

Many foraging experiments have found that subjects are suboptimal in foraging tasks, waiting out delays longer than they should given the reward structure of the environment. Additionally, theories of decision-making suggest that actions arise from interactions between multiple decision-making systems and that these systems should depend on the availability of information about the future. To explore suboptimal behavior on foraging tasks and how varying the amount of future information changed behavior, we ran rats on two matching neuroeconomic foraging tasks, Known Delay (KD) and Randomized Delay (RD), with the only difference between them being the certainty of the cost of future opportunities. Rats' decision-making strategies differed significantly based on the amount of future certainty. Rats on both tasks still showed suboptimality in decision-making through a sensitivity to sunk costs; however, rats on KD showed significantly less sensitivity to sunk costs than rats on RD. Additionally, on neither task did the rats account for travel and postreward lingering times as heavily as prereward foraging times providing evidence problematic for the Marginal Value Theorem model of foraging behavior. This suggests that while future certainty reduced decision-making errors, more complex decision-making processes unaffected by future certainty were involved and likely produced these decision-making errors within subjects on these foraging tasks.

Keywords: decision-making, foraging, behavior, prefrontal cortex

Foraging is a complex neural process that depends on information processing comparing current and future (expected) options. Theoretical analyses treat foraging in terms of expected uncertainty, comparing a sure opportunity with the probability of encountering future opportunities (Charnov, 1976; Stephens & Krebs, 1986). Many foraging experiments have found that subjects are suboptimal in foraging tasks, waiting out delays longer than they should given the reward structure of the environment (Carter & Redish, 2016; Constantino & Daw, 2015; Hayden et al., 2011; Nonacs, 2001; Wikenheiser et al., 2013). Some theories have suggested that these suboptimalities arise because humans and other animals are generally risk-averse, preferring consistent choices over options with an equivalent mean but higher variance (Kacelnik & Bateson, 1996; Shafir et al., 1999) and suggest that foraging behavior should depend on the uncertainty of that future outcome.

Theories of decision-making suggest that actions arise from an interaction between different decision-making systems, which differ in their underlying algorithms (Daw et al., 2005; Dickinson, 1985; O'Keefe & Nadel, 1978; Redish, 1999, 2013; van Der Meer et al., 2012). These decision processes react to information in different ways and arrive at decisions at different speeds (Kahneman et al., 2011; O'Doherty et al., 2017; Redish, 2013; van Der Meer et al., 2012). Moreover, they depend on different neural circuits (Balleine et al., 2009; Rusu & Pennartz, 2020; van Der Meer et al., 2012) and have varying behavioral consequences (Doll et al., 2011; Everitt & Robbins, 2016), some of which can produce suboptimality in some environments. Some of those algorithms depend on planning, and some are susceptible to sunk costs, but the relationship between planning and sunk costs is currently unknown. These decision-making theories suggest that these decision systems should depend on the availability of information about future expectations.

To explore these questions, we ran rats on a pair of neuroeconomic foraging tasks varying only in the level of immediate future certainty. Previous research investigating decision-making has developed neuroeconomic foraging tasks that require subjects to make serial stay/skip decisions to earn their food for the day. These tasks are neuroeconomic because the subjects have a limited amount of time to gather food, making time an economically-limited commodity (Steiner & Redish, 2014; van Wingerden et al., 2015). Other experiments investigating decision-making in humans and other animals have used similar foraging tasks to better understand how temporally constrained decisions differ depending on the many factors that can be varied in a decision-making scenario (Carter et al., 2015; Constantino & Daw, 2015; Garrett & Daw, 2020; Stephens, 2008; Trapanese et al., 2019). Understanding how information availability alters behavior is important to further understand the underlying algorithms of decision-making systems and elucidate how decision-making can go wrong. We tested rats on a pair of tasks that provided different amounts of future information.

The Restaurant Row task is a topological loop maze with evenly spaced food reward sites. Subjects traverse the maze, encountering four "restaurant" sites serially in order to gain reward. Upon entering a "restaurant," a tone indicating how long the rat would have to wait

This article was published Online First June 24, 2021.
Anneke A. Duin https://orcid.org/0000-0002-2738-6711
London Aman https://orcid.org/0000-0002-5284-2926
Brandy Schmidt https://orcid.org/0000-0001-7755-1722
A. David Redish https://orcid.org/0000-0003-3644-9072
This research was funded by NIH MH080318 and NIH MH112688. We thank Kelsey Seeland, Chris Boldt, and Ayaka Sheehan for technical help, and the rest of the Redish lab.
Correspondence concerning this article should be addressed to A. David Redish, Department of Neuroscience, University of Minnesota. Email: redish@umn.edu

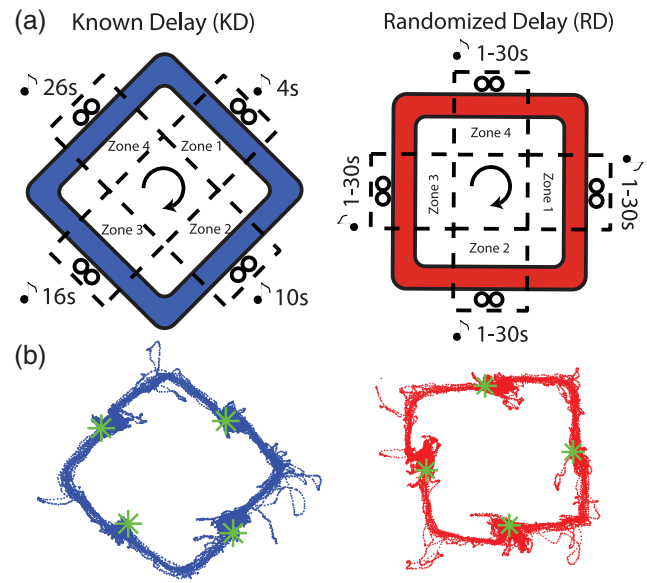
to earn reward would begin to countdown (tones indicating countdown descended in pitch by 250 Hz every second, within the audible range of the rats, (Fay, 1988), allowing the subject to make a decision either to wait out the delay and receive reward or leave and visit the next offer zone. Each day, subjects would run on the task for a set amount of time to earn their food for the day, making this task neuroeconomic; waiting for food at one restaurant was time spent out of the time budget for that day. In Restaurant Row, the subject is uncertain about the cost of future offers because the offer delays are unknown until the subject enters the zone (Schmidt et al., 2019; Steiner & Redish, 2014; Sweis et al., 2018). In this experiment, the Randomized Delay (RD) task was modeled after the Restaurant Row task (see Methods section).

The Time Out task is similar to the Restaurant Row task in that subjects traversed a topological loop maze to forage for food at equally spaced feeder locations, making decisions of whether to wait out the delay for food or to continue on to the next reward location (Wikenheiser et al., 2013). As with Restaurant Row, the rat had a set amount of time to spend foraging for food on the track. However, unlike in the Restaurant Row task, in the Time Out task, the delays at a given feeder location remained constant throughout a session (although it changed between sessions). This meant that after one loop around the track, the subjects had the ability to know what an upcoming delay would be before encountering the offer. In Time Out, the subject is certain about the cost of future offers (Wikenheiser et al., 2013). In this experiment, the Known Delay (KD) task was modeled after the Time Out task (see Methods section).

In both the RD and KD tasks, upon encountering an offer, the rat could choose to wait out the delay and earn reward (stay) or leave the wait zone and proceed to the next offer (skip). Since rats did not receive food except on the task, it was in their best interest to earn as much food as possible during each session. Previous experiments using Restaurant Row and Time Out have found that rodents exhibit sub-optimal behavior on these tasks. We set out to determine if this suboptimality was modulated by knowledge of the future.

Figure 1

The rats were Run on Two Mazes, Known Delay (KD) and Randomized Delay (RD)



Note. (a) Each maze consisted of four food reward zones. On both mazes, upon entering an offer zone, the rats heard a tone indicating the duration of the delay (delay tone decreased in 250 Hz increments, the higher the delay tone pitch, the longer the delay). The rat could choose to wait out the delay countdown and accept food reward (stay) or exit the zone and proceed to the next zone (skip). The KD task had set delays during a session but varying delays between sessions, whereas the RD task had pseudorandom delays between 1 and 30 s. During testing, each rat was run on both tasks each day (30 min per task). The order of the tasks switched between days. (b) Example tracking data of each task (green asterisks indicate reward zone location). See the online article for the color version of the figure.

Methods

Subjects

First-generation Fisher-Brown Norway hybrid (FBNF-1) rats (n = 8, four male, four female) were the subjects of this experiment, aged 7–12 months at the start of the experiment, bred in-house. Rats were housed under a 12 hr dark/12 hr light unshifted light cycle and were run during the light cycle. However, before running, rats were conditioned for three days before starting the task to expect their complete daily ration at their personal running time, which is known to shift circadian cycles (Froy, 2007). During behavioral training and testing, rats received their full daily food complement on the task. Rats were maintained above 80% of their free-feeding weight and had access to water ad-libitum. All procedures were approved by the University of Minnesota Institute for Animal Care and Use Committee (IACUC).

Behavior

Tasks

This experiment consisted of two tasks: RD and KD, see Figure 1. These tasks were modeled after Restaurant Row (Schmidt et al., 2019; Steiner & Redish, 2014; Sweis et al., 2018) and Time Out (Wikenheiser et al., 2013) and are neuroeconomic foraging tasks where rats make serial stay/skip decisions to earn food. The rats were trained to run clockwise around square mazes with four offer zones placed at the center of each side. For each rat, one maze was identified as RD and a separate maze as KD. Male and female rats ran on different physical mazes. Squares measured 77 cm on each side and consisted of 16 cm elevated tracks of width 9 cm. Upon entering a reward zone, a tone played indicating how long the animal would have to wait to receive reward, with higher frequencies corresponding to higher delays. The tone would immediately begin to count down by descending in pitch 250 Hz per second until either the rat had waited out the delay and earned reward (*stay*, earning 2 plain food pellets, 45 mg each, Research Diets, New Brunswick, NJ) or the rat left the reward zone and proceeded to the next zone (*skip*). If the rat left the reward zone, the delay tone countdown stopped, the offer was rescinded, and the next feeder in the sequence was primed for when the rat arrived at the next reward zone. On RD, delays were randomized between 1 and 30 s upon entry (uniform distribution cycling fully without replacement before restarting a new randomly ordered cycle), so the cost of an offer was unknown to the subject until entering the reward zone. On KD, each reward zone had a set delay throughout a session, but the delays

varied between sessions. The reward zones were primed in serial order which forced the rats to encounter the restaurants in the same order throughout each session. Rat position on each maze was recorded from a light-emitting diode (LED) that was strapped to a backpack worn by the rat. A Cheetah Digital Lynx SX system (Neuralynx) was used to record video time-stamped tracking data at 60 Hz. The tasks were controlled by software written in-house in MATLAB R2012a (The MathWorks, Natick, MA).

Differences Between RD, KD, and Other Tasks

The RD and KD tasks differed from the tasks they were modeled after in order to better match them to test how knowledge of the future affects decision-making. RD differed from previous iterations of the Restaurant Row task in that the food pellets used were all the same flavor, instead of four different flavors, and reward sites were on the sides of the square maze, instead of being on the end of spokes attached to the corners of the square maze (Schmidt et al., 2019; Steiner & Redish, 2014; Sweis et al., 2018). Additionally, other variations of the Restaurant Row task have included a separate offer zone and wait zone, where the delay to reward would be presented to the animal in the offer zone, however, the delay would not start descending to reward unless the rat chose to enter the wait zone. In RD only a wait zone existed. KD had differences from the previous iteration of the Time Out task; KD used a square maze with four reward zones whereas Time Out was a circular maze with three reward zones (Wikenheiser et al., 2013). These modifications allowed KD and RD to differ only in the stability of the delays on a given day: in RD, every delay was unknown until entry (random 1–30 s), while on KD, the delays could be known once the rat had made one full loop around the track.

KD Delay Sets

Five delay sets were used for testing on KD: [4 s, 10 s, 16 s, 36 s], [4 s, 12 s, 18 s, 26 s], [4 s, 14 s, 24 s, 32 s], [6 s, 12 s, 20 s, 28 s], [6 s, 12 s, 28 s, 36 s]. One additional delay set ([1 s, 4 s, 16 s, 36 s]) was run for only one session for one rat due to it being a typo from one of the other delay sets. It is included in the data set, but its inclusion did not change the results and does not affect our conclusions. KD was set up so the rat would encounter the delays from a given set in increasing order around one loop of the track.

Training Pre-Surgery

Rats received one 45-min training session each day. See Figure 2. Training began with habituation, where each feeder had a set delay of 1 s. Habituation lasted until the rat made 100 feeder entries in the time allotted, which usually took about seven days. Given the 1 s delay during habituation, rats received reward on all feeder entries. Following habituation, the rats continued training in a counterbalanced manner by either learning KD first (Training method 1, 2M, 2F) or by learning RD first (Training method 2, 2M, 2F). In Training method 1, rats received 10 days of KD with delay sets that increased from small to medium in leanness (see Environmental Leanness below). The rats then received 20 additional days of training on KD with delay sets with leanness' similar to the delay sets used for experimental testing. Note that the delay at each feeder was rotated between sessions, meaning even if the same delay set was used two consecutive days, no feeder had the same delay two sessions in a row to avoid the rat associating any one physical feeder with only low delays. The rats then proceeded to learn RD by running RD for 20 days with 1–30 s random delays at each feeder. In Training method 2, after habituation, rats proceeded with training on RD with 5 days of 1–5 s random delays at each feeder. The rats continued to run on RD for an additional five days of 1–15 s random delays at each feeder. The rats then received 20 additional days of training on RD with 1-30 s random delays at each feeder. The rats then proceeded to learn KD with 20 days of training using delay sets with leanness' similar to the delay sets used for experimental testing. We did extensive training on both the KD and RD tasks in order to ensure that lack of training was not an issue if there ended up being no difference in behavior between the different tasks. After completing either Training method 1 or 2, the animals swapped between running either KD or RD each day for 10 days (5 days KD, 5 days RD), using delay sets for KD that would be used for experimental testing and using 1–30 s random delays for RD. The animals were then put back on free food and underwent surgery. To ensure any behavioral differences observed between RD and KD was not due to the manner in which the rats were trained, we analyzed the experimental testing data from the two cohorts of rats testing on either training method and found no differences in the overall trends of the results.

Surgery and Retraining

Surgery was conducted as described in Schmidt et al. (2019). Each rat was implanted with a 64-channel dual bundle hyperdrive in CA1 of hippocampus and a 32-channel silicon probe in bilateral medial prefrontal cortex. Rats were given saline and baytril for 6 days following surgery and were given three days of recovery before retraining. Rats were then retrained swapping between KD and RD tasks each day (45 min on the task per day), until their behavior was consistent and the hyperdrive began recording an adequate number of hippocampal cells. The silicon probe and hyperdrive had their depth adjusted over the recovery period and training until the cells and signal recorded were deemed adequate to begin testing. (Neurophysiological data is still being processed and is not reported here.)

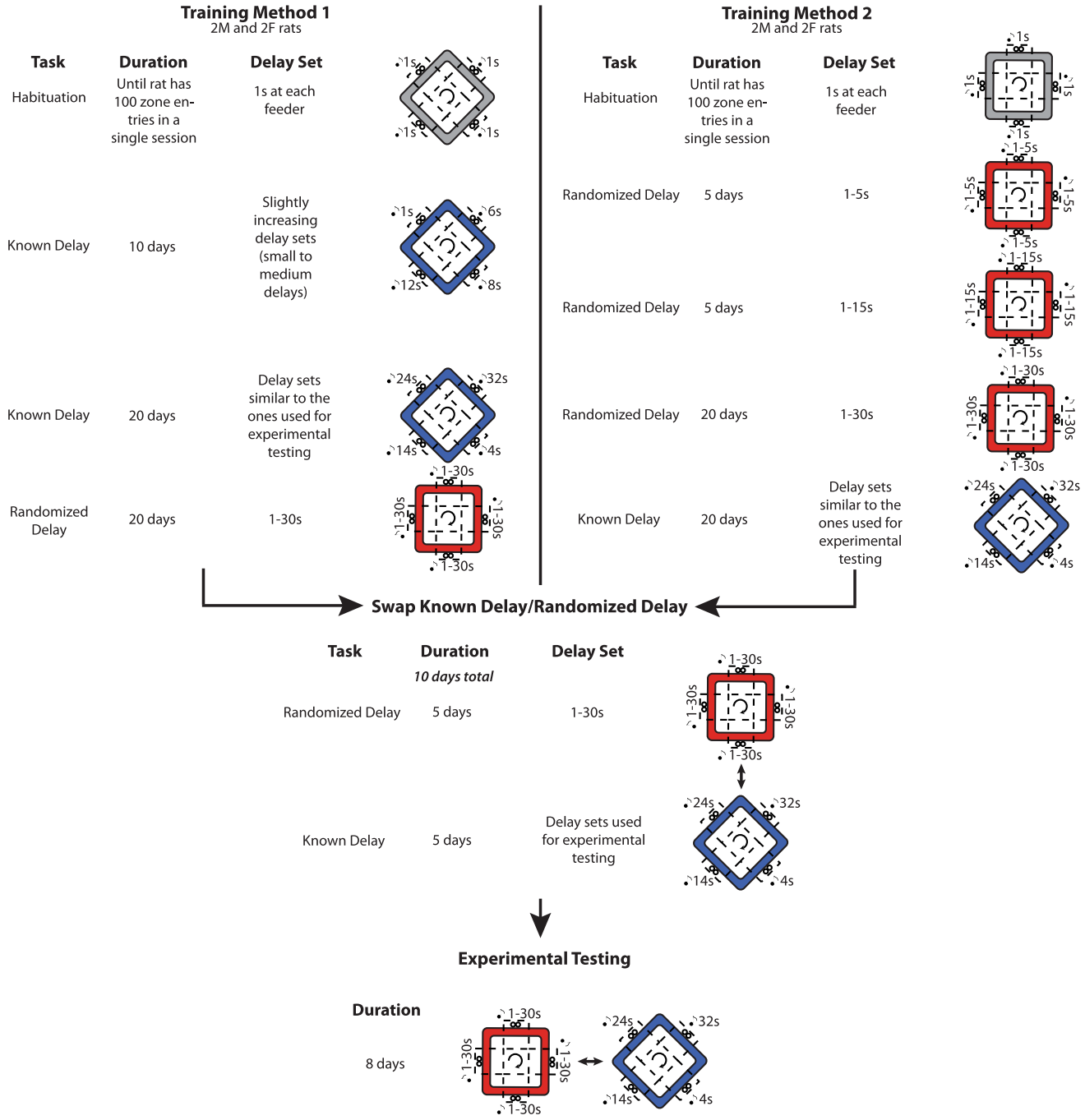
Experimental Sequence

The testing sequence lasted 8 days. Rat behavior during the experimental sequence was tracked via an LED attached to their head implant. Rats ran one session of each task for 30 min per task to gain their food for the day, making KD and RD economic tasks in which time is the commodity. The rats' neural activity was recorded while resting near the maze for 5 min preceding each task (PRE) and 5 min following (POST). Between switching tasks, the rats were given 30 min break. The order of the tasks the rats ran on during the experimental sequence was counterbalanced throughout testing. The neurophysiology of the rats during testing was recorded using Cheetah Digital Lynx SX system (Neuralynx) which also recorded video time-stamped tracking data.

DREADDS Protocol

Following the discoveries reported in this paper, we reanalyzed data from Schmidt et al. (2019), see corresponding methods for in-depth experimental protocol. Twelve Brown-Norway rats aged

Figure 2
Training Schema



Note. Rats were first habituated to the task environment and then were split into two cohorts (2M/2F on each training method) where they either learned KD then RD, or RD then KD. Training methods were counterbalanced to ensure that behavioral differences did not arise based on order of which tasks were learned. After learning both tasks, rats entered the swap phase where they swapped between running on KD or RD each day for 10 days. The rats were then experimentally tested on both tasks each day for 8 days (30 min on each task with a 30 min rest in between). See the online article for the color version of the figure.

10–14 months were used in that experiment. They were transfected with an mCitrine- ($n = 4$) or mCherry- ($n = 8$) tagged AAV8-CaMKIIα-hM4Di virus (DREADDS, University of North Carolina Vector Core, Chapel Hill, NC) bilaterally into the prelimbic cortex. The targeted coordinates for the medial prefrontal cortex viral infusion were 3.0 mm anterior/posterior, 0.7 mm medial/lateral, and

3.6 mm dorsal/ventral. All rats underwent a 20-day sequence of either clozapine *N*-oxide (CNO, 5 mg/kg sc, NIMH Chemical Synthesis and Drug Supply Program) or vehicle (VEH) injections 20 minutes before they started the task. The CNO was dissolved in dimethylsulfoxide (DMSO; Fisher Scientific, Pittsburg; PA) and 0.9% saline to yield a DMSO concentration of 10%. VEH injections also contained 10% DMSO. Experimenters were blind to the identity of the solution at the time of testing and the blind was broken after data had been collected. h4MDi DREADDs silence neuronal activity by suppressing presynaptic vesicular release and by hyperpolarizing the cell (Zhu & Roth, 2014). CNO administration activates DREADDs (Mahler et al., 2014; Stachniak et al., 2014). We interpret these manipulations as most likely impacting mPFC functionality, but CNO is known to be back-metabolized to clozapine, which can bind to serotonin and dopamine receptors and has been shown to increase behavioral flexibility in animal models (Ilg et al., 2018; MacLaren et al., 2016). There were controls done in Schmidt et al. (2019), however, the control task design did not allow for sunk cost analyses.

Analysis

All data were processed in Matlab 2017a with in-house code (The MathWorks, Natick, MA).

Total Reward

Total reward measured how many grams of food the rat earned over the course of a session.

Environmental Leanness

Leanness is defined as how much reward was available to the subject during a session. It is an indicator of how much reward one could obtain if all one did was wait out delays in the environment. Leanness can be derived as the inverse of the potential rate of reward, meaning, the more reward one can receive in one's environment per unit time, the lower the leanness of the environment. Since at each feeder site on both tasks the rats receive two food pellets if they wait out reward (reward/encounter same at each feeder), leanness can then be described as the mean of the encountered delays during a session. A large leanness indicates a scarcer environment and a small leanness indicates a richer environment. For both tasks, the environmental leanness was calculated based on experienced leanness, or the delays each rat encountered during the session.

$$\text{Environmental Leanness} = \text{mean}(\text{delays encountered during session}). \quad (1)$$

Running Speed

The running speed of the rats was calculated by first finding the change of the x and y tracking coordinates (dx/dt, dy/dt) using best-fit velocity vectors and taking the square root of the sum of the square of the dx data and square of the dy data. Derivatives were calculated using the Janabi-Sharifi algorithm (Janabi-Sharifi et al., 2000).

Sunk Cost Data Analysis

Sunk cost analyses entailed computing linear regression models of the probability of earning a reward as a function of the time remaining in the delay offer. Slopes of each of the linear regressions as a function of time waited were found in addition to the control data slopes (from when the rat enters an offer and has waited zero seconds). The differences in these two sets of data indicate whether sunk cost behavior was present. Because all lines were constrained to be 1.0 at 0 s delay (as the reward has arrived), slopes flatter than the control data indicate the rat was more likely to take a delay once they had waited varying amounts of time. The difference between the sunk cost data and control data was found, with any values above zero indicating the presence of sunk cost behavior.

Delay Thresholds

The rats made a decision at each feeder on RD and KD and each entry into a feeder zone was treated as independent. Waiting out the delay at any reward site dispensed the same amount of reward (two 45 mg plain-flavored food pellets) in each task. On RD, since every zone encounter had a delay between 1 and 30 s, we treated all feeders the same when calculating the threshold for RD because at each feeder throughout a session the rat should theoretically have the same probability of staying or skipping an offer. On KD, each feeder had a set delay throughout a session which meant the rat did not have the same probability of staying or skipping at each feeder. So, instead of calculating delay thresholds for KD, we found the probability of waiting (p_{wait}) at each separate feeder. The threshold for RD and the p_{wait} for each feeder for KD were calculated in order to determine the value of a given delay offer. Threshold for RD is a number between 1 and 30 and signifies the delay at which the rat was more likely to skip the offer than stay for the offer. In contrast, because the delays are constants, on KD, rats should express a constant probability of waiting at a given reward site, which we define as p_{wait} for that reward site (Wikenheiser et al., 2013). The p_{wait}'s are values between 0 and 1 indicating how likely a rat would be to stay on each encounter (1 being stay and 0 being skip).

Efficiency

For each task, a base optimal reinforcement rate equation and base observed reinforcement rate equation (Equations 3 and 4) was modeled after previous literature which accounted for potential earnings given the rats' defined thresholds (Stephens & Krebs, 1986; Wikenheiser et al., 2013). The observed reinforcement rate was then divided by the optimal reinforcement rate to yield the efficiency (Equation 2).

$$\text{Efficiency} = \frac{\text{observed reinforcement rate}}{\text{optimal reinforcement rate}}. \quad (2)$$

Reinforcement rate is the (real or potential) reward received divided by the time spent. Theories of reinforcement rate optimization have suggested that suboptimality of reinforcement rate arises because foraging animals only take some time components into account while ignoring others (Bateson & Kacelnik, 1995, 1996; Gallistel, 1990; Kalenscher & Pennartz, 2008; Mazur, 1985; Stephens & Krebs, 1986). In order to understand the efficiency of the rats' choices, we sought to determine how changing what time components were included in the calculation would change the derived optimal and observed reinforcement rates and the corresponding efficiency. In order to compare the optimal rates to the observed rates, the rats were assumed to have encountered the same offers in the optimal situation but to have made the choices that

would maximize rate of reward (given the integrated time factors included in the calculation). Note that reinforcement rate is thus defined as (number of rewards received)/(sum of time spent in time factors included in the calculation).

We started with a base observed reinforcement rate assuming rats attended only to the time spent waiting for reward: the amount of reward obtained on a session (based on their own behavioral thresholds) divided by all accepted delay time less than threshold. The base optimal reinforcement rate differed in that the thresholds were calculated based on what threshold should have been, given the delays encountered, in order to gain the maximum reward.

In order to better understand the rats' strategies on the tasks, we explored time components added to the denominator of the optimal and observed reinforcement rates. If the amount of time spent doing a behavior differed between tasks, the efficiency would reflect the differences in the medians of the efficiencies compared to the base efficiency value. The different behaviors for which time was accounted for included the time they took to decide to skip (decision time, DT), traveling from feeder to feeder (travel time, TT) and lingering after receiving reward (lingering time, LT). Equations 3 and 4 are the base reinforcement rate equations. For KD, optimal and observed rates differed in that the probability of waiting out a delay (p_{wait}) was determined either after the session or based on the rats' actual behavior, respectively (Wikenheiser et al., 2013). In RD, the observed rate was calculated based on the rats' behavior for each zone entry (isStay = 1, isSkip = 0), whereas the optimal rate was calculated using the threshold the rat should have used to maximize its rate of reward intake ($d <$ thresh $= 1$, $d >$ thresh $= 0$). In the base reinforcement equations below, each entry into an offer zone was treated as independent. There are four p_{wait} values for KD and only one threshold value for RD because each feeder could be treated as the same on RD but not KD. Therefore, in the equations below, p_{wait} values for each feeder are a value between 0 and 1 and isStay/isSkip is a Boolean value of either 1 (stay) or 0 (skip) making these reinforcement rate equations comparable. When calculating the optimal reinforcement rate for RD, isStay was replaced with ($d <$ thresh) and isSkip was replaced with ($d >$ thresh) indicating the behavior the rat should have done given the delays encountered. isStay and isSkip for each zone encounter was used to calculate RD's observed reinforcement rate (instead of $d <$ thresh/ $d >$ thresh) to better encapsulate the rats' behavior. Delay is represented as d in the equations below for conciseness.

$$\text{KD base reinforcement rate} = \frac{2 * \sum_{i=1}^{4} p_{\text{wait}_i}}{\sum_{i=1}^{4} d_i * p_{\text{wait}_i}}, \tag{3}$$

$$\text{RD base reinforcement rate} = \frac{2 * \sum(\text{isStay})}{\sum[d * (\text{isStay})]}. \tag{4}$$

Equations 5 and 6 adds skip decision time (DT) into the reinforcement rate calculation, accounting for the time it takes the rats to skip an offer on either task. Because there was no separate offer zone on these tasks, staying out the delay, d, includes both the decision time and the rest of the waiting time, while skipping needs to identify the time spent before leaving the zone as decision time (DT).

$$\text{KDDT reinforcement rate} = \frac{2 * \sum_{i=1}^{4} p_{\text{wait}_i}}{\sum_{i=1}^{4}[(d_i * p_{\text{wait}_i}) + (\text{DT}*(1 - p_{\text{wait}_i}))]}, \tag{5}$$

$$\text{RDDT reinforcement rate} = \frac{2 * \sum(\text{isStay})}{\sum[d * (\text{Stay}) + \text{DT}*(\text{isSkip})]}. \tag{6}$$

Equations 7 and 8 adds travel time to the denominator of the base reinforcement rate equation accounting for the travel time between feeders. Because travel times postskip and poststay could have been different, we separated them in this analysis (TT_{stay}, travel time after a stay; TT_{skip}, travel time after a skip). Lingering time after receiving reward was not included in the travel time after a stay.

$$\text{KDTT reinforcement rate} = \frac{2 * \sum_{i=1}^{4} p_{\text{wait}_i}}{\sum_{i=1}^{4}[(d_i * p_{\text{wait}_i}) + (p_{\text{wait}_i}*\text{TT}_{\text{stay}}) + ((1 - p_{\text{wait}_i})*\text{TT}_{\text{skip}})]}, \tag{7}$$

$$\text{RDTT reinforcement rate} = \frac{2 * \sum(\text{isStay})}{\sum[(d + \text{TT}_{\text{stay}}) * (\text{isStay}) + \text{TT}_{\text{skip}} * (\text{isSkip})]}. \tag{8}$$

Equations 9 and 10 take the base reinforcement rate equation and adds the time spent lingering after reward (LT) into the denominator.

$$\text{KDLTreinforcement rate} = \frac{2 * \sum_{i=1}^{4} p_{\text{wait}_i}}{\sum_{i=1}^{4}[(d_i * p_{\text{wait}_i}) + (p_{\text{wait}_i}*\text{LT})]}, \tag{9}$$

$$\text{RDLT reinforcement rate} = \frac{2 * \sum(\text{isStay})}{\sum[(d + \text{LT}) * (\text{isStay})]}. \tag{10}$$

Statistics

All statistical analyses were two-sided. Behavioral data were compared using either an ANOVA, ANCOVA, or Wilcoxon signed rank tests of significance. For the reanalysis of the Schmidt et al. (2019) data set, data were compared using Wilcoxon rank sum tests of significance.

Data and Code Availability

All data and code from this experiment are available from the corresponding author upon reasonable request.

Results

The rats were behaviorally tested on two tasks, RD and KD, seen in Figure 1. Each consisted of a slightly elevated square maze with four feeders located at the center of each square's side. While these tasks were identical in shape and structure, they differed in that RD offered pseudorandom delays from 1 to 30 s at each offer zone encounter, while KD offered a set delay at each feeder within session, but which varied between sessions. On both tasks the animals knew the locations of where the next reward could come from; however, cost of the future rewards was certain in KD but uncertain in RD.

Comparing the overall performance of the rats on each task, rats gained more reward on KD than on RD (Wilcoxon signed rank test, $p = 5.8\text{e}{-05}$) (Figure 3a), even as the environment became more lean (ANCOVA, a main effect of behavior, $F(1, 124) = 24$,

$p < 1e{-}20$, a main effect of leanness, $F(1, 124) = 11$, $p = .0014$, and no interaction between leanness and behavior, $F(1, 124) = 0.14$, $p = .71$) (Figure 3b). Along with earning more reward (Figure 3a), rats on KD also ran more laps, encountering more feeder sites (Figure 4a) (Wilcoxon signed rank $p = 3.1e{-}9$). However, the average amount of reward per entry to an offer zone was greater on RD than KD (Wilcoxon signed rank test, $p = .03$), indicating that rats were skipping more offers when on KD (Figure 4b).

Behavioral Differences in the Tasks With Varying Levels of Immediate Future Certainty

The key question for these tasks is whether the rats' behavior differed based on the knowledge they were given of future opportunities. In order to address this question, we first set out to determine if the rats showed behavior that would indicate an understanding that the KD task provided more certain future outcomes than the RD task. We first looked at the rats' behavior when within a reward zone, specifically whether the rat waited out an offer (stay) receiving food reward, or left before the countdown (skip), forgoing the reward. In both tasks, rats were more likely to stay for shorter delays than for longer delays, producing a sigmoidal curve of responses consistent with previous experiments (Schmidt et al., 2019; Steiner & Redish, 2014; Sweis et al., 2018). However, the extent of this decrease with delay varied between the tasks—rats showed a stronger effect of delay on the KD task than on the RD task (Figure 5a, ANOVA, a main effect of task, $F(1, 12581) = 73$, $p < 1e{-}20$, a main effect of delay, $F(1, 12581) = 5{,}970$, $p < 1e{-}20$, and an interaction between delay and task, $F(1, 12581) = 162$, $p < 1e{-}20$). In general, subjects are more likely to accept offers below threshold and reject offers above threshold the more certain a task is (Carandini & Churchland, 2013; Raposo et al., 2012), which was seen in the difference of behavior between the rats on KD and RD. This suggests that the rats were behaving differently on the two tasks and that the rats understood that KD was a more certain environment than RD.

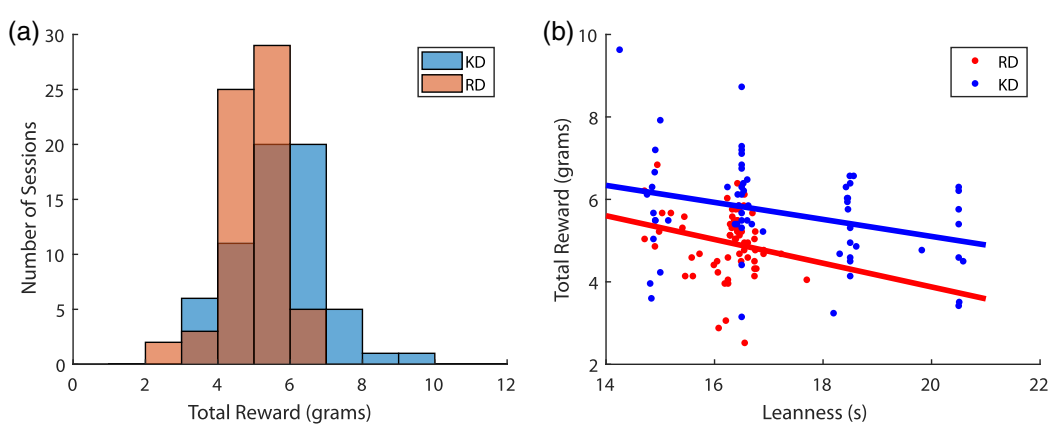
An effect of this response should be a decreased rate of staying on leaner sessions (which include more high-delay conditions). Consistent with this observation, we found that rats were more likely to skip on average over the entire task as the leanness of the task increased (ANCOVA, a trend of behavior, $F(1, 124) = 3.9$, $p = .051$, a main effect of leanness, $F(1, 124) = 18$, $p < 1e{-}20$, and no interaction between behavior and leanness, $F(1, 124) = 0$, $p = .98$) (Figure 5b). This suggests that the rats were responding to a leaner environment similarly on both tasks, by being more particular about what delays they stayed for in order to get as much reward as possible.

Efficiency Under the Normative Assumption of Maximizing Reward Rate

Both Sweis et al. (2018) and Wikenheiser et al. (2013) investigated the optimality of the behavior of animals on the analogous Restaurant Row and Time Out tasks. In order to test how task uncertainty impacted the efficiency of the rats' behavior in this study, we calculated the observed reinforcement rate and optimal reinforcement rate (see Methods section). These rates depend on which time factors are included in the denominator. Based on the presented offers and the time they were given to forage, optimal thresholds or probabilities of waiting to gain the most reward were calculated. To the temporal component of both rates, different time-consuming behaviors on the maze could be accounted for to better understand the rats' behavioral strategy to gain reward. In the base efficiency condition (Equations 3 and 4), following Wikenheiser et al. (2013), only the delays were taken into account, but we later additionally accounted separately for skip decision time (Equations 5 and 6), travel time between sites (Equations 7 and 8), and lingering time after reward (Equations 9 and 10). The efficiency was calculated as the observed reinforcement rate over the optimal reinforcement rate. A high-efficiency score (close to 1) indicates that the rat's behavior was very close to what they "should have" ideally done, meaning their behavior was well accounted for. Consistent with previous studies, we found that the base efficiency of rats on

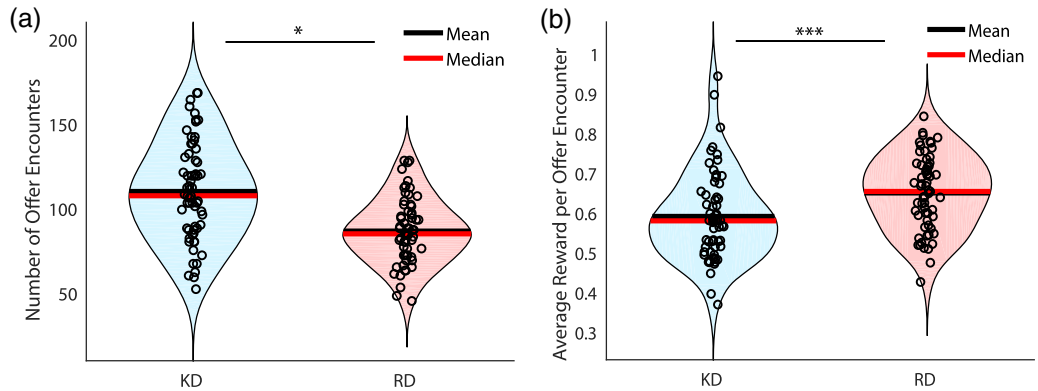
Figure 3

Rats Earned More Reward on the KD Than RD Task, Even With Changing Leanness of the Environment



Note. (a) Rats received more food on the KD task than on the RD task. (b) Rats received less reward on both tasks as leanness increased, however, across leanness rats received more reward on KD than RD. See the online article for the color version of the figure.

Figure 4
Rats Skipped More Offers on KD



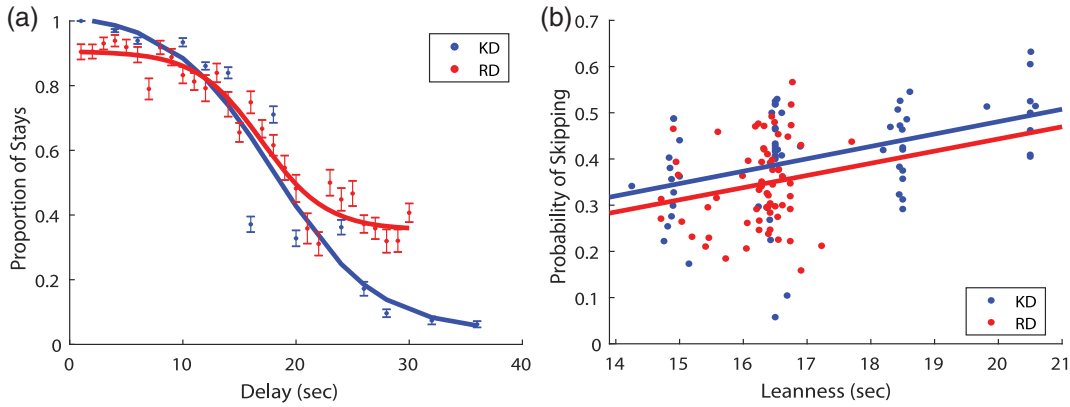
Note. (a) Rats had more offer encounters on KD than on RD. (b) The average reward per offer encounter was greater when rats ran on RD than KD , ($^{*}p < 0.05$; $^{***}p < 0.001$). See the online article for the color version of the figure.

both KD and RD were far from 100% efficient, with rats behaving more efficiently on KD than rats on RD (Figure 6a). Note that the KD base efficiencies replicated Wikenheiser et al. (2013) findings on a similar task to KD, with the efficiencies of the animals having a similar distribution between 20% and 70%.

Most theories assume that animals are normatively trying to maximize the rate of reward, but also that they do not necessarily account for all time in their temporal budget (Gallistel, 1990; Stephens & Krebs, 1986). To investigate this, we examined how changing the time components accounted for in the reinforcement rates changed the efficiency of the rats on the two tasks, simulating possible temporal budgets the rats were using to maximize their reward. Previous studies have found that subjects often ignore intertrial interval time, or the time spent after reward delivery until offer of next reward, when computing their expected rate of reward income (Bateson & Kacelnik, 1995; Gallistel, 1990; Kacelnik & Bateson, 1996; Kalenscher & Pennartz, 2008; Mazur, 1985; Stephens, 2002). This literature suggests that rats may be optimizing their behavior to the time spent upon reward offer, such as skip decision time, ignoring other behaviors such as travel time and lingering time that occur after reward outcome. For example, in previous literature looking at the efficiency behavior of rats on a similar task to KD, behavior was rate-maximized if one included an aversion-to-leave parameter, occurring upon reward offer, however, adding travel time did not drastically improve the subjects' efficiency (Wikenheiser et al., 2013, see additional unpublished note).

Factoring travel time into the reinforcement rates found that while the efficiency of the rats on KD and RD improved from the base condition, they were still not close to 100% efficient (Figure 6b). The difference between the median efficiency of KD and RD with travel time factored in decreased in comparison to the base efficiency condition due to rats on RD spending more time traveling between zones than on KD (Wilcoxon signed rank test $p = 3.1\text{e}{-05}$). This suggests that with more certainty in the task, rats run faster, and while

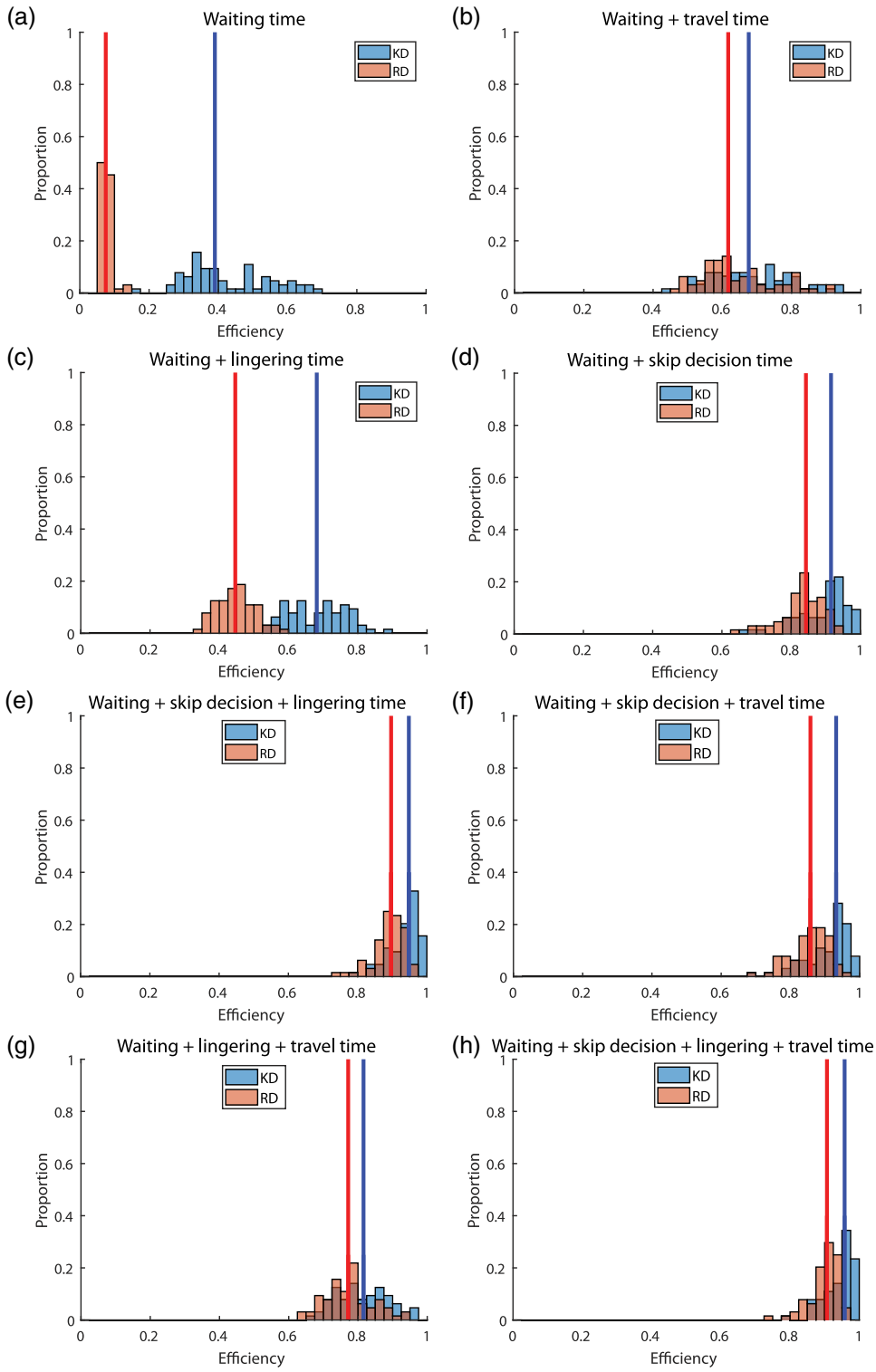
Figure 5
Rats Showed Similar Behavior Trends on Both Tasks, but the Extent of This Behavior Differed Between Tasks



Note. (a) On both tasks, the proportion of stays decreased as delay increased, however, the extent to which they decreased differed between the tasks. (b) On both tasks, the probability of skipping increased as the environmental leanness increased. Error bars represent standard error. See the online article for the color version of the figure.

Figure 6

Rats Optimized Their Behavior to Time Spent Pre-reward Outcome



Note. (a) The rats' base efficiency is higher on KD than RD, but far from 100% efficient. (b) Accounting for travel time and (c) lingering time increased the efficiency metric on both tasks, but was still far from 100% efficient. (d) Accounting for skip decision time caused the efficiency for rats on KD and RD to near 100%. Additional efficiency metrics were determined by accounting for either (e) skip decision and lingering time, (f) skip decision and travel time, (g) lingering and travel time, and (h) all three additional components into the reinforcement rates. Red and blue vertical lines indicate median. See the online article for the color version of the figure.

travel time does impact how efficiently the rats behaved, it does not explain their behavior on the maze well.

Factoring lingering time into the reinforcement rates found that while the efficiency of the rats increased on both tasks, their median efficiencies were still not close to 100% efficient (Figure 6c). The median efficiency ratio of KD and RD with lingering time factored in did not change compared to the base condition. This was due to rats on KD and RD having no significant difference in the amount of time spent lingering (KD median linger time = 8.32 s, RD median linger time 8.55 s, Wilcoxon signed rank test, $p = .32$). This suggests that time spent lingering is not correlated with future knowledge on a task and instead is likely related to reward receipt. While factoring lingering time into the efficiency analysis did improve the efficiency of the rats on both tasks compared to the base condition, their behavior was still far from efficient and not well explained.

When we instead factored in skip decision time into the efficiency analysis, we found that both rats on KD and RD were behaving close to 100% efficient (Figure 6d). This suggested that the rats' reward maximization strategy temporally accounted for the time spent making a decision before reward delivery much more heavily than the time after reward outcome such as time spent lingering or traveling. Consistent with previous studies (Bateson & Kacelnik, 1995, 1996; Kalenscher & Pennartz, 2008; Mazur, 1985; Stephens, 2002), these data suggest that rats were accounting for prereward foraging times when optimizing reward rate, and not including travel or lingering time as heavily in their behavioral optimization of reward rate. This is a particularly interesting finding given the importance of travel time in normative theories of foraging optimization (Charnov, 1976; Stephens & Krebs, 1986).

To further examine how rats were optimizing epochs of the task, we looked at the efficiency of the rats when accounting for skip decision and lingering time (Figure 6e), when accounting for skip decision and travel time (Figure 6f), when accounting for lingering and travel time, but not skip decision time (Figure 6g) and when accounting for all three additional components (Figure 6h). Accounting for lingering time increased the efficiency metric of the rats on both KD and RD more than travel time, however, neither of these components changed the efficiency metric very much relative to when only delays and skip decision time were accounted for. These data suggest that lingering time was accounted for more so than travel time in the rats' rate maximization strategies, although neither components were accounted as heavily as skip decision time.

Planning Behaviors

With future delays being known on KD but unknown on RD, we set out to determine if this difference could allow rats on KD to plan future decisions or show evidence of automation behavior. To explore this, we looked at the encounters where rats skipped and asked how long the rats took to decide to skip and proceed to the next offer. Over all sessions, the time it took a rat to decide to skip was independent of delay length on KD but increased with delay on RD (ANCOVA, a main effect of behavior, $F(1, 4951) = 103$, $p = 4.9e{-}24$, a main effect of delay, $F(1, 4951) = 106$, $p = 1.1e{-}24$, and an interaction between behavior and delay, $F(1, 4951) = 104$, $p = 3.8e{-}24$) suggesting that rats were using different decision processes on the two tasks (Figure 7).

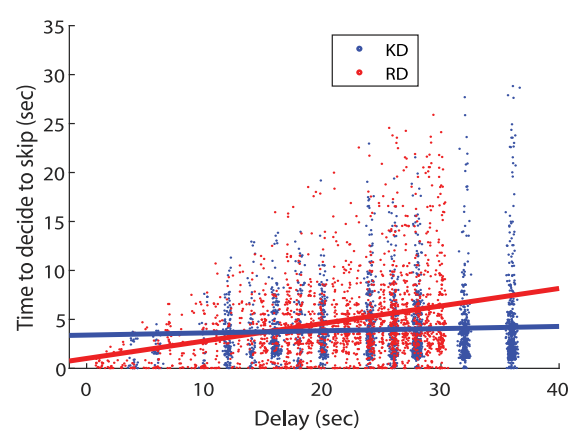
There are several hypotheses that could potentially explain the time a rat takes to decide to skip on these tasks. First, rats could have been waiting for a specific tone before skipping, which would have yielded a line parallel to the diagonal, which is not what was seen in either task (Figure 7). Second, rats could have required a constant perceptual decision time to decide to skip, which would have yielded a line parallel to the x-axis. Although a constant decision time almost parallel to the *x*-axis was seen in KD, this was not observed on the RD task (Figure 7), making this hypothesis improbable. Instead, rats were deciding to skip or not by some other, more complex processes which appears to be impacted by the certainty of the future. This behavior can be further examined by looking at the speed of the rats during different actions on the tasks. If the rats on KD are using the knowledge of the future to prepare future actions, there may be evidence in how the rats handle delays of known length on the task in comparison to when running on RD.

To better analyze the rats' behavior, we categorized patterns of stay/skip on each task as either taking the next offer (1-step), skipping one offer and taking the next (2-step), skipping two offers and taking the following (3-step) or skipping three offers and taking the following (4-step) (Figure 8). There were very few occurrences of 4-step actions, so they were omitted from further analysis. The likelihood of skipping more than one offer in a row on RD fell off as expected (each delay is random, thus, on average over all encounters, each encounter has a constant probability of being skipped, and skipping two offers in a row is a product of the two independent probabilities). However, the four delays in the KD task were always presented to the rat in increasing order. This meant that the rats encountered the two longest delays in order, which made it more likely for the rats to skip two feeders than to skip only one.

To compare speed data across the varying times it took rats to run around the maze, the time it took the rats to travel between reward zones was binned into 20 spatial bins per reward-site encounter. On journeys where the rats accepted the next offer (a 1-step journey), we found that they decreased their speed as they approached and stopped at the offer (Figure 9a–c) on both tasks.

When the rats on RD skipped one or two offers (2-step, 3-step), regardless of skipping, rats slowed upon encounter of a new offer in

Figure 7
Planning Behaviors on KD and RD

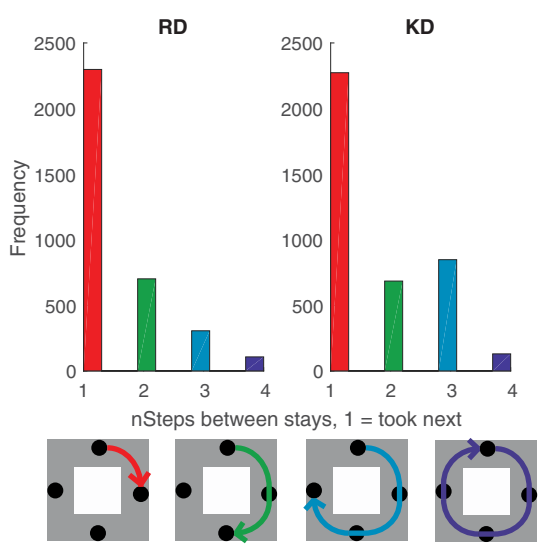


Note. As the delay increased, the rats took longer to decide to skip on RD. On KD, the time to decide to skip was independent of increasing delay. Three incorrectly tracked data points were removed. See the online article for the color version of the figure.

the same manner as if they were to accept the offer (Figure 9d, f, g, and i). Additionally, this behavior stayed constant throughout each session (Figure 9e and h). However, while rats on KD skipping one or two offers also slowed down on passing through the zones, we found that after about 20–25 encounters, rats on the KD tasks showed less or no slowing as they passed through the zones they were going to skip (Figure 9d–i). This suggests that as the session progressed, rats on KD showed a switch to automated behavior whereas on RD they did not. This further illustrates that when rats were given knowledge of the cost of future options, their behavior changed to show expectations of future outcomes.

At the start of a KD session, rats did not know the delays at each feeder until they encountered them. Thus, we might expect rats to slow down on approaching encounters early in the session, but to maintain speed on later encounters that they know they are going to skip. Because RD delays are unknown until encounter, we would expect RD behavior to remain constant throughout the session. To investigate this, we looked at the first and last 25 encounters on the KD and RD tasks to see if there was any difference in the rats' skip decision time. We found that on KD, the rats' decision to skip as a function of delay flattened from beginning to end of the session suggesting that while early behavior required deliberation, later behavior became more automated (ANCOVA, a main effect of behavior (KD beginning vs. KD end), $F(1, 1368) = 59$, $p < 1e{-}20$, a main effect of delay, $F(1, 1368) = 10$, $p = .0016$, and an interaction between behavior and delay, $F(1, 1368) = 4.2$, $p = .039$) (Figure 10). This was not seen in the RD condition suggesting that the difference in future uncertainty changed the strategies that the rats used on the tasks.

Figure 8
Rats' Behavior Was Categorized into Patterns of Stay/Skip Decisions



Note. Rats either took the next offer (1-step), skipped one offer and accepted the following (2-step), skipped two offers and accepted the following (3-step), or skipped three offers and accepted the following (4-step). On RD, the rats skipping between stays decreased exponentially. On KD, many 3-steps were observed due to the configuration of the task. See the online article for the color version of the figure.

These data suggest that changing the knowledge of future opportunities on a neuroeconomic foraging task does change decision-making behavior in rats. On KD, rats showed behavior indicative that they knew what the future offers would be before encountering them by taking less time to skip delays as the session progressed, having less change in their speed before encountering an offer when skipping, and demonstrating a different strategy of staying versus skipping various offers than on RD. Seeing these differences in decision-making behavior, we then sought to investigate whether the ability to plan affected whether rats experienced a sensitivity to sunk costs, a phenomenon known to impact decision-making.

Sunk Costs

The sunk cost fallacy is a decision-making phenomenon that is well known in humans, and has also been observed in rats, mice, and other non-human animals (Dawkins & Brockmann, 1980; Magalhães & White, 2016; Pattison et al., 2012; Sweis et al., 2018). Sunk costs refer to investments made into a course of action that cannot be recovered and according to standard economic theory, should not be included in decisions because decisions should be made based on future opportunity. Therefore, a sensitivity to sunk costs is often seen as an error in decision-making (Thaler, 1980). The sunk cost fallacy arises when one allows past investments to escalate one's commitment into a decision that may not be in one's best interest to continue pursuing (Kahneman et al., 1991; Sleesman et al., 2012; Staw, 1976; Staw & Ross, 1989). It is thought that susceptibility to sunk costs increases when the future is more uncertain (Magalhães & White, 2016), therefore, since we have seen evidence that rats on KD have knowledge of what is ahead, we investigated whether the differences of uncertainty between KD and RD would lead to a difference in sunk cost sensitivity.

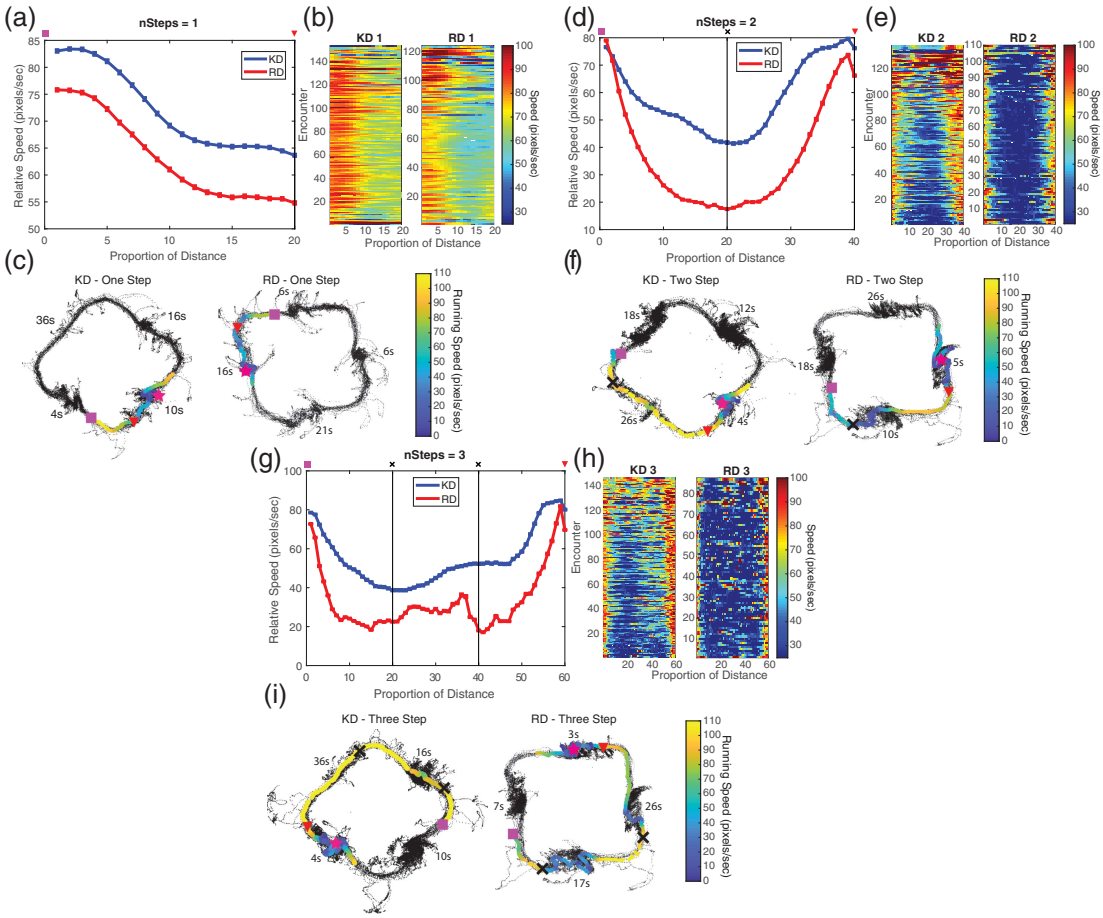
In previous research, experiments using Restaurant Row and Time Out found that subjects were not optimal in their behavior and found evidence of sunk cost sensitivity (Restaurant Row, Sweis et al., 2018) or an aversion to leave behavior (Time Out, Wikenheiser et al., 2013). We sought to investigate if a sensitivity to sunk costs could be partly responsible for the suboptimal behavior seen on RD and KD.

A sensitivity to sunk costs was seen on both KD and RD (KD: ANOVA, interaction between time spent already waiting and the difference between observed and control slopes, $F(25, 696) = 3.4$, $p = 7.0e{-}7$; even after controlling for main effects of task, $F(1, 696) = 375$, $p = 3.7e{-}67$, of time spent waiting, $F(25, 696) = 14$, $p = 2.6e{-}46$, and of the difference between observed and control slopes, $F(1, 696) = 19$, $p = 1.7e{-}5$). There was also an interaction effect between the tasks ($F(25, 696) = 4$, $p = 3.1e{-}10$) with more sunk costs seen on the RD than the KD task (see Figure 11). This suggests that adding future certainty into the task reduced sunk cost behavior but did not eliminate it completely.

Additionally, on RD, sensitivity to sunk costs appeared to be delayed for the first five seconds of waiting upon encountering the delay (Figure 11c). This was not seen on KD, which appeared to show sunk costs starting immediately, reminiscent of effects seen in Sweis et al. (2018). Rats on RD did not have knowledge of the future, which meant that they did not know the delay to reward until entering a reward zone. This meant that they would need time upon entering the reward zone to perceive what the delay was. Previous iterations of Restaurant Row included an offer zone and a wait zone (Sweis et al., 2018). In those versions, upon entering the offer zone, a tone indicating the length

Figure 9

Rats on KD Showed Knowledge of Future Expectations



Note. (a) Average binned speed on the 1-step condition. Rats decreased their speed upon encountering a delay which they would accept on both tasks. Error bars represent standard error. (b) Average binned speed over encounters for RD and KD. (c) Example tracking with speed plotted for a 1-step condition for both RD and KD. (d) The average binned speed on the 2-step condition. Rats on KD showed less slowing before encountering a delay they would skip. (e) This behavior developed around the 15–20 encounter mark on KD. On RD, rats slowed before encountering an offer they would skip. (f) Example tracking with speed plotted for a 2-step trial. Four extraneous tracking points were removed from KD 2-step. (g) The average binned speed on the 3-step condition showed rats on RD slowing before skipping each offer. Rats on KD showed less or no slowing before encountering an offer they would skip. (h) This behavior developed around the 20–25 encounter mark. (i) Example tracking with speed plotted for a 3-step trial. A few extraneous tracking points were removed from KD 3-step. Purple squares indicate where the rat exited the first zone in the step conditions, black X's indicate zone entrances during step conditions, and the red triangle indicates last zone entry of step conditions. Pink star indicates a feeder fire on the example tracking figures. See the online article for the color version of the figure.

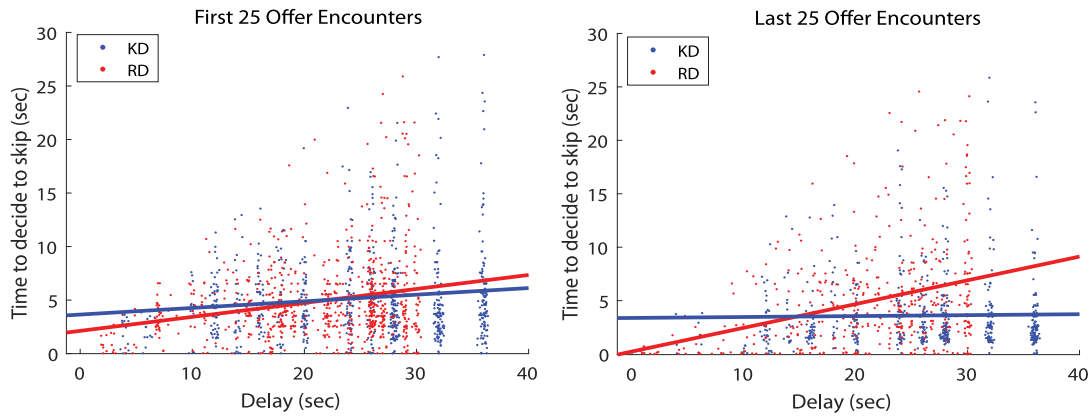
of the delay to reward would be emitted but would not start counting down. At that point, the subject could choose to either enter the wait zone, where the delay tone would begin descending to reward, or the subject could skip and proceed to the next offer. This differs from Randomized Delay where there only exists a wait zone; upon entering the wait zone a delay begins counting down to reward. Previous research has found that on the Restaurant Row with an offer zone, sunk costs do not accumulate in the offer zone but immediately start accumulating if they enter the wait zone. This research has also shown that on Restaurant Row with only a wait zone, sunk cost accumulation was delayed (Sweis et al., 2018, see Supp. Figure 7). Sweis et al. (2018) have suggested that decision processes used to perceive and react to the delay are not accounted for by the same neural system that adds time investments into the rats' internal sunk cost calculation, which would predict that the rats' escalation of commitment would be delayed when they could not plan, leading to differences between KD and RD.

Sunk Costs Under mPFC Disruption

Seeing that sunk cost behavior was different on tasks where rats could or could not plan, this drove us to revisit an old data set published in Schmidt et al. (2019) to look at sunk costs when the rats' medial prefrontal cortex (mPFC) was disrupted using h4MDi

Figure 10

Rats Decision to Skip on KD Significantly Flattened Throughout Each Session



Note. The time rats took to skip was plotted versus delay for the first 25 and last 25 encounters of each session. One data point was incorrectly tracked and was removed from the first 25 offer encounters figure. See the online article for the color version of the figure.

DREADDS because the mPFC is well-established as having a role in planning behavior (Euston et al., 2012). In that study, rats were virally transfected with h4MDi DREADDS with a CAMKIIα promoter and were run on a variation of the Restaurant Row task where clozapine N-oxide (CNO, disrupted) was used to activate the DREADDS; vehicle control injections were used as a control (VEH, intact, see Methods). The Restaurant Row task differed from the RD task in that the reward zones were located at the end of four spokes and different flavored pellets were used (cherry, chocolate, banana, and plain) (Figure 12a). Comparing the disrupted to the intact condition, rats under the disrupted condition ran more laps (Figure 12b, Wilcoxon rank sum test $p = .0014$) and gained more reward (Figure 12c, Wilcoxon rank sum test $p = .00029$).

On the Restaurant Row tasks, planning can be measured through vicarious trial and error behavior (VTE; Steiner & Redish, 2014; Sweis et al., 2018). VTE is a behavioral indication of deliberative decision-making and is calculated as the z-scored integrated angular velocity measure of the rats' head at a decision point during a task (Muenzinger, 1938; Muenzinger & Gentry, 1931; Papale et al., 2012; Redish, 2016; Schmidt et al., 2013, 2019; Tolman, 1939). The rats under the disrupted condition showed less VTE behavior suggesting that DREADDS caused a disruption of deliberative behavior with prefrontal disruption [(probability of VTE, pVTE; Figure 12d), Wilcoxon rank sum test $p = 3e{-}6$].

Since behavioral changes were seen with prefrontal disruption, we were curious to see whether rats under the disruption condition would show an increased sensitivity to sunk costs. While rats in both the disrupted and intact conditions showed a sensitivity to sunk costs (Figure 12e), rats in the disrupted condition showed an increased sensitivity to sunk costs (Figure 12f, ANOVA, interaction between task and observed versus control slopes, $F(1, 791) = 26$, $p < 1e{-}20$, even after controlling for a main effect of task, $F(1, 791) = 4.2$, $p = .04$, main effect of time spent, $F(23, 791) = 96$, $p < 1e{-}20$, main effect of observed versus control slopes, $F(1, 791) = 554$, $p < 1e{-}20$, and interaction between time spent and observed versus control slopes, $F(23, 791) = 20$, $p < 1e{-}20$). As noted in the methods, these data cannot preclude the possibility that CNO may have affected the sunk costs directly rather than through the medium of disrupting mPFC. Nevertheless, our results do find a difference between CNO and VEH in these virus-transfected rats that supports an inverse relationship between VTE and sunk costs.

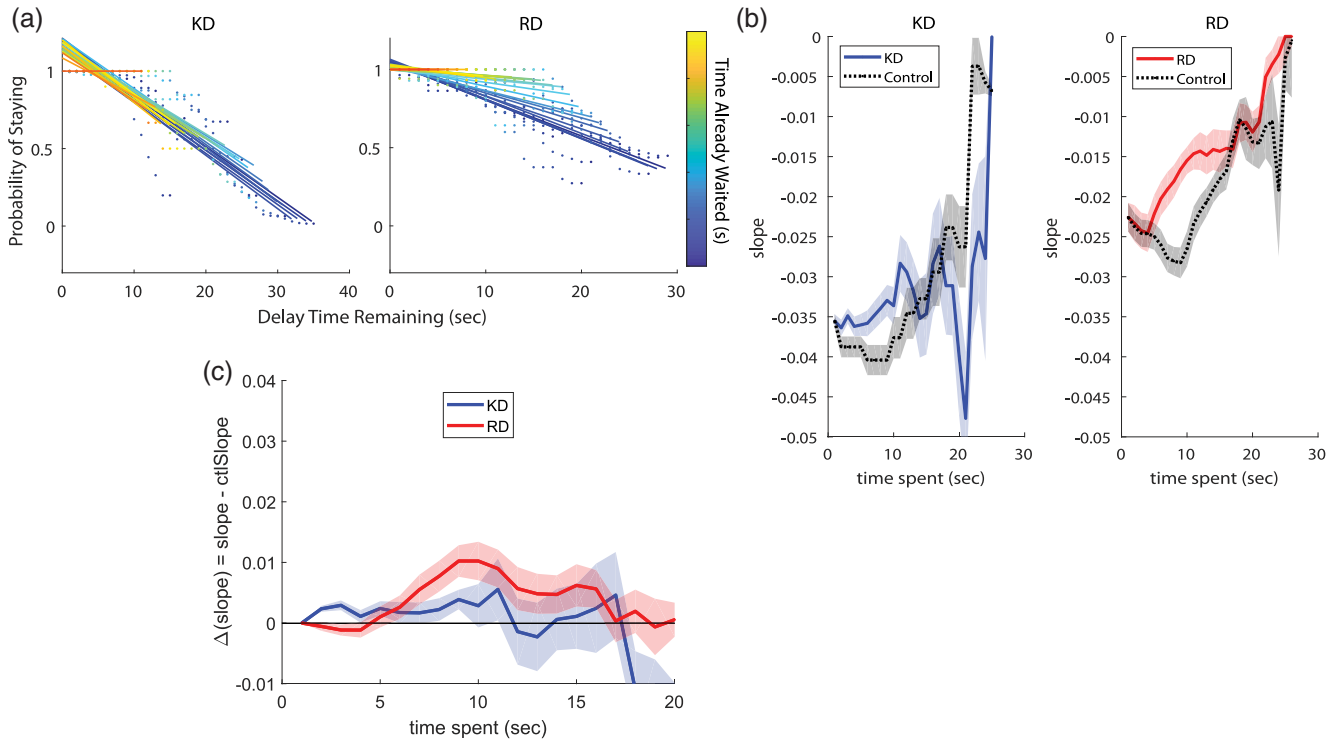
Previous literature has found that VTE behavior allows subjects to evaluate whether an offer is worth pursuing (Sweis et al., 2018). These studies found that this time spent deliberating was not included in the accumulation of sunk costs (Sweis et al., 2018). This suggests that because rats in the disrupted condition showed less VTE behavior, instead of evaluating a decision before entering the spoke, they likely made a snap judgement to enter the spoke and subsequently had to reevaluate the decision once committed. By the time the mPFC-disrupted rats were reevaluating their decision to stay, they had already made the decision to enter the spoke whereas rats in the control condition used deliberative strategies to decide whether to pursue or avoid the problematic offer. This could explain the increased sensitivity to sunk costs in the disrupted condition; by disrupting VTE, the mPFC-disrupted rats were left only with reevaluation of their decisions, causing them to accumulate more sunk costs than rats with VTE intact. These data suggest an inverse relationship between the presence of VTE behavior and sunk cost sensitivity—when rats were unable to use VTE behavior to evaluate whether to pursue or avoid a problematic decision, they had to make the decision to quit as a reevaluation and became more susceptible to sunk costs.

Discussion

We sought to investigate how decision-making behavior would change when rats were given varying knowledge of the future in a neuroeconomic foraging environment. We found that when rats were given knowledge of future conditions, they were able to prepare for future actions. On KD, the rats showed behavior indicating that they understood future offers before encountering them by taking less time to skip bad offers as the session progressed, maintaining speed when encountering an offer they would skip, and having a more exploitative stay/skip response curve on KD than RD. We then investigated how having the ability to plan would affect errors in decision-making such as showing a sensitivity to sunk

Figure 11

Sunk Cost Behavior Immediately Accrued on KD But Not on RD



Note. (a) An example from one rat of the probability of staying for a reward in the offer zone on KD and RD (R397). Dark blue data points indicate trials in which the subject had just entered the offer zone (control data). Colored data points indicate delay time remaining after the subject had already waited varying amounts of time. (b) Slope of points from probability of staying as a function of time remaining plotted against the amount of time waited. Colored lines represent the sunk cost trials, the black dashed line represents the control data. Mean over all rats, shaded errorbars indicate standard error over rats. (c) Difference in sunk cost slope to control slope in (b) versus time spent waiting for a delay. Values above zero indicate sunk cost behavior was present. Rats showed sunk cost behavior on both KD and RD, however rats on RD showed more sunk costs than on KD. Additionally, the accrual of sunk costs was delayed on RD but not on KD. See the online article for the color version of the figure.

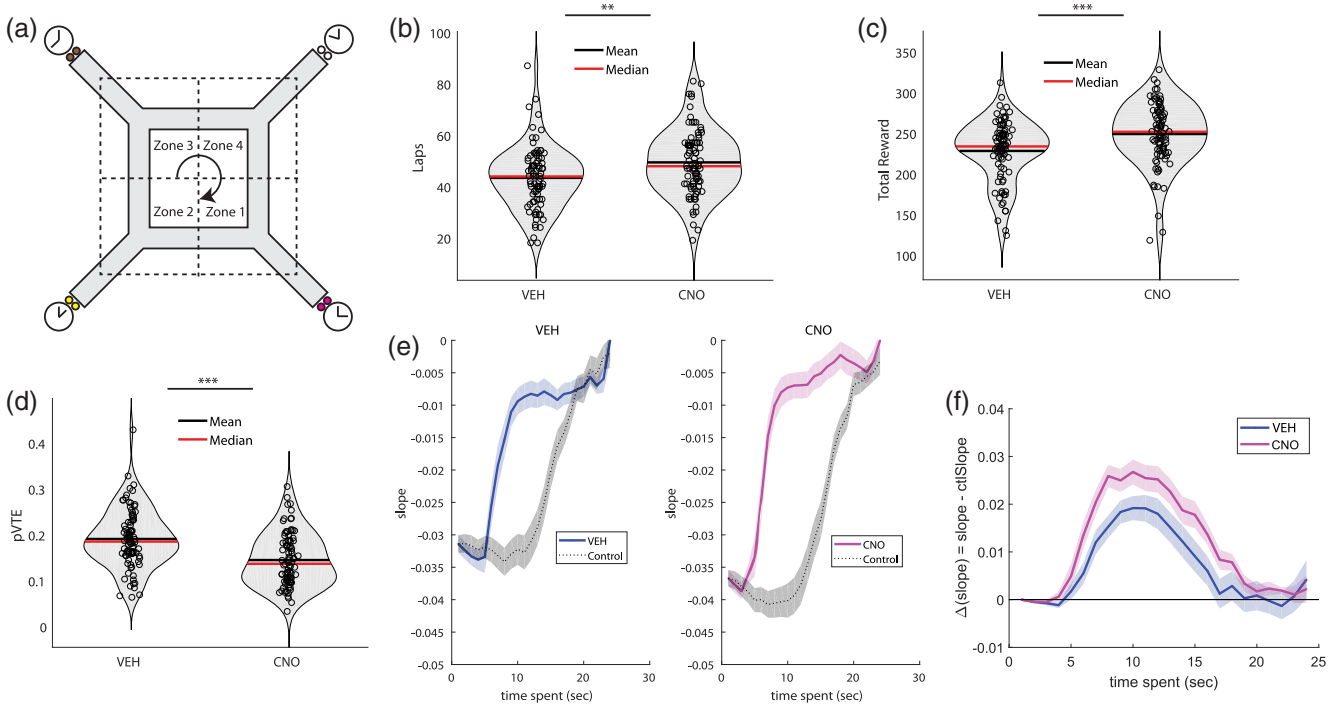
costs. We found that while rats on both tasks showed a sensitivity to sunk costs, rats on KD showed significantly less than on RD, suggesting that while future knowledge reduced the escalation of commitment from past investments of time, it did not completely eliminate it. We also found from the Schmidt et al. (2019) reanalysis that when the rats' planning was disrupted, indicated by a reduction in VTE behavior, sunk cost sensitivity increased. These data suggest that rats use a complex decision-making process where multiple factors such as past investments and future potential earnings had varying levels of impact on the neural systems contributing to decision-making, depending on the decision-making systems activated to solve a given task.

This study aimed to extend two previous studies, Wikenheiser et al. (2013) and Sweis et al. (2018), where they found suboptimality in the behavior of animals on tasks similar to KD and RD. On the Time Out task, Wikenheiser et al. (2013) was able to explain the observed inefficiency by adding an aversion to leave a presented offer as a parameter to the reinforcement rate equation. On the Restaurant Row task, Sweis et al. (2018) observed a sensitivity to sunk costs and suggested the suboptimality was due to multiple, parallel valuation algorithms that are differently susceptible to sunk costs. Our data suggest that the aversion to leave hypothesis proposed by Wikenheiser et al. (2013) may be the same behavioral process as the sunk cost fallacy, with the differences in sensitivity to sunk costs on the tasks explained by the knowledge of the future the subject has.

Many rational choice theories assume that decision-makers aim to rate-maximize (Stephens & Krebs, 1986). As such, rate maximization equations predict that animals' preferences and decisions will reflect choosing rewards that will increase the ratio of reward per unit time, where time represents all time in an environment where rewards are able to be earned. However, in many situations it has been observed that subjects' decisions are determined not by this ratio but only by the waiting time preceding rewards (Gallistel, 1990; Kacelnik & Bateson, 1996; Kalenscher & Pennartz, 2008; Stephens & Anderson, 2001). This was replicated in this study, where the efficiency of the subjects was not well explained by including travel time or lingering time in the rate-maximizing equations. So, if the rats were making their decisions and maximizing their behavior primarily on the time before reward delivery, this time may have more weight in the rats' decision-making processes than other times on the maze. If that is the case, then when rats on either task enter a reward zone for a delay above threshold and pause for some amount of time, that time may be processed differently than

Figure 12

hM4Di DREADDs Rats Showed Behavioral Changes Under CNO Including an Increased Sensitivity to Sunk Costs



Note. (a) Schematic of the Restaurant Row task. Different color pellets indicate different flavors (brown = chocolate, white = plain, yellow = banana, pink = cherry). (b) Number of laps ran during the CNO and VEH conditions. (c) Reinforcement rate for the CNO and vehicle (VEH) conditions, measured in 45 mg pellets received during the one hour session. (d) Probability of VTE during the CNO and VEH conditions. (e) and (f) Sunk cost sensitivity for CNO and VEH conditions. Shaded error bars represent standard error. (** $p < 0.01$; *** $p < 0.001$). See the online article for the color version of the figure.

time elsewhere on the maze, because that time means more to the rats.

Additionally, our results are problematic for the Marginal Value Theorem (MVT). MVT is an optimality model that is used to describe rate-maximizing behavior in patch environments much like the foraging tasks used in this experiment (Charnov, 1976). MVT assumes that when an animal is harvesting resources at a patch, the resources available will decrease over time, meaning the value of staying at that patch decreases over time. The MVT assumes the subject compares the value of staying at the current patch with the rate of reward in the environment (including the time it takes to travel from the current patch to search for a new patch) when deciding whether to remain at the current patch or search for a new one (Charnov, 1976). However, MVT is problematic in our analysis because MVT assumes that the subject is assessing the rate of reward correctly, by accounting for time-consuming factors like travel time into their foraging strategy. We found that the rats in this experiment were not interpreting all time as the same when making their decisions, accounting for prereward foraging times more heavily than postreward lingering or traveling times. Our data suggest that the MVT may need to take context and experimental conditions into account in order to accurately predict behavior on a foraging task.

Why did future knowledge not completely abolish sunk cost sensitivity in rats on KD? Sunk cost sensitivity arises when the subject's past investments, in this case time, influences their current decisions. Theory suggests making decisions requires choosing something that will maximize your future earnings, not choosing something based on investments already lost (Kalenscher & Pennartz, 2008; Stephens & Krebs, 1986). When it comes to making a decision, many authors have argued that a reevaluation of one's current action followed by a change in course of action can actually be rational behavior especially in situations where past investments exist (Arkes & Ayton, 1999; Karlsson et al., 2005; Strotz, 1955). One example of this is when a rat enters a reward zone with a delay above threshold. The offer is not in their best interest to take, so choice theory suggests that it is in the subjects' best interest to reevaluate and change its course of action regardless of time already waited. Many argue that a susceptibility to sunk costs arises in scenarios where the outcome value has significant uncertainty (Arkes & Blumer, 1985), so one could argue that sunk cost sensitivity is seen in rats on RD because, although the location of their next reward is certain, if they choose to skip a bad deal they could encounter an even worse deal at the next reward zone, making outcome value uncertain. However, this does not explain why rats on KD also showed sunk cost sensitivity.

On KD, rats had complete knowledge of their potential future earnings which theoretically would indicate that because there is no uncertainty of future outcome value there would be no need to let past investments influence their decisions. However, this was not

seen. Instead, a sensitivity to sunk costs was also observed in rats on KD.

Some theories suggest that the reason one can get trapped in escalating their commitment is uncertainty of the future (Magalhães & White, 2016); other theories have suggested that one can get trapped in escalation due to overattention to immediately present situations (Kahneman et al., 1991; Staw & Ross, 1989). Our data suggest both components are involved in sunk-cost behavior. Uncertainty played more of a role in RD than KD, however, overattention to immediately present situations would appear in both tasks and could explain why some sunk costs remained in KD. The multiple systems theory suggests that behavior arises from interacting decision-making systems. Our data suggest that knowledge of the future reduces but does not eliminate sunk costs. The evidence that VTE, known to be an indicator of deliberative decision-making processes (Redish, 2016), is inversely related to sunk costs indicates that some aspects of sunk costs arise from competing decision-making systems, most likely Pavlovian systems attending to immediate situations, as seen in conditioned place preference (Huston et al., 2013) and the endowment effect (Hutcherson et al., 2015; Kahneman et al., 1991).

Providing subjects with knowledge of the future could reduce uncertainty and change how the systems interact to reduce behavioral errors like sunk costs; however, even under future certainties, one should not expect to fully remove sensitivities to sunk costs. We saw in this experiment that rats on KD knew of the future delays to reward, but still showed a sensitivity to sunk costs. These data suggest that subtle changes in task structure can have dramatic effects on how decision-making systems interact to control behavior. How you ask the question changes the subject's response.

References

Arkes, H. R., & Ayton, P. (1999). The sunk cost and concorde effects: Are humans less rational than lower animals? *Psychological Bulletin*, *125*(5), 591–600. https://doi.org/10.1037/0033-2909.125.5.591

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, *35*(1), 124–140. https://doi.org/10.1016/0749-5978(85)90049-4

Balleine, B. W., Daw, N. D., & O'Doherty, J. P. (2009). Multiple forms of value learning and the function of dopamine. In P. W. Glimcher, C. F. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Decision making and the brain* (p. 367–387). Elsevier. https://doi.org/10.1016/B978-0-12-374176-9.00024-5

Bateson, M., & Kacelnik, A. (1995). Preferences for fixed and variable food sources: Variability in amount and delay. *Journal of the Experimental Analysis of Behavior*, *63*(3), 313–329. https://doi.org/10.1901/jeab.1995.63-313

Bateson, M., & Kacelnik, A. (1996). Rate currencies and the foraging starling: The fallacy of the average revisited. *Behavioral Ecology*, *7*(3), 341–352. https://doi.org/10.1093/beheco/7.3.341

Carandini, M., & Churchland, A. K. (2013). Probing perceptual decisions in rodents. *Nature Neuroscience*, *16*(7), 824–831. https://doi.org/10.1038/nn.3410

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. https://doi.org/10.1037/xge0000196

Carter, E. C., Pedersen, E. J., & McCullough, M. E. (2015, February). Reassessing intertemporal choice: Human decision-making is more optimal in a foraging task than in a self-control task. *Frontiers in Psychology*, *6*, Article 5. https://doi.org/10.3389/fpsyg.2015.00095

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, *9*(2), 129–136. https://doi.org/10.1016/0040-5809(76)90040-X

Constantino, S. M., & Daw, N. D. (2015). Learning the opportunity cost of time in a patch-foraging task. *Cognitive, Affective & Behavioral Neuroscience*, *15*(4), 837–853. https://doi.org/10.3758/s13415-015-0350-y

Dawkins, R., & Brockmann, H. J. (1980). Do digger wasps commit the concorde fallacy? *Animal Behaviour*, *28*(3), 892–896. https://doi.org/10.1016/S0003-3472(80)80149-7

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–1711. https://doi.org/10.1038/nn1560

Dickinson, A. (1985). Actions and habits: The development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *308*(1135), 67–78. https://doi.org/10.1098/rstb.1985.0010

Doll, B. B., Hutchison, K. E., & Frank, M. J. (2011). Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(16), 6188–6198. https://doi.org/10.1523/JNEUROSCI.6486-10.2011

Euston, D. R., Gruber, A. J., & McNaughton, B. L. (2012). The role of medial prefrontal cortex in memory and decision making. *Neuron*, *76*(6), 1057–1070. https://doi.org/10.1016/j.neuron.2012.12.002

Everitt, B. J., & Robbins, T. W. (2016). Drug addiction: Updating actions to habits to compulsions ten years on. *Annual Review of Psychology*, *67*, 23–50. https://doi.org/10.1146/annurev-psych-122414-033457

Fay, R. R. (1988). *Hearing in vertebrates: A psychophysics databook*. Hill-Fay Associates.

Froy, O. (2007). The relationship between nutrition and circadian rhythms in mammals. *Frontiers in Neuroendocrinology*, *28*(2–3), 61–71; https://doi.org/10.1016/j.yfrne.2007.03.001

Gallistel, C. R. (1990). The organization of learning. *The organization of learning*. The MIT Press.

Garrett, N., & Daw, N. D. (2020). Biased belief updating and suboptimal choice in foraging decisions. *Nature Communications*, *11*(1), Article 3417. https://doi.org/10.1038/s41467-020-16964-5

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. https://doi.org/10.1038/nn.2856

Huston, J. P., Silva, M. A. D. S., Topic, B., & Müller, C. P. (2013). What's conditioned in conditioned place preference? *Trends in Pharmacological Sciences*, *34*(3), 162–166. https://doi.org/10.1016/j.tips.2013.01.004

Hutcherson, C. A., Bushong, B., & Rangel, A. (2015). A neurocomputational model of altruistic choice and its implications. *Neuron*, *87*(2), 451–462. https://doi.org/10.1016/j.neuron.2015.06.031

Ilg, A. K., Enkel, T., Bartsch, D., & Bähner, F. (2018). Behavioral effects of acute systemic low-dose clozapine in wild-type rats: Implications for the use of DREADDs in behavioral neuroscience. *Frontiers in Behavioral Neuroscience*, *12*, Article 173. https://doi.org/10.3389/fnbeh.2018.00173

Janabi-Sharifi, F., Hayward, V., & Chen, C. S. J. (2000). Discrete-time adaptive windowing for velocity estimation. *IEEE Transactions on Control Systems Technology*, *8*(6), 1003–1009. https://doi.org/10.1109/87.880606

Kacelnik, A., & Bateson, M. (1996). Risky theories – The effects of variance on foraging decisions. *American Zoologist*, *36*(4), 402–434. https://doi.org/10.1093/icb/36.4.402

Kahneman, D., Knetsch, J., & Thaler, R. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias . *The Journal of Economic Perspectives*, *5*(1), 193–206. https://doi.org/10.1257/jep.5.1.193

Kahneman, D., Lovallo, D., & Sibony, O. (2011). The big idea: Before you make that big decision . . . *Harvard Business Review*, *89*(6), 50–60. https://hbr.org/2011/06/the-big-idea-before-you-make-that-big-decision

Kalenscher, T., & Pennartz, C. M. A. (2008). Is a bird in the hand worth two in the future? The neuroeconomics of intertemporal decision-making. *Progress in Neurobiology*, *84*(3), 284–315. https://doi.org/10.1016/j.pneurobio.2007.11.004

Karlsson, N., Gärling, T., & Bonini, N. (2005). Escalation of commitment with transparent future outcomes. *Experimental Psychology*, *52*(1), 67–73. https://doi.org/10.1027/1618-3169.52.1.67

MacLaren, D. A. A., Browne, R. W., Shaw, J. K., Radhakrishnan, S. K., Khare, P., España, R. A., & Clark, S. D. (2016). Clozapine N-oxide administration produces behavioral effects in Long-Evans rats: Implications for designing DREADD experiments. *eNeuro*, *3*(5), https://doi.org/10.1523/ENEURO.0219-16.2016

Magalhães, P., & White, K. G. (2016). The sunk cost effect across species: A review of persistence in a course of action due to prior investment. *Journal of the Experimental Analysis of Behavior*, *105*(3), 339–361. https://doi.org/10.1002/jeab.202

Mahler, S. V., Vazey, E. M., Beckley, J. T., Keistler, C. R., McGlinchey, E. M., Kaufling, J., Wilson, S. P., Deisseroth, K., Woodward, J. J., & Aston-Jones, G. (2014). Designer receptors show role for ventral pallidum input to ventral tegmental area in cocaine seeking. *Nature Neuroscience*, *17*(4), 577–585. https://doi.org/10.1038/nn.3664

Mazur, J. E. (1985). Probability and delay of reinforcement as factors in discrete-trial choice. *Journal of the Experimental Analysis of Behavior*, *43*(3), 341–351. https://doi.org/10.1901/jeab.1985.43-341

Muenzinger, K. F. (1938). Vicarious trial and error at a point of choice: I. A general survey of its relation to learning efficiency. *The Pedagogical Seminary and Journal of Genetic Psychology*, *53*(1), 75–86. https://doi.org/10.1080/08856559.1938.10533799

Muenzinger, K. F., & Gentry, E. (1931). Tone discrimination in white rats. *Journal of Comparative Psychology*, *12*(2), 195–206. https://doi.org/10.1037/h0072238

Nonacs, P. (2001). State dependent behavior and the marginal value theorem. *Behavioral Ecology*, *12*(1), 71–83. https://doi.org/10.1093/oxfordjournals.beheco.a000381

O'Doherty, J. P., Cockburn, J., & Pauli, W. M. (2017). Learning, reward, and decision making. *Annual Review of Psychology*, *68*, 475–484. https://doi.org/10.1146/annurev-psych-010416-044216

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Clarendon Press; http://hdl.handle.net/10150/620894

Papale, A. E., Stott, J. J., Powell, N. J., Regier, P. S., & Redish, A. D. (2012). Interactions between deliberation and delay-discounting in rats. *Cognitive, Affective & Behavioral Neuroscience*, *12*(3), 513–526. https://doi.org/10.3758/s13415-012-0097-7

Pattison, K. F., Zentall, T. R., & Watanabe, S. (2012). Sunk cost: Pigeons (*Columba livia*), too, show bias to complete a task rather than shift to another. *Journal of Comparative Psychology*, *126*(1), 1–9. https://doi.org/10.1037/a0023826

Raposo, D., Sheppard, J. P., Schrater, P. R., & Churchland, A. K. (2012). Multisensory decision-making in rats and humans. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(11), 3726–3735. https://doi.org/10.1523/JNEUROSCI.4998-11.2012

Redish, A. D. (1999). *Beyond the cognitive map: From place cells to episodic memory*. MIT Press.

Redish, A. D. (2013). *The mind within the brain: How we make decisions and how those decisions go wrong*. Oxford University Press.

Redish, A. D. (2016). Vicarious trial and error. *Nature Reviews Neuroscience*, *17*(3),147–159. https://doi.org/10.1038/nrn.2015.30

Rusu, S. I., & Pennartz, C. M. A. (2020). Learning, memory and consolidation mechanisms for behavioral control in hierarchically organized cortico-basal ganglia systems. *Hippocampus*, *30*(1), 342–359. https://doi.org/10.1002/hipo.23167

Schmidt, B., Duin, A. A., & Redish, A. D. (2019). Disrupting the medial prefrontal cortex alters hippocampal sequences during deliberative decision making. *Journal of Neurophysiology*, *121*(6), 1981–2000. https://doi.org/10.1152/jn.00793.2018

Schmidt, B., Papale, A., Redish, A. D., & Markus, E. J. (2013). Conflict between place and response navigation strategies: Effects on vicarious trial and error (VTE) behaviors. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *20*(3), 130–138. https://doi.org/10.1101/Lm.028753.112

Shafir, S., Wiegmann, D. D., Smith, B. H., & Real, L. A. (1999). Risk-sensitive foraging: Choice behaviour of honeybees in response to variability in volume of reward. *Animal Behaviour*, *57*(5), 1055–1061. https://doi.org/10.1006/anbe.1998.1078

Sleesman, D. J., Conlon, D. E., McNamara, G., & Miles, J. E. (2012). Cleaning Up the big muddy: A meta-analytic review of the determinants of escalation of commitment. *Academy of Management Journal*, *55*(3), 541–562. https://doi.org/10.5465/amj.2010.0696

Stachniak, T. J., Ghosh, A., & Sternson, S. M. (2014). Chemogenetic synaptic silencing of neural circuits localizes a hypothalamus→midbrain pathway for feeding behavior. *Neuron*, *82*(4), 797–808. https://doi.org/10.1016/j.neuron.2014.04.008

Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance*, *16*(1), 27–44. https://doi.org/10.1016/0030-5073(76)90005-2

Staw, B. M., & Ross, J. (1989). Understanding behavior in escalation situations. *Science*, *246*(4927), 216–220. https://doi.org/10.1126/science.246.4927.216

Steiner, A. P., & Redish, A. D. (2014). Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. *Nature Neuroscience*, *17*(7), 995–1002. https://doi.org/10.1038/nn.3740

Stephens, D. W. (2002). Discrimination, discounting and impulsivity: A role for an informational constraint. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *357*(1427), 1527–1537. https://doi.org/10.1098/rstb.2002.1062

Stephens, D. W. (2008). Decision ecology: Foraging and the ecology of animal decision making. *Cognitive, Affective and Behavioral Neuroscience*, *8*(4). https://doi.org/10.3758/CABN.8.4.475

Stephens, D. W., & Anderson, D. (2001). The adaptive value of preference for immediacy: When shortsighted rules have farsighted consequences. *Behavioral Ecology*, *12*(3), 330–339. https://doi.org/10.1093/beheco/12.3.330

Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton University Press.

Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, *23*(3), 165–180. https://doi.org/10.2307/2295722

Sweis, B. M., Abram, S. V., Schmidt, B. J., Seeland, K. D., MacDonald, A. W., III, Thomas, M. J., & Redish, A. D. (2018). Sensitivity to "sunk costs" in mice, rats, and humans. *Science*, *361*(6398), 178–181. https://doi.org/10.1126/science.aar8644

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, *1*(1), 39–60. https://doi.org/10.1016/0167-2681(80)90051-7

Tolman, E. C. (1939). Prediction of vicarious trial-and-error by means of the schematic sow-bug. *Psychological Review*, *46*(4), 318–336. https://doi.org/10.1037/h0057054

Trapanese, C., Meunier, H., & Masi, S. (2019). What, where and when: Spatial foraging decisions in primates. *Biological Reviews of the Cambridge Philosophical Society*, *94*(2), 483–502. https://doi.org/10.1111/brv.12462

van Der Meer, M., Kurth-Nelson, Z., & Redish, A. D. (2012). Information processing in decision-making systems. *Neuroscientist*, *18*(4), 342–359. https://doi.org/10.1177/1073858411435128

van Wingerden, M., Marx, C., & Kalenscher, T. (2015). Budget constraints affect male rats' choices between differently priced commodities. *PLOS ONE*, *10*(6), Article e0129581. https://doi.org/10.1371/journal.pone.0129581

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), 8308–8313. https://doi.org/10.1073/pnas.1220738110, see additional unpublished note: http://redishlab.neuroscience.umn.edu/papers/2013-PNAS-Commentary-WebLetter-2013-08-23.pdf

Zhu, H., & Roth, B. L. (2014). Silencing synapses with DREADDs. *Neuron*, *82*(4), 723–725. https://doi.org/10.1016/j.neuron.2014.05.002

Received October 27, 2020
Revision received January 24, 2021
Accepted January 28, 2021 ■